آمار و احتمال مهندسی
فصل اول آمار توصـــیفی

مقدمه
امروزه در علوم پایه و مهندسی به دفعات نـاز به جمع آوری اطلاعات در مورد مجموعه‌های از اشیاء و یا انسانها داریم که این امر بر عهده علم آمار می‌باشد و با کمک آمار توصیفی می‌توان اطلاعات جمع آوری شده را به صورتی منظم گرد آوری نمود بطوریکه بتوان با یک نگاه اجمالی به نتایج بدست آمده یک دید کلی نسبت به کل داده‌ها بدست آورد.
در این فصل به چگونگی جمع آوری، اطلاعات و تجزیه و تحلیل آن با کمک نمودارها و پارامترهای مرکزی و پراکندگی می‌پردازیم.

۱-۱ تعاریف اولیه
جامعه آماری : به مجموعه‌ای از اشیاء یا افراد که حداقل یک ویژگی مشترک آنها مورد مطالعه قرار می‌گـیرد. جامعه آماری می‌گوییم. ویژگی‌های مشترک یک جامعه آماری از عضوی به عضو دیگر تغییر می‌کند. که آنها را متغیر می‌نامند. متغیرها معمولاً به دو نوع تقسیم می‌شوند:
متغیر کمی : متغیرهایی می‌باشند که معمولاً قابل اندازه‌گیری هستند و می‌توان مقدار آنها را به صورت عددی نمایش داد مثل مقدار وزن، قد، حجم و ....
۱- متغیر کیفی : متغیرهایی می‌باشند که مستقیماً توسط اعداد و ارقام قابل اندازه‌گیری نیستند. مثل گروه خونی، شغل، رنگ چشم و ... که برای اندازه‌گیری این متغیرها به آنها عددی نسبت می‌دهیم.
متغیرهای کمی خود بر دو نوع هستند.
۱- گسسته : متغیرهایی که بین دو مقدار متصور آنها هیچ عدد دیگری وجود نداشته باشد.
۲- پیوسته : متغیرهایی که بین هر دو مقدار متصور آنها همواره عددی دیگری وجود دارد. مثل وزن یا طول و قد افراد.
پس از جمع آوری داده‌ها برای رسیدن به اهداف مورد نیاز به بررسی و تجزیه و تحلیل داده‌ها داریم که برای این منظور ابتدا داده‌ها را در یک جدول تنظیم و طبقه‌بندی می‌کنیم و سپس با استفاده از نمودارهای آماری نحوه توزیع داده‌ها را نمایش می‌دهیم و در نهایت داده‌ها را با کمک چند عدد به نام شاخص یا آماره خلاصه می‌کنیم.

۱-۲ جدول آماری
در جداول آماری علاوه بر داده‌ها، تعداد و درصد تکرار آنها نمایش داده می‌شود بطوریکه با یک نگاه به جدول می‌توان اطلاعات مفیدی در مورد پراکندگی و توزیع داده‌ها بدست آورد. یکی از متداول‌ترین جداول آماری جدول فراوانی است در جداول فراوانی همواره موارد زیر را خواهیم داشت:
۱- فراوانی: با شمردن تعداد دفعات تکرار هر داده در میان تمامی داده‌ها فراوانی آن داده بدست می‌آید. فرض کنید N داده داشته باشیم با قرار دادن داده‌های مشابه در یک دسته، در نهایت K طبقه خواهیم داشت $(\mathrm{K} \le \mathrm{N})$ که تعداد داده‌ها در هر دسته را فراوانی آن داده می‌نامیم. فراوانی طبقه i ام را با $\mathrm{f_i}$ نمایش می‌دهیم که $1 \le i \le k$ و $1 \le f_i \le N$ و $\sum_{i=1}^{k} f_i = N$ (N تعداد کل داده‌ها)

۲- فراوانی نسبی : عبارتست از حاصل تقسیم فراوانی هر طبقه بر تعداد کل داده‌ها که آنرا با $i = 1, \cdots k$ ، $r_i = \dfrac{f_i}{N}$ نمایش می‌دهیم.

$$\sum_{i=1}^{k} r_i = \sum_{i=1}^{k} \frac{f_i}{N} = \frac{1}{N} \sum_{i=1}^{k} f_i = 1$$

هم چنین اگر فراوانی نسبی را در عدد ۱۰۰ ضرب کنید درصد وقوع هر داده را در میان کل داده‌ها بدست می‌آورید.

۳- فراوانی تجمعی نسبی : حاصل جمع فراوانی هر طبقه با طبقات قبل از آنرا فراوانی تجمعی می‌نامیم و به $(1 \le i \le k)$ g_i نمایش می‌دهیم.

$$\text{فراوانی تجمعی طبقه iام} = g_i = f_1 + f_2 + \cdots + f_1 = \sum_{j=1}^{i} f_j$$

به همین ترتیب حاصل جمع فراوانی نسبی هر طبقه با طبقات قبل آنرا فراوانی تجمعی نسبی می‌نامیم و به $(1 \le i \le k)$ s_i نمایش می‌دهیم

$$\text{فراوانی تجمعی طبقه iام} = s_i = r_1 + r_2 + \cdots + r_1 = \sum_{j=1}^{i} s_j = s_i = \frac{1}{N} g_i$$

توجه : در محاسبه g_i و s_i اگر x_i نماینده طبقه iام باشد می‌بایستی داده‌ها به صورت مرتب و از کوچک به بزرگ در جدول قرار داده شوند بطوریکه داشته باشیم $x_1 < x_2 < \cdots x_n$.

۱-۲-۱ جدول فراوانی برای داده های گسسته

در مثال زیر چگونگی تشکیل جدول فراوانی برای داده‌های گسسته را بیان می‌کنیم.

مثال ۱ : داده‌های زیر تعداد فرزندان تحت پوشش بیمه در ۳۰ خانواده را نشان می‌دهد.

5 , 4 , 2 , 1 , 1 , 1 , 3 , 3 , 3 , 2 , 0 , 5 , 0 , 2 , 2 , 2 ,2 , 4 , 4 , 1 , 0 , 1 , 1 , 3 , 2 , 5 , 2 , 1 , 0 , 3

ابتدا داده‌ها را از کوچک به بزرگ برای تشکیل جداول فراوانی مرتب می‌کنیم:

0 , 0 , 0 ,0 , 1 , 1 , 1 , 1 , 1 , 1 , 1 , 2 , 2 , 2 , 2 , 2 ,2 , 2 , 2 , 3 , 3 , 3 , 3 , 4 , 4 , 4 , 4 , 5 , 5 , 5

در این صورت ۶ طبقه بدست خواهد آمد. نماینده هر طبقه در ستون x_i و فراوانی ان در ستون f_i نوشته می‌شود به این ترتیب جدول زیر را خواهیم داشت:

x_i	f_i	r_i	g_i	s_i
۰	۴	۰/۱۳	۴	۰/۱۳
۱	۷	۰/۲۳	۱۱	۰/۳۶
۲	۸	۰/۲۶	۱۹	۰/۶۲
۳	۴	۰/۱۳	۲۳	۰/۷۵
۴	۴	۰/۱۳	۲۷	۰/۸۸
۵	۳	۰/۱	۳۰	۱
جمع	۳۰	۱		

با توجه به جدول می‌توان به نتایج زیر رسید:

۱- با توجه به عدد ۰/۲۶ در ستون فراوانی نسبی می‌توان نتیجه گرفت که ۲۶ درصد از خانوارها دارای ۲ فرزند می‌باشند.

۲- با توجه به عدد ۰/۷۵ در ستون فراوانی تجمعی نسبی می‌توان نتیجه گرفت که ۷۵ درصد خانوارها حداکثر دارای ۳ فرزند می‌باشند.

۲-۲-۱ جدول فراوانی برای داده‌ها پیوسته

مثال ۲ : داده‌های زیر طول عمر ۴۰ عدد لامپ را نشان می‌دهند که به نزدیک‌ترین عدد صحیح گرد شده‌اند.

11	9	12	15	20	13	14	17	23	22
8	16	17	21	11	18	21	12	11	10
14	13	19	16	15	17	20	8	7	13
15	17	16	14	22	12	11	9	18	19

زمانی که با داده‌های پیوسته کار می‌کنیم داده‌ها را به رده‌های (فاصله‌ها) با طول مساوی تقسیم می‌کنیم و فراوانی داده‌ها را در هر رده بدست می‌آوریم. در این حالت نیاز به ثبت تمامی داده‌ها در جدول نمی‌باشد بلکه هر رده را به صورت مرتب از کوچک به بزرگ در جدول ثبت می‌کنیم.

در این حالت روند به این صورت است که :

۱- از آنجا که هر عدد به نزدیکترین عدد صحیح گرد شده است پس مثلاً عدد ۱۲ در واقع عددی بین (۱۲/۵ و ۱۱/۵) بوده است برای اینکه این مقدار نیز در عملیات وارد شود از مفهوم میزان تغیر پذیری داده‌ها که با s نمایش می‌دهیم استفاده می‌کنیم.

$$S=\frac{\text{واحد گرد شده داده‌ها}}{2}=\frac{1}{2}=0/5$$

۲- دامنه واقعی داده‌ها و کوچکترین و بزرگترین داده را به این ترتیب محاسبه می‌کنیم.

$$\min=\text{کوچکترین داده}-S=7-0/5=6/5$$

$$\max=\text{بزرگترین داده}+S=23+0/5=23/5$$

$$\text{تغیرات داده‌ها}\quad\text{دامنه}\quad R=\max-\min=23/5-6/5=11$$

۳- با توجه به دامنه داده‌ها می‌توان طول هر رده را معین نمود. برای این می‌بایستی دامنه را بر تعداد رده‌ها (که به صورت دلخواه قابل انتخاب است) تقسیم نمود.

$$\text{طول رده}=\omega=\frac{R}{K}\qquad K=\text{تعداد رده‌ها}$$

معمولاً تعداد رده‌ها طوری انتخاب می‌شوند که هر رده حداقل پنج داده را در برداشته باشد می‌توان از فرمول $K=1+3/322\log_{10}^{N}$ برای تعیین تعداد رده استفاده نمود بنابر این:

$$K=1+3/3222\log_{10}^{40}=6/322\cong 7\ \Rightarrow\ \omega=\frac{17}{7}=2/4\cong 3$$

۴- حالا کافیست ۷ رده به طول ۳ در ستون اول جدول فراوانی ایجاد کنیم و مجدداً مطابق مثال قبل جدول را کامل کنیم. اولین رده از کوچکترین عنصر آغاز می‌شود.

رده‌ها	خط و نشان	x_i	fi	r_i	g_i	s_i
$6/5-9/5$	𝍸	۸	۵	$0/125$	۵	$0/125$
$9/5-12/5$	𝍸 ‌‌|||	۱۱	۸	$0/2$	۱۳	$0/325$
$12/5-15/5$	𝍸 ||||	۱۴	۹	$0/225$	۲۲	$0/55$
$15/5-18/5$	𝍸 ||||	۱۷	۹	$0/225$	۳۱	$0/775$
$18/5-21/5$	𝍸 |	۲۰	۶	$0/15$	۳۷	$0/925$
$21/5-24/5$	|||	۲۳	۳	$0/075$	۴۰	$1/0$
جمع			۴۰			

توجه شماره ۱ :

- x_i معرف نماینده هر رده است که عضو میانی رده می‌باشد.

- چون آخرین رده شامل هیچ عضوی نبود آنرا حذف می‌کنیم مثلاً در این مثال رده ۲۴/۵-۲۷/۵ شامل هیچ عضوی نیست بنابر این حذف می‌شود.

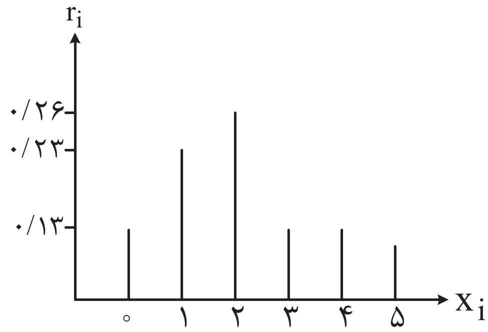
۱-۳ نمودارهای آماری

یکی دیگر از روشهای مناسب برای خلاصه نمودن داده‌های آماری استفاده از نمودار می‌باشد. نمودارها بهترین ابزار برای نمایش نحوه توزیع داده‌ها می‌باشند که در ذیل معروفترین آنها را معرفی می‌کنیم.

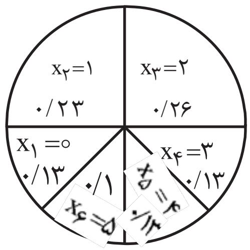
۱-۳-۱ نمودارهای آماری برای داده‌های گسسته

۱- نمودارهای میله‌ای:

در این نمودار محور x ها نمایش دهنده مقادیر داده‌ها و محور y ها نمایش دهنده فراوانی نسبی می‌باشد. چون فراوانی نسبی داده‌های هر جامعه آماری مقادیر بین صفر و یک را می‌گیرد بنابر این می‌توان چندین نمودار را با یکدیگر مقایسه نمود. شکل نمودار میله‌ای را برای مثال ۱ نمایش می‌دهد.



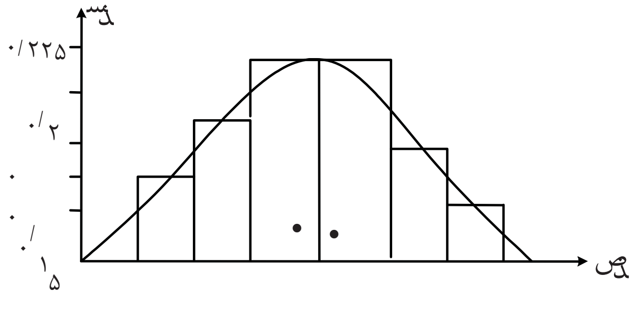
۳- نمودار دایره‌ای : این نمودار درون دایره‌ای رسم می‌شود که به قطاع‌هایی تقسیم شده است و هر قطاع متناسب با مساحت خود معرف یکی از فراوانی‌های نسبی در جدول فراوانی می‌باشد. شکل زیر نمایش دهنده نمودار دایره‌ای برای مثال ۱ می‌باشد.



۱-۳-۲ نمودارهای آماری برای داده‌های پیوسته

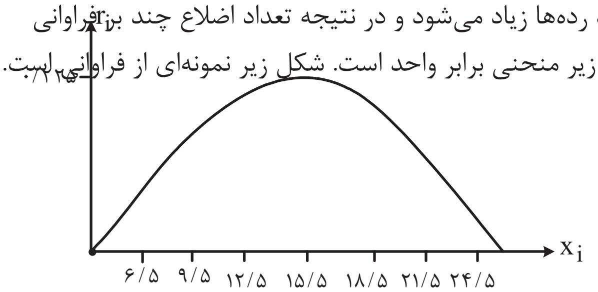
همانطور که در جدول فراوانی مشاهده کردید تفوت عمده رده‌های پیوسته با گسسته در بکارگیری رده‌ها می‌باشد این امر در نمودارها نیز کاملاً صادق است. در این قسمت سه نمودار را که برای نمایش داده‌های پیوسته بکار می‌روند معرفی می‌کنیم.

۱- هیستوگرام (نمودار ستونی) : این نمودار کاملاً مشابه نمودار میله‌ای می‌باشد با این تفاوت که محور xها نمایش دهنده رده‌های جدول فراوانی است و به ازای هر رده مستطیلی که عرض آن یک واحد و ارتفاع آن معادل فراوانی نسبی آن رده است رسم می‌شود. بنابر این مجموع مساحتهای مستطیل‌های رسم شده برابر واحد می‌باشد که همان مجموع کل فراوانی نسبی‌ها است. شکل زیر نمودار هستیوگرام برای مثال ۲ می‌باشد.



۲- چند برابر فراوانی: اگر در نمودار هستیوگرام وسط قاعده‌های بالای مستطیل‌ها را به یکدیگر توسط خطوطی به صورت متوالی متصل کنیم نمودار چند برابر فراوانی بدست می‌آید که در شکل بالا نیز نشان داده شده است. ابتدای خطوط به وسط رده ماقبل و انتهای خطوط به وسط رده مابعد مستطیلهای هستیوگرام متصل می‌شوند به این ترتیب مساحت زیر نمودار چند بر فراوانی مجدداً برابر واحد خواهد بود.

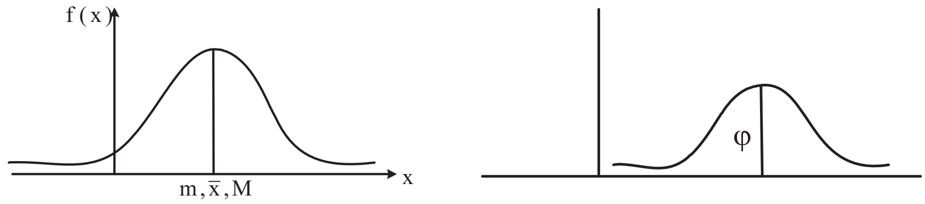
۳- منحنی فراوانی: در صورتی که تعداد داده‌ها زیاد باشد و طول رده‌ها کوچک، تعداد رده‌ها زیاد می‌شود و در نتیجه تعداد اضلاع چند بر فراوانی افزایش یافته و در نهایت مشابه یک منحنی خواهد شد که در این حالت نیز مساحت زیر منحنی برابر واحد است. شکل زیر نمونه‌ای از فراوانی است.



متغیر تصادفی نرمال: اگر مجموعه داده‌های X دارای میانگین $\overline{x}$ و واریانس σ^2 باشند و نمودار آنها از تابع

$$(-\infty < x < +\infty) \qquad f_{(x)} = \frac{1}{\sqrt{2\pi\sigma}}\, e^{\frac{-(x-\overline{x})^2}{2^2}}$$

تبعیت کند می‌گوییم متغیر تصادفی X نرمال می‌باشد. نمودار متغیر تصادفی نرمال نسبت به خط $x = \overline{x}$ متقارن می‌باشد. به شکل زیر توجه کنید.



همچنین با افزایش مقدار δ نمودار در راستای محور xها کشیده تر می‌شود.
با افزایش مقدار $\overline{x}$ نمودار به سمت راست و با کاهش آن به سمت چپ جابجا می‌شود.

۴-۱ پارامترهای مرکزی و پراکندگی

از آنجا که با مطالعه یک جامعه آماری تعداد زیادی داده بدست می‌آوریم و مطالعه روی تک تک یا قسمتی از این داده‌ها مشکل و حتی غیر ممکن است همواره علاقه داریم این داده‌ها را با کمک شاخص‌ها و پارامترهایی خلاصه کنیم تا بتوان با یک نگاه اجمالی به آن یک دید کلی نسبت به کل داده‌ها و جامعه آماری بدست آورد. بنابر این پارامترهای مرکزی و پراکندگی را معرفی می‌کنیم.

۱-۴-۱ پارامترهای مرکزی

عموماً داده‌ها در جامعه آماری یکنوع تجمع .و فشردگی حول یک مقدار خاص از صفت مورد مطالعه را بوجود می‌آورند که این مقدار خاص به عنوان یک پارامتر مرکزی معرفی می‌شود. مهمترین پارامترهای مرکزی عبارتند از: میانگین، میانه و مد (نما)

۱- میانگین: میانگین بر چند نوع است که معروفترین آنها میانگین حسابی، خندسی، همساز (هارمونیک) و درجه دوم می‌باشد.

الف) میانگین حسابی: اگر داده های $x_1, x_2, \ldots, x_k$ به ترتیب دارای فراوانی $f_1, f_2, \ldots, f_k$ باشند و تعداد کل این داده‌ها N باشد $\overline{X}$ یا میانگین حسابی به صورت زیر تعریف می‌شود:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{K} f_i\, x_i$$

برای داده‌های پیوسته x_i را نماینده هر رده در نظر می‌گیریم. به عنوان مثال میانگین حسابی برای مثال ۱ و ۲ عبارتست از:

مثال 1 $\quad : \overline{X} = \frac{1}{30}\,(4\times0+7\times1+8\times2+4\times3+4\times4+3\times5) = 2/2$

مثال 2 $\quad : \overline{X} = \frac{1}{40}\,(8\times5+11\times8+14\times9+17\times9+20\times6+23\times3) = 14/9$

ب) میانگین هندسی: در صورتی که همگی داده‌ها مثبت باشند میانگین هندسی به صورت زیر محاسبه می‌شود.

$$G=\left(\prod_{i=1}^{N} x_i^{f_i}\right)^{\frac{1}{N}}$$

ج: میانگین همساز: (همنوا یا هارمونیک) : اگر هیچکدام از داده ها صفر نباشد می‌توان میانگین همساز را از طریق فرمول محاسبه نمود.

$$H=\frac{N}{\sum_{i=1}^{n}\frac{f_i}{x_i}}$$

در حالت کلی رابطه $H \le G \le \overline{X}$ بین این سه میانگین برقرار است. که حالت تساوی زمانی رخ می دهد که همگی داده‌ها برابر باشند.

۲- میانه: با مرتب نمودن داده‌ها به صورت غیر نزولی عدد m را بعنوان میانه آنها در نظر می‌گیریم اگر تقراِیباً نیمی از داده‌ها سمت چپ آن و نیمی دیگر در سمت راست آن قرار داشته باشند.

همچنین با توجه به جدول فراوانی، میانه کوچکترین مقدار x است که فراوانی تجمعی آن بیشتر یا مساوی با $\frac{N}{۲}$ باشد.

- برای داده‌های گسسته در صورتی که داده‌ها را به صورت غیر نزولی مرتب کنیم اگر تعداد داده‌ها فرد باشد میانه داده وسطی یا به عبارتی $m=x_{(\frac{n+1}{2})}$ خواهد بود و اگر تعداد داده‌ها زوج باشد، میانه میانگین دو داده وسطی می‌باشد

$$m=\frac{x_{\left(\frac{n}{2}\right)}+x_{\left(\frac{n}{2}+1\right)}}{2} \qquad ; m=\frac{x_{\left(\frac{n}{2}\right)}+x_{\left(\frac{n}{2}+1\right)}}{2}=\frac{2+2}{2}=4$$

- برای داده‌های پیوسته ابتدا می‌بایستی رده میانه دار را پیدا نمود. رده میانه دار اولین رده‌ای است که فراوانی تجمعی نسبی آن از ۰/۵ بیشتر است. میانه عددیست که در این رده قرار دارد. حال برای محاسبه آن از فرمول زیر استفاده می‌کنیم:

$$m=L_{0/5}+\frac{(0/5N-g_{0/5})\omega}{f_{0/5}}$$

که در آن :

$L_{0/5}$: کران پایین رده میانه دار

N : تعداد داده ها

$g_{0/5}$: فراوانی تجمعی رده قبل از رده میانه دار

$f_{0/5}$: فراوانی رده میانه دار

ω : طول رده

به عنوان مثال میانه برای مثال ۱ با توجه به تعداد زوج داده‌ها عبارتست از:

برای مثال ۲ نیز میانه برابر است با:

$$m=12/5+\frac{(0/5\times 40-13)3}{9}=14/8$$

۳- مد یا نما: داده‌ای که فراوانی آن از سایر داده‌ها بیشتر باشد مد یا نما گفته می‌شود و با M نمایش داده می‌شود.

الف) روش محاسبه مد برای داده‌های گسسته: پس از بدست آوردن فراوانی داده، داده‌های که فراوانی آن از بقیه بیشتر باشد مد می‌باشد در این صورت سه حالت زیر پیش می‌آید: (که در مثال زیر به آن می‌پردازیم)

مثال ۳ : برای داده‌های مثال ۱ تنها عدد ۲ دارای بیشترین فراوانی می‌باشد. پس عدد ۲ مقدار مد خواهد بود. اما برای داده‌های ۵ و ۴ و ۴ و ۲ و ۲ و ۲ و۱ و ۱ و ۱ چون فراوانی ۱ و ۲ برابر است و این دو عدد به صورت متوالی قرار گرفته‌اند در این حالت مد برابر است با میانگین این دو عدد:

$$M=\frac{1+2}{2}=\frac{3}{2}$$

اما برای داده‌های ۴ و ۴ و۲ و ۲ و ۱ و ۱، فراوانی ۱ و ۴ مساوی می‌باشد ولی مجاور یکدیگر نیستند دو مقدار برای مد خواهیم داشت که عبارتند از:

$$M_1 = 1 \qquad M_2 = 4$$

توجه کنید در صورتی که فراوانی همه داده‌ها برابر باشد داده‌ها بدون مد در نظر گرفته می‌شوند.

ب) محاسبه مد برای داده‌های پیوسته: رده‌ای که فراوانی آن از سایر رده‌ها بیشتر است را به عنوان رده نمایی انتخاب می‌کنیم در این حالت می‌توان نماینده رده را به عنوان مد یا نما انتخاب کرد. اما برای محاسبه دقیق تر از فرمول زیر نیز می‌توان استفاده نمود:

$$M = L_M + \left(\frac{D_1}{D_1 + D_2}\right)\omega$$

L_M: کران پایین رده نمایی.

D_1: اختلاف فراوانی‌های نسبی رده نمایی و رده قبل از آن.

D_2: اختلاف فراوانی‌های نسبی رده نمایی و رده بعد از آن.

ω: طول رده.

در مثال ۲ داریم:

$$M_1 = 14 \quad ; \quad M_2 = 17$$

$$M = \frac{14+17}{2} = 15/5$$

و یا بصورت دقیق تر:

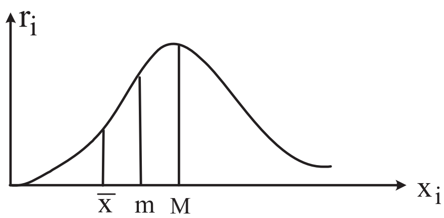
$$M_1 = 12/5 + \left(\frac{0/025}{0/025+0}\right) \times 3 = 15/5$$

$$M_2 = 15/5 + \left(\frac{0}{0+0/075}\right) \times 3 = 15/5$$

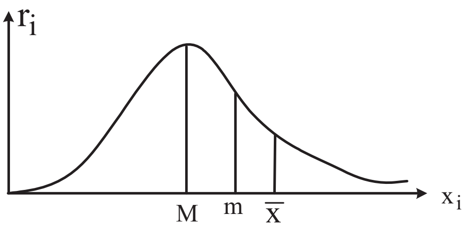
$$\Rightarrow M = 15/5$$

۵-۱ چولگی توزیع فراوانی

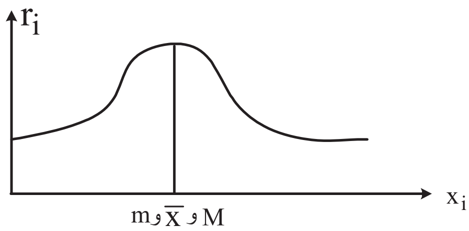
با رسم نمودار فراوانی داده‌های پیوسته عموماً یکی از اشکال زیر بدست می‌آید. که چگونگی قرار گرفتن میانگین و میانه و مد را در اشکال زیر مشاهده می‌شود.



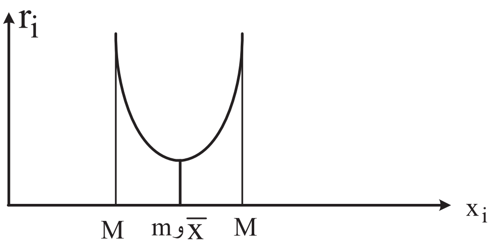
چوله به‌چپ $\bar{X} < m < M$



چوله به‌راست $M < m < \bar{X}$



متقارن $M = \bar{X} = m$



متقارن $m = \bar{X} = M$

برای محاسبه میزان عدم تقارن منحنی، شاخصی به نام B_1 (ضریب چولگی) وجود دارد که در فصل‌های بعدی با آن آشنا می‌شوید.

در توزیع‌هایی که چولگی زیاد نباشد رابطه تجربی زیر که به رابطه پیرسن معروف است برقرار می‌باشد.

$$\overline{X}-M \cong 3(\overline{X}-m)$$

مقدار عدم تقارن منحنی که با ضریب چاولگی سنجیده می‌شود را با نمودار نرمال مقایسه می‌کنیم برای بدست آوردن ضریب چاولگی از مفهوم گشتاور مرتبه kام طول میانگین (گشتاور مرکزی مرتبه kام) استفاده می‌کنیم. که عبارتست از :

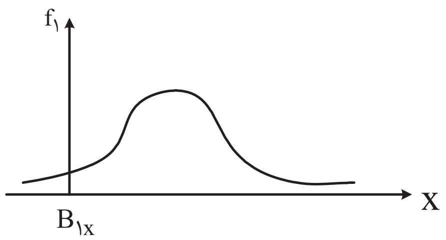
$$\mu_k=\frac{1}{N}\sum_{i=1}^{n} f_i\ (x_i-\overline{x})^k$$

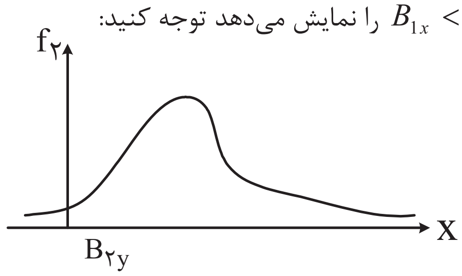
ضریب چاولگی که با B_1 نمایش داده می‌شود میزان عدم تقارن منحنی را نسبت به منحنی نرمال نمایش می‌دهد که به شکل زیر تعریف می‌شود:

$$B_1=\frac{\mu_3}{\sqrt{\mu_2^3}}=\frac{\mu_3}{\delta^3}$$

در صورتی که منحنی متقارن باشد مقدار B_1 برابر صفر می‌باشد و هر چه B_1 مقدار بزرگتری داشته باشد نشان دهنده افزایش عدم تقارن می‌باشد.

به شکل زیر که دو مقدار $B_{1x}<B_{1y}$ را نمایش می‌دهد توجه کنید:



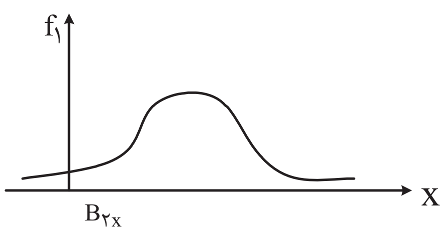


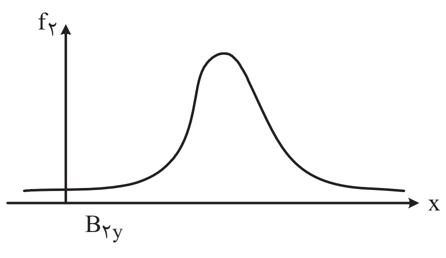
ضریب کشیدگی : نشان دهنده میزان کشیدگی منحنی می‌باشد و به صورت زیر تعریف می‌شود.

$$B_2=\frac{\mu_4}{\mu_2^2}=\frac{\mu_4}{\delta^4}$$

دو منحنی زیر متقارن می‌باشند اما ضریب کشیدگی در آنها متفاوت است.

$B_{2x}<B_{2y}$





۶-۱ پارامترهای پراکندگی

شاخص های مرکزی تنها یک منطقه را به عنوان محل تمرکز داده ها معرفی می کنند حال آنکه ممکن است دو دسته داده با پراکندگی های متفاوت دارای میانگین برابری باشند به عبارتی نیاز به شاخصی برای نمایش میزان پراکندگی داده ها خواهیم داشت که در ادامه به معرفی این پارامترها می پردازیم.

۱- دامنه داده‌ها: حاصل تفاضل بزرگترین داده از کوچکترین داده را R یا دامنه داده‌ها گویند.

$$R = X_N - X_۱$$

که در آن داده‌ها به فرم $X_{(1)} < X_{(2)} < \cdots < X_N$ با فرض اینکه مرتب شده‌اند.

این شاخص معیار خوبی برای محاسبه میزان پراکندگی داده‌ها نیست. زیرا در محاسبه تنها کوچکترین و بزرگترین داده وارد می‌شود.

مثال ۴ : نمرات ۱۰ دانش آموز در دو کلاس متفاوت در درس ریاضی به قرار زیر است:

کلاس A :۰ , 4 , 4 , 12 , 12 , 12 , 14 , 16 , 16 , 20

کلاس B :8 , 8 , 9 , 9 , 12 , 12 , 12 , 13 , 13 , 14

با محاسبه مقادیر $\overline{X}$ و m و M دیده می‌شود که برای هر دو کلاس داریم:

$$\overline{X} = ۱۱ \quad , \quad M = ۱۲ \quad , \quad m = ۱۲$$

اما با توجه به مقادیر تک تک نمرات واضح است که میزان پراکندگی نمرات در دو کلاس کاملاً متفاوت است و برای این منظور نیاز به شاخص‌های پراکندگی می‌باشد تا این مطلب را بتوان با مقایسه آنها نشان داد.

برای دو کلاس مقدار دامنه را محاسبه می‌کنیم:

کلاس A : $R = ۲۰ - ۰ = ۲۰$

کلاس B : $R = ۱۴ - ۸ = ۶$

۲- میانگین انحرافات: فاصله داده x_i از میانگین را انحراف از میانگین داده x_i گویند که به صورت $|x_i - \overline{x}|$ محاسبه می شود. اگر این مقدار را برای تمامی داده‌ها محاسبه کنیم و از نتیجه میانگین بگیریم میانگین انحرافات بدست خواهد آمد که عبارتست از : $D = \frac{۱}{N}\sum_{i=۱}^{K} f_i\, |x_i - \overline{x}|$.

از آنجا که میانگین انحرافات به تمام داده‌ها وابسته است معیار مناسبی برای سنجش پراکندگی داده‌ها محسوب می‌شود اما بدلیل وجود قدر مطلق در فرمول، محاسبه آن مشکل است و نمی‌توان آنرا ساده نمود بنابر این از واریانس و انحراف استاندارد استفاده می‌کنیم.

۳- واریانس و انحراف استاندارد: میانگین مجذور انحرافات را واریانس می‌نامیم و با نماد ∂ یا $S_b^۲$ نمایش می‌دهیم. که عبارتست از:

$$S_b^۲ = \frac{۱}{N}\sum_{i=۱}^{K} f_i\,(x_i - \overline{x})^۲$$

در مبحث استنباط آماری واریانس را از مجموع مجذور انحرافات داده ها تقسیم بر $N-۱$ بدست می‌آورند و آنرا با $S^۲$ نمایش می‌دهند.

$$S^۲ = \frac{۱}{N-۱}\sum_{i=۱}^{K} f_i\,(x_i - \overline{x})^۲$$

در مباحث این درس هر جا صحبت از واریانس می‌کنیم منظور $S^۲$ می‌باشد.

اگر از واریانس جذر بگیریم یعنی $S = \sqrt{s^۲}$ در این صورت S را انحراف استاندارد می‌نامیم که مغیار مناسبی برای سنجش پراکندگی می‌باشد.

همچنین $S^۲$ را می‌توان از فرمول زیر محاسبه نمود:

$$S^۲ = \frac{۱}{N-۱}\left[\sum_{i=۱}^{K} f_i\,x_i^۲ - \frac{۱}{n}\left(\sum_{i=۱}^{K} f_i\,x_i\right)^۲\right] = \frac{۱}{n-۱}\left[\sum_{i=۱}^{K} f_i\,x_i^۲ - n\,\overline{x}^۲\right]$$

اثبات:

$$S^{۲} = \frac{۱}{N-۱}\sum_{i=۱}^{N} f_i\,(x_i-\bar{x})^{۲} = \frac{۱}{N-۱}\sum_{i=۱}^{n} f_i\,(x_i^{۲}-۲\bar{x}_i\,\bar{x}+\bar{x})$$

$$=\frac{۱}{N-۱}\,(\sum_{i=۱}^{n} f_i\,x_i^{۲})-۲n\,\bar{x}^{۲}+n\,\bar{x}^{۲} = \frac{۱}{N-۱}\,(\sum_{i=۱}^{n} f_i\,x_i^{۲})-n\,\bar{x}^{۲}$$

مثال ۵: مقدار واریانس را برای مثال ۲ محاسبه کنید:

$$\bar{X} = ۱۴/۹ \qquad N=۴۰$$

$$S^{۲} = \frac{۱}{۴۰-۱}\left[(۵۰\times ۶۴ + ۸\times ۱۲۱ + ۹\times ۱۹۶ + ۹\times (۱۷)^{۲} + ۶\times ۴۰ + ۳\times (۲۳)^{۲} - ۴۰\times ۱۴۹)^{۲}\right]$$

۴- ضریب تغییرات : در محاسبه میزان پراکندگی داده‌ها همواره با داده‌هایی سروکار داریم که با مقیاس‌های مختلفی اندازه‌گیری شده‌اند بنابر این برای مقایسه میزان پراکندگی داده‌های بدست آمده از دو جامعه آماری که با مقیاس‌های مختلفی اندازه‌گیری شده‌اند استفاده از واریانس مناسب نمی‌باشد زیرا واریانس به مقیاس اندازه‌گیری وابسته می‌باشد بنابر این از مقیاس مناسبتری به نام ضریب تغییرات استفاده می‌کنیم که از رابطه $CV = \frac{S}{\bar{X}}$ بدست می‌آید و معمولاً با ضریب آن در عدد صد بر حسب درصد بیان می‌شود.

مثال ۶: یک کارخانه تولید لاستیک دو نوع محصول A و B تولید می‌کند. لاستیک نوع A دارای میانگین طول عمر ۲۰۰۰۰ کیلو متر و انحراف استاندارد ۲۰۰۰ کیلو متر می‌باشد و لاستیک نوع B دارای میانگین طول عمر ۱۸۰۰۰ کیلو متر و انحراف استاندارد ۲۰۰ کیلو متر می‌باشد، کدام نوع لاستیک برای خرید مناسب‌تر می‌باشد؟

$$X_A = ۲۰۰۰۰ \qquad \bar{X}_B = ۱۸۰۰۰ \qquad \Rightarrow \qquad CV_A = \frac{۲۰۰۰}{۲۰۰۰۰} = ۰/۱$$

$$S_A = ۲۰۰۰ \qquad S_B = ۲۰۰ \qquad \Rightarrow \qquad CV_B = \frac{۲۰۰}{۱۸۰۰۰} = ۰/۰۱$$

$$CV_A = ۰/۱\times ۱۰۰ = \%۱۰$$
$$CV_A = ۰/۰۱\times ۱۰۰ = \%۱$$

همانطور که ملاحظه می‌کنید میانگین طول عمر لاستیک دوم از لاستیک اول کمتر است ولی با توجه به اینکه ضریب تغییرات لاستیک دوم کمتر از لاستیک اول است خرید لاستیک دوم به صرفه‌تر می‌باشد.

۱-۷ تغییر مقیاس و مبدأ

داده‌ها را با واحدهای متفاوتی می توان از جامعه آماری جمع آوری نمود به عنوان مثال فرض کنید داده‌های مربوط به وزن ۴۰ نفر از دانشجویان یک کلاس را با واحد کیلوگرم جمع آوری کرده باشید و بخواهید مقادیر میانگین و واریانس را بر حسب پاوند بدست بیاورید برای این منظور نیازی به محاسبه مجدد میانگین و واریانس نمی‌باشد بلکه کافیست از روش تغییر مقیاس و مبدأ استفاده کنید.

۱-۷-۱ تغییر مقیاس

اگر تمامی داده‌ها در عدد a ضرب شوند در این صورت داریم:

$$x_۱, x_۲, \ldots, x_n \quad\rightarrow\quad ax_۱, ax_۲, \ldots, ax_n$$

$$\Rightarrow \bar{X} = \frac{۱}{N}\sum_{i=۱}^{n} x_i \quad\rightarrow\quad \bar{X}_a = \frac{۱}{N}\sum_{i=۱}^{n} ax_i = a(\frac{۱}{N}\sum_{i=۱}^{n} x_i) = a\bar{X} \quad\Rightarrow\quad \bar{X}_a = a\bar{X}$$

به همین ترتیب برای محاسبه واریانس بدست می‌آید:

تمرین ۱ :

$$S_a^{۲} = a^{۲} S^{۲} \quad \rightarrow \quad S_a = |a|\, S$$

$$S_a = \frac{۱}{n-۱} \sum_{۱}^{n} (a x_i - a \bar{x})^{۲} = a^{۲} \frac{۱}{n-۱} \sum_{۱}^{n} (x_i - \bar{x})^{۲} = a^{۲} S^{۲}$$

$$\Rightarrow S_a = \sqrt{a^{۲} S^{۲}} = |a|\, S$$

تمرین ۲ :

$$X \rightarrow X + b$$

$$\bar{X}_b = \bar{X} + b$$

$$\bar{X}_b = \frac{۱}{n} \sum_{۱}^{n} (x_i + b) = \frac{۱}{n} \left[\sum_{۱}^{n} x_i + \sum_{۱}^{n} b \right] = \frac{۱}{n} \sum_{۱}^{n} x_i + \frac{b}{n} \sum_{۱}^{n} (۱) = \bar{X} + b \frac{n}{n} = \bar{X} + b$$

$$S_b^{۲} = S^{۲}$$

$$S_b^{۲} = \frac{۱}{n-۱} \sum_{۱}^{n} (x_i + \cancel{b} - \bar{X} - \cancel{b})^{۲} = \frac{۱}{n-۱} \sum_{۱}^{n} (x_i - \bar{x})^{۲} = S^{۲}$$

تمرین ۳ –

$$X \rightarrow Y = aX + b$$

$$\bar{Y} = a\bar{X} + b \quad , \quad S_Y^{۲} = a^{۲} S^{۲}$$

اثبات : $\bar{Y} = \frac{۱}{n} \sum_{۱}^{n} (a x_i + b) = \frac{a}{n} \sum_{۱}^{n} x_i + \frac{bn}{n} = a\bar{X} + b$

$$S_Y^{۲} = \frac{۱}{n-۱} \sum_{۱}^{n} (a x_i + \cancel{b} - a\bar{x} - \cancel{b})^{۲} = \frac{a^{۲}}{n-۱} \sum_{۱}^{n} (x_i - \bar{x})^{۲} = a^{۲} S^{۲}$$

استاندارد سازی

مثال : نمره علی از امتحان فیزیک و ریاضی به ترتیب برابر ۴۰ و ۶۰ شده است اگر میانگین نمرات امتحان فیزیک و ریاضی به ترتیب برابر ۲۰ و ۵۰ باشد و انحراف معیار امتحان فیزیک و ریاضی به ترتیب برابر ۱ و ۲ باشد علی کدام درس را بهتر امتحان داده است.

حل: برای اینکه بتوان نمرات دو درس را با یکدیگر مقایسه نمود می‌بایستی ابتدا نمرات را استاندارد سازی نمود و سپس آنها را با یکدیگر مقایسه نمود.

نمرات استاندارد شده علی در درس ریاضی $= \frac{۶۰ - ۵۰}{۲} = ۵$

نمرات استاندارد شده علی در درس فیزیک $= \frac{۴۰ - ۲۰}{۱} = ۲۰$

با وجود اینکه نمره علی در درس فیزیک کمتر از ریاضی می‌باشد اما با استاندارد نمودن نمره دو درس مشاهده می‌کنیم که نمره وی در درس فیزیک بالاتر از درس ریاضی می‌باشد به عبارتی علی درس فیزیک را بهتر از درس ریاضی امتحان داده است.

۱-۷-۲ تغییر مبدأ

در صورتی که به تمام داده‌ها مقدار b را اضافه یا کم کنیم می‌توان نشان داد که مقادیر $\overline{X}$ و $S^{۲}$ جدید از روابط زیر محاسبه می‌شوند:

$$\overline{X}_b = \overline{X} + b$$

$$S_b^{۲} = S^{۲}$$

تغییر مبدأ روی واریانس بی تأثیر است

با اعمال همزمان تغییر مبدأ و مقیاس خواهیم داشت:

$$\overline{Y} = a\,\overline{X} + b$$

$$S_a^{۲} = a^{۲}\,S^{۲}$$

مطالب فوق برای میانه و مد نیز صادق می‌باشند و داریم:

$$m^{۱} = am + b$$

$$M^{۱} = a\,m + b$$

۱-۱۳ استاندارد سازی

از یک جامعه آماری n نمونه $x_{۱}, x_{۲}, \ldots x_n$ بصورت تصادفی انتخاب می‌کنیم بطوریکه میانگین و واریانس نمونه‌ها بترتیب $\overline{x}$ و $S_x^{۲}$ می‌باشد. با توجه به تغییر مبدأ و مقیاس مقدار هر نمونه را از میانگین نمونه‌ها کم می‌کنیم و حاصل را بر S_x تقسیم می‌کنیم بنابر این داده‌های $y_n = \dfrac{x_n - \overline{x}}{S_x}$, $y_{۲} = \dfrac{x_{۲} - \overline{x}}{S_x}$, $y_{۱} = \dfrac{x_{۱} - \overline{x}}{S_x}$ را خواهیم داشت.

با محاسبه مقایر میانگین و واریانس داده‌های جدید داریم: $S_y^{۲} = (\dfrac{۱}{S_x^{۲}})\ S_x^{۲} = ۱$; $\overline{y} = \dfrac{\overline{x} - \overline{x}}{S_x} = ۰$

همانطور که ملاحظه می‌کنید داده‌های جدید دارای میانگین صفر و واریانس ۱ می‌باشند که به آنها داده‌های استاندارد شده می‌گوییم. همینطور اگر داده‌های $x_{۱}$ تا x_n را با متغییر تصادفی X نمایش دهیم در این صورت $Y = \dfrac{X - \overline{x}}{S_x}$ فرم استاندارد شده یا صورت معیاری متغیر تصادفی X می‌باشد.

مسائل فصل اول :

۱- دو جامعه با اندازه‌های میانگین $\overline{X}_{۱}, \overline{X}_{۲}$ و انحراف معیار $S_{۱}, S_{۲}$ را با یکدیگر ادغام می‌کنیم ثابت کنید میانگین و انحراف معیار جدید از روابط زیر بدست می‌آید:

$$\overline{X} = \frac{N_{۱}\,\overline{X}_{۱} + N_{۲}\,\overline{X}_{۲}}{N_{۱} + N_{۲}}$$

$$S^{۲} = \frac{N_{۱}\,S_{۱}^{۲} + N_{۲}\,S_{۲}^{۲}}{N_{۱} + N_{۲}} + \frac{N_{۱}\,N_{۲}}{(N_{۱} + N_{۲})^{۲}} \quad (\overline{X}_{۱} - \overline{X}_{۲})$$

۲- میانگین و واریانس ۲ داده به ترتیب ۱۵ و ۵ می‌باشد. اگر به جای عدد ۲۵ اشتباهاً عدد ۱۵ را در محاسبات اعمال کرده باشیم میانگین و واریانس جدید را بدست بیاورید.

۳- نشان دهید تغیر مقیاس داده بر روی مقدار ضریب تغیرات $CV = \dfrac{S}{X}$ بی‌اثر می‌باشد، آیا این مطلب در مورد تغیر مبدأ نیز صادق است؟

۴- ثابت کنید برای میانگین حسابی، هندسی و همساز رابطه زیر برقرار است.

$$\overline{X}_H \le \overline{X}_G \le \overline{X}$$

۵- اگر میانگین را از داده‌های یک خامعه آماری کم کنیم و نتیجه را بر انحراف معیار تقسیم کنیم (یعنی $y_i = \frac{x_i - \overline{X}}{S}$) نشان دهید میانگین و انحراف معیار جدید به ترتیب صفر و یک می‌باشد.

۶- برای بدست آوردن یک معیار پراکندگی جدید داده‌ها را دو به دو با یکدیگر مقایسه می‌کنیم و میانگین $n^۲$ داده جدید $(X_i - X_j)$ را با $S_{ij}^۲$ نمایش می‌دهیم:

$$S_{ij}^۲ = \frac{۱}{n^۲} \sum_{i=۱}^{n} \sum_{j=۱}^{n} (X_i - X_j)^۲$$

نشان دهید $S_{ij}^۲ = ۲S^۲$ (راهنمایی: داخل پرانتز مقدار $\overline{X}$ را اضافه و کم کنید)

۷- نشان دهید میانگین حسابی و واریانس نخستین n عدد طبیعی به ترتیب $\frac{n+۱}{۲}$ و $\frac{n^۲-۱}{۱۲}$ می‌باشد.

۸- جدول زیر را برای داده‌ها و فراوانی آنها در نظر بگیرید:

x	۰	۱	۲	…	n
f	$\binom{n}{۰}$	$\binom{n}{۱}$	$\binom{n}{۲}$	…	$\binom{n}{n}$

نشان دهید میانگین و واریانس این داده‌ها به ترتیب عبارتست از:

$$\overline{X} = \frac{n}{۲} \quad , \quad S^۲ = \frac{n}{۴}$$

۹- اگر متحرکی مسافت $x_۱$ را با سرعت $v_۱$ و … و مسافت x_n را با سرعت v_n طی کنید ثابت کنید سرعت متوسط این متحرک با استفاده از رابطه میانگین همساز یا هارمونیک بدست می‌آید که در آن v_i معادل مقادیر داده‌ها و x_i معادل فراوانی آنهاست.

۱۰- عدد QP که $۰ < P < ۱$ را چندک Pام داده‌ها تعریف می‌کنیم هر گاه فراوانی تجمعی نسبی (r_i) آن بزرگتر یا مساوی با عدد P باشد. به عبارت دیگر هر گاه XP ۱۰۰٪ از داده‌ها قبل از آن قرار گیرند. به عنوان مثال $Q_{۰/۵}$ که به آن چارک دوم می‌گوییم همان میانه می‌باشد چرا که ۵۰٪ داده‌ها قبل از آن قرار دارند. در حالت کلی $Q_{۰/۲۵}$ و $Q_{۰/۵}$ و $Q_{۰/۷۵}$ را به ترتیب با $Q_۱, Q_۲, Q_۳$ نمایش می‌دهیم و به آنها چارکهای اول و دوم و سوم می‌گوییم.

الف) اگر برای داده‌های گسسته x_i داشته باشیم $x_۱ < x_۲ < \cdots < x_i < \cdots x_N$ انگاه نشان دهید $Q_P = (۱-\omega)\, x_r + \omega x_{r+۱}$ که در آن $r = (n+۱)P \quad , \quad \omega = (n+۱)P - r$.

ب) برای داده‌های پیوسته رده‌ای که فراوانی تجمعی آن بزرگتر یا مساوی با عددP باشد را رده QP می‌نامیم
نشان دهید چندک Pام برای داده‌های پیوسته از رابطه زیر بدست می‌آید:

$$Q_P = L_p + \frac{(np - g_p)}{f_p}$$

که در آن :

L_P : کران پایین رده QP

g_P : فراوانی تجمعی رده قبل از رده QP

f_P : فراوانی رده QP

ω : طول رده QP

۱۱- جدول زیر تعداد کتب فروخته شده توسط کتابفروشی را در طول ۳۰ روز نمای‌ش می‌دهد

۱۵	۱۰	۷	۲۰	۱۱	۱۳	۱۸	۶	۵	۴
۱۱	۱۹	۱۲	۱۶	۹	۱۰	۲۱	۱۳	۸	۱۴
۱۷	۲۰	۱۰	۱۲	۱۶	۱۳	۱۱	۱۲	۷	۱۱

برای داده‌های فوق :

الف: یک جدول فراوانی تشکیل دهید و نمودار میله‌ای داده‌ها را رسم کنید.

ب: میانگین داده‌ها $\overline{X}$ ، مد M و میانه m را بدست آورید.

ج: دامنه داده‌ها R، میانگین انحرافات، واریانس $S^{۲}$ و ضریب تغیر را بدست آورید.

د: اگر به بزرگترین داده مقدار x $[x \geq ۰]$ واحد اضافه کنیم کدامیک از مقادیر مد یا میانه بدون تغیر باقی می‌مانند.

ه: چارک اول و سوم را بدست آورید و دامنه چارکها را محاسبه کنید .(دامنه چارکها : $Q_{۳} - Q_{۱}$)

و: مقدار دهک چهارم را محاسبه کنید . (دهک چهارم همان $Q_{۰/۴}$ می‌باشد)

۱۲- در یک شهر میزان درجه حرارت در طول ۳۰ روز به قرار زیر است:

۵	۷	۸/۳	۱۰	۱۱	۱۲/۵	۱۳/۸	۱۳	۱۲	۱۳/۱
۱۴	۱۵/۲	۱۵/۶	۱۶	۱۵/۴	۱۵	۱۶/۵	۱۷	۱۹	۱۹/۷
۲۰/۶	۲۱	۲۱/۳	۲۰/۵	۲۲	۲۲/۸	۲۱/۷	۲۳	۲۴/۱	۲۵

الف: برای داده‌های فوق یک جدول فراوانی تشکیل دهید و هستیوگرام و چند بر فراوانی را رسم کنید

ب: مقادیر میانگین، میانه و مد را محاسبه کنید.

ج: مقادیر واریانس و ضریب تغیر را محاسبه کنید.

د: چند درصد داده‌ها در فاصله $(\overline{X} - S \,,\, \overline{X} + S)$ و چند درصد داده‌ها در فاصله $(\overline{X} - ۲S \,,\, \overline{X} + ۲S)$ قرار دارند؟

ه: چند درصد داده‌ها بین چارک اول و سوم قرار دارند؟

فصل دوم
آنالیز ترکیبی و احتمال

در این فصل ابتدا با روشها و قواعد شمارش آشنا می‌شوید. این قوانین به شما کمک می‌کنند که تعداد دفعات قوع یک حالت خاص از بین تمتمی حالات را به دست بیاورید. در ادامه با استفاده از نتایج بدت آمده و بکارگیری قوانین احتمالات، می‌توانید احتمال وقع یک حالت خاص را بدست آورید.

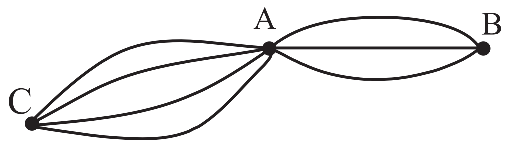
۲-۱ آنایز ترکیبی

یک سکه را در نظر بگیرید با پرتاب این سکه دو حالت ممکن است رخ بدهد. شیر یا خط حال اگر بخواهیم بدانیم با پرتاب ۳ عدد سکه چند حالت رخ می‌دهد می‌بایستی از قوانین شمارش استفاده کنیم.

به طور کلی تمامی قوانین شمارش مبتنی بر دو اصل ضرب و جمع می‌باشند. که با استفاده از این دو اصل می‌توان تمامی مسایل شمارش را به سادگی حل نمود.

اصل جمع: اگر کاری را بتوان به $n_۱$ طریق و کار دیگری را به $n_۲$ طریق انجام داد و این دو کار را نتوان همزمان انجام داد، آنگاه کار اول یا کار دوم را می‌توان به $n_۱ + n_۲$ طریق انجام داد.

مثال ۱: فرض کنیم از شهر A بتوان به سه طریق به شهر B و به چهار طریق به شهر C سفر نمود. به چند طریق می‌توان از شهر A به شهر B یا شهر C سفر نمود؟

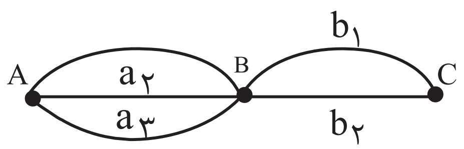


حل: کار اول را سفر از شهر A به B در نظر بگیرید که به $n_۱ = ۳$ طریق قابل انجام است همینطور کار دوم را سفر از شهر A به C در نظر بگیرید که به $n_۲ = ۴$ طریق قابل انجام است به وضوح هر دو کار نیز همزمان قابل انجام نیستند پس به $n = n_۱ + n_۲ = ۷$ طریق می‌ت‌وان از شهر A به B یا C سفر نمود.

توجه: اصل جمع را می‌توان به n کار نیز تعمیم داد. به شرطی که هیچ دو جفت کاری را نتوان همزمان انجام داد.

اصل ضرب: اگر کاری را بتوان به $n_۱$ طریق و کار دیگری را به $n_۲$ طریق انجام داد واین دو کار را بتوان بصورت همزمان و یکی پس از دیگری انجام داد، آنگاه هر دو کار را می‌توان به $n_۱ \times n_۲$ طریق انجام داد.

مثال ۱: فرض کنید از شهر A به سه طریق بتوان به شهر B سفر نمود. و از شهر B نیز به دو طریق به شهر C سفر نمود. به چند طریق می‌توان از شهر C سفر نمود؟



با توجه به شکل می‌بایستی ابتدا از شهر A به B و سپس به C سفر نمود که این دو عمل می‌بایستی ابتدا از شهر A به B و سپس به c سفر نمود مه این دو عمل می‌بایسی یکی پس از دیگری انجام شوند تا کل کار (سفر از شهر A بع C) انجام پذیرد. بنابر این کار اول سفر از شهر A به B که به $n_۱ = ۳$ طریق قابل انجام است و کار دوم سفر از شهر B به C $n_۲ = ۲$ طریث قابل انجام است را خواهیم داشت که کل کار را می‌توان به $n = n_۱ \times n_۲ = ۳ \times ۲ = ۶$ طریق می‌وان انجام داد.

اگر راهها از شهر A به B را با a_i و از شهر B به C را با b_i نامگذاری کنیم این ۶ طریق را می‌توان به صورت زیر لیست نمود:

۱) $a_۱ b_۱$	۳) $a_۲ b_۱$	۵) $a_۳ b_۱$
۲) $a_۱ b_۲$	۴) $a_۲ b_۲$	۶) $a_۳ b_۲$

توجه: اصل ضرب را می‌وان به n کار نیز تعمیم داد به شرطی که تمتم کارها را یکی پس از دیگری و همزمان بتوان انجام داد.

مثال ۳: پرتاب ۳ سکه بصورت همزمان چند حالت ممکن را در پی دارد؟
پرتاب یک سکه به دو حالت ممکن شیر و خط امکان‌پذیر است حال سه کار را در نظر بگیرید که عبارتند از پرتاب سکه که هر کدام به دو طریق قابل انجام هستند و بناست این سه کار همزمان انجام شوند بنابراین کل حالات ممکن عبارتست از:

$$n=n_۱\times n_۲\times n_۳=۲\times۲\times۲=۸$$

این حالات را می‌توان در یک جامعه لیست نمود که به آن فضای نمونه گویند.

خط : خ　　　　شیر : ش

$S=\{$(ش و خ و ش) و (خ و ش و خ) و (ش و خ و خ) و (ش و ش و خ) و(خ و ش و ش) و (خ و خ و خ) و (خ و خ و ش) و (ش و ش و ش)$\}$

با توجه به مثال فوق به تعریف زیر توجه کنید:

تعریف: مجموعه تمام حالات ممکن از انجام یک آزمایش تصادفی را فضای نمونه نامند و با S نمایش می‌دهند.

اغلب در مسایل به تعداد اعضای مجموعه فضای نمونه یعنی $|S|$ نیازمندیم زیرا در بسیاری از حالات تعداد اعضا بسیار زیاد یا حتی نامتناهی هستند در نتیجه نمی‌توان اعضای مجموعه را لیست نمود.

برای روشن شدن مطلب به مثال زیر توجه منید:

مثال ۴: اگر یک سکه را آنقدر پرتاب کنیم تا برای اولین بار خط بیاید، فضای نمونه مربوط به آزمایش را نمایش دهید:
بوضوح ممکن است یک سکه در اولین پرتاب خط بیاید و یا ممکن است پس از چندین بار شیر، خط بیاید که با در نظر گرفتن این حالات با یک فضای نمونه نامتناهی روبرو هستیم زیرا ممکن است حتی پس از هزار بار پرتاب باز هم شیر بیاید و خط را مشاهده نکنیم !
که البته احتمال وقع چنین حالتی را در ادامه این فصل محاسبه می‌کنیم.

$S=\{$(خ) و(خ و ش) و (خ و ش و ش) و (خ و ش و ش و ش) و ...$\}$

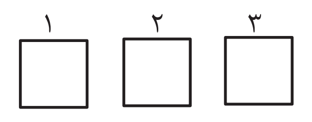
در ادامه برای روشهای شمارش فنونی را با استفاده از اصل ضرب می‌آوریم.

۲-۱-۱ جایگشت

در بسیاری از مسایل قصد داریم چندین شیء را به ترتیب در یک ردیف قرار دهیم. برای حل این گونه مسایل ابتدا مفهوم جایگشت را معرفی می‌کنیم.
تعریف: یک جایگشت از n شئی عبارتست از قرار دادن آنها در یک صف یا ردیف با رعایت یک نظم و ترتیب مشخص.

مثال ۵: حروف A و l و i را به عنوان سه شئی در نظر بگیرید در این صورت کلمه Ali یک جایگشت از این سه حرف می‌باشد. توجه کنید که ترتیب قرار گرفتن حروف مهم است مثلاً ilA یک جایگشت دیگر از این حروف می‌باشد و با جایگشت قبلی متمایز است.

مثال ۶: به چند طریق می‌توان سه عدد کتاب متمایز را در یک قفسه قرار داد؟
ابتدا توجه کنید که کتابها متمایزند پس آنها را $A_۱, A_۲, A_۳$ نامگذاری می‌کنیم از طرفی قرار گرفتن کتابها در قفسه به این معنی است که ترتیب قرار گیری برای ما مهم است. بنابراین می‌بایستی تمام جایگشتهای سه شئی (کتاب) را بدست بیاوریم.
برای هر کتاب در قفسه یک مکان در نظر بگیرید مطابق شکل زیر:



در هر یک از مکانها می‌توانیم هر یک از سه کتاب را قرار دهیم. در مکان اول یکی از سه کتاب را قرار دهیم از آنجا که در هر صورت یک کتاب در مکان اول قرار می‌گیرد در مکان دوم می‌توان یکی از دو کتاب باقیمانده را قرار داد و به همین ترتیب در مکان آخر تنها کتاب باقیمانده قرار می‌گیرد.

از آنجا که این کار را می‌توان در سه مرحله و به صورت پیاپی انجام داد بطوریکه مرحله اول ۳ حالت مرحله دوم ۲ حالت و مرحله سوم ۱ حالت دارد. بنابراین با توجه به اصل ضرب داریم:

$۳\times۲\times۱=۶=$ تعداد کل جایگشتها

که فضای نمونه نیز عبارتست از:

$$S=\{(A_۱,A_۲,A_۳),(A_۱,A_۳,A_۲),(A_۲,A_۱,A_۳),(A_۲,A_۳,A_۱),(A_۳,A_۱,A_۲),(A_۳,A_۲,A_۱)\}$$

در حالت کلی تعداد جایگشتهای n شئی متمایز در یک ردیف (رعایت ترتیب) عبارتست از:

$$n\times(n-۱)\times(n-۲)\cdots(۲)(۱)=n!$$

در نتیجه مثال قبل را می‌توان به صورت خلاصه جایگشت سه شئی در یک ردیف در نظر گرفت که می‌شود:

$$۳!=۶$$

توجه: اگر در محاسبه جایگشت‌ها تکرار اشیاء مجاز باشد، می‌توانیم در تمام مکان‌ها از تمام n شئی استفاده کنیم و در نتیجه تعداد کل جایگشت‌ها عبارتست از: $n\times n\times n\cdots n=n^n$.

مثال ۷: اگر در مثال ۶ از هر سه کتاب به تعداد دلخواه داشته باشیم تعداد کل جایگشت‌ها عبارتست از: $۳\times۳\times۳=۳^۳=۲۷$

۲-۱-۲ ترتیب‌های n شئی به گروههای r تایی (P_r^n)

یک حالت کلی‌تر از جایگشت‌ها قرار دادن n شی متمایز در r $(r\le n)$ مکان، به ترتیب می‌باشد که با توجه به روش بدست آوردن تعداد کل جایگشت‌ها که گفته شد در این حالت نیز تعداد کل جایگشتهای n شی متمایز در r مکان عبارتست از:

$$n(n-۱)(n-۲)\cdots(n-r+۱)=\frac{n!}{(n-r)!}$$

که آنرا با (P_r^n) می‌دهیم. توجه کنید که $P_n^n=n!$ که همان تعداد کل جایگشت‌هاست. و اگر تکرار اشیاء مجاز باشد در این صورت تعداد حالات ممکن عبارتست از:

$$\underbrace{n\times n\times\cdots\times n}_{\text{r بار}}=n^r$$

مثال ۸: به چند طریق می‌توان با حروف A , B , C , D , E یک کلمه سه حرفی ساخت؟

$$P_۳^۵=\frac{۵!}{(۵-۳)!}=\frac{۵!}{۲!}=۵\times۴\times۳=۶۰$$

توجه کنید که حروف متمایز می‌باشند.

مثال ۹: از بین ۱۵ تیم شرکت کننده در مسابقه فوتبال به چند طریق سه تیم رتبه‌های اول و دوم و سوم را بدست می‌آورند؟

$$P_۳^{۱۵}=\frac{۱۵!}{(۱۵-۳)!}=\frac{۱۵!}{۱۲!}=۱۵\times۱۴\times۱۳=۲۷۳۰$$

مثال ۱۰: اگر در مثال ۸ تکرار حروف مجاز باشند تعداد حالات ممکن چند تاست؟

$$n^r:۵\times۵\times۵=۵^۳=۱۲۵$$

حال حالت خاصی از جایگشت را در نظر می‌گیریم که در آن تمامی اشیاء متمایز نیستند.

مثال ۱۱: تعداد جایگشت‌های حروف کلمه book را بدست آورید.

در این حالت حرف o دوبار تکرار شده است بنابراین تمامی اشیاء متمایز نیستند اما مجدداً با استفاده از اصل ضرب می‌توانیم تعداد حالات را محاسبه کنیم.

ابتدا تعداد کل حالات را بدون در نظر گرفتن عدم تمایز اشیاء بدست می‌آوریم که برابر است با:

$$n!=۴!=۲۴$$

دقت کنید که در فضای نمونه مثلاً دو جایگشت obok و obok دوبار تکرار می‌شوند.

زیرا فرض شده است تمام اشیاء (حروف) متمایز باشند اما چون حرف o دوبار تکرار شده استو مسلماً با جابجایی دو حرف o در تمام حالات تغییری در کلمه بوجود نمی‌آید بنابراین طبق اصل ضرب !۲ اضافه در عدد !۴ ضرب شده است که برای حذف آن می‌بایستی !۴ را بر !۲ تقسیم کنیم بنابراین تعداد کل حالات عبارتست از:

$$\frac{۴!}{۲!}=\frac{۲۴}{۲}=۱۲$$

حالت کلی این مساله بصورت زیر است:

اگر از n شئی، $n_۱$ تای آنها مشابه یکدیگر و $n_۲$ تای آنها مشابه یکدیگر و ..... و n_k تای آنها مشابه یکدیگر باشند تعداد حالاتی که می‌توان این n شی را در یک ردیف مرتب کرد عبارتست از:

$$\frac{n!}{n_۱!n_۲!\cdots n_k!}\ ,\quad \sum_{i=۱}^{k} n_i=n:$$

که آنرا با $\left(n_۱,n_۲,\cdots,n_k\right)$ یا $c_n^{n_۱,n_۲,\ldots,n_k}$ نمایش دهیم.

مثال ۱۲: به چند طریق می‌توان دانشجویان یک کلاس ۲۰ نفری را به دسته‌های ۳ و ۴ و ۶ و ۷ نفری تقسیم نمود؟

در این مثال با وجود اینکه تمام اشیاء (دانشجویان) متمایز هستند اما چون هدف تقسیم آنها به چهار دسته ۳ و ۴ و ۶ و ۷ نفری است در نتیجه دانشجویان تخصیص داده شده به هر دسته مثل اشیاء مشابه در نظر گرفته می‌شوند به عبارت بهتر بین تخصیص دانشجوی iام به دسته jام تفاوت قایل هستیم (رعایت ترتیب) اما بین دانشجویان تخصیص داده شده به یک دسته مشخص تفاوتی قایل نیستیم (تشابه) بنابراین تعداد حالات ممکن عبارتست از:

$$\frac{۲۰!}{۳!\,۴!\,۶!\,۷!}$$

توجه کنید که می‌بایستی حتماً مجموع تعداد اعضای دسته‌ها برابر تعداد کل دانشجویان باشد»

$$۳+۴+۶+۷=۲۰$$

۲-۱-۳ ترکیب و مسایل انتخاب

تا بحال با اشیاء متمایز و نظو در ترتیب قرارگیری سروکار داشتیم اما اگر بخواهیم از ترتیب قرار گیری صرف نظر کنیم این حالت به مساله انتخاب r شئی از بین n شئی متمایز تبدیل می‌شود که به آن ترکیب r شئی می‌گوییم.

مثال ۱۳: از بین ۵ مدل اتومبیل سه مدل را انتخاب کنیم به چند طریق این امر امکان‌پذیر است؟

اتومبیل‌ها را از ۱ تا ۵ شماره‌گذاری می‌کنیم و سه جایگاه برای سه مدل اتومبیل انتخابی در نظر می‌گیریم بنابراین طبق جایگشت‌های ۳ شئی از ۵ شئی تعداد کل حالات عبارتست از:

$$P_۳^۵=\frac{۵!}{(۵-۳)!}=۶۰$$

اما چون مساله تنها انتخاب می‌باشد و ترتیب قرار گیری مورد نظر نیست بنابراین به عنوان مثال حالت‌های زیر تکراری محسوب می‌شوند:

$(۱,۲,۳)\quad(۳,۲,۱)\quad(۲,۱,۳)\quad(۲,۳,۱)$

$(۱,۳,۲)\quad(۳,۱,۲)$

به تعداد جایگشت‌های مکان‌های سه اتومبیل، جایگشت تکراری داریم که طبق اصل ضرب یعنی عدد !۳ اضافه در عدد $P_{۳}^{۵}$ ضرب شده است و در نتیجه تعداد کل انتخاب‌ها عبارتست از:

$$\frac{P_{۳}^{۵}}{۳!}=\frac{۵!}{۳!\,۲!}=۱۰$$

در حالت کلی تعداد حالات انتخاب r شئی از بین n شئی متمایز عبارتست از:

$$c_n^r=\binom{n}{r}=\frac{P_r^n}{r!}=\frac{n!}{r!(n-r)!} \qquad (o\le r\le n)$$

که آنرا با $c(n,r)$ یا $\binom{n}{r}$ نمایش می‌دهند.

در مثال بعد حالتی را بررسی می‌کنیم که بتوان از اشیاء تکراری نیز استفاده نمود:

مثال ۱۴: می‌خواهیم از سه دسته اسکناس ۱۰۰ و ۲۰۰ و ۵۰۰ تومانی ۵ عدد اسکناس انتخاب کنیم، این کار به چند طریق ممکن است؟

همانطور که در این مساله مشاهده می‌کنید اشیاء اسکناس ۱۰۰ و ۲۰۰ و ۵۰۰ تومانی می‌باشند که متمایز می‌باشند. اما می‌توانیم از آنها به تعداد دلخواه انتخاب کنیم. مثلاً حالت‌های زیر ممکن است اتفاق بیفتد:

AAABB یا ABCBC یا BBBBB

A= تومانی ۱۰۰

B= تومانی ۲۰۰

C= تومانی ۵۰۰

توجه: می‌توانیم از یک اسکناس اصلاً انتخاب نکنیم مثل حالت BBBBB که از اسکناس ۱۰۰ و ۵۰۰ تومانی استفاده نشده است.

برای حل مساله حالت زیر را در نظر بگیرید:

$$ABCBC \rightarrow ABBCC \rightarrow X|XX|XX$$

زیرا ترتیب مهم نیست

با توجه به حالت بالا A را قبل از پرانتز اول و B را بین دو پرانتز و نهایتاً C را بعد از پرانتز آخر می‌آوریم بنابراین تمام حالت‌ها را می‌توان با قرار دادن دو پرانتز نمایش داد.

پس تعداد کل حالات ممکن از قرار دادن ۵ حرف X و دو پرانتز بدست می‌آید که برابر است با:

$$\binom{۳+۵-۱}{۵}=\binom{۷}{۵}=\frac{۷!}{۵!\,۲!}$$

در حالت کلی انتخاب r شئی از بین n شئی متمایز با تکرار برابر است با:

$$\binom{n+r-۱}{r}$$

۲-۲ فضای نمونه و پیشامد

مجموعه تمام حالت‌های ممکن را از انجام یک آزمایش تصادفی فضای نمونه نامیدیم و آنرا با S نمایش داده‌ایم حال به معرفی پیشامد می‌پردازیم.

تعریف: هر یک از زیر مجموعه‌های فضای نمونه را یک پیشامد می‌نامیم و در صورتی که نتیجه آزمایش تصادفی عضوی از پیشامد E باشد می‌گوییم پیشامد E به وقوع پیوسته است.

مثال ۱۵: تعداد اعضای فضای نمونه و تعداد پیشامدهای ممکن به ازای پرتاب یک عدد تاس را بدست آورید:

حل: پرتاب یک عدد تاس منجر به شش حالت ۱ تا ۶ می‌شود که در نتیجه فضای نمونه عبارتست از $S=\{۱,۲,۳,۴,۵,۶\}$ که ۶ عضو دارد اما چون هر کدام از زیر مجموعه‌های فضای نمونه می‌تواند به عنوان یک پیشامد در نظر گرفته شوند و تعداد کل زیر مجموعه‌های یک مجموعه ۲^N ($n=$ تعداد اعضای مجموعه) می‌باشد بنابراین تعداد پیشامدهای ممکن $۲^۶=۶۴$ می‌باشد.

به عنوان مثال هر کدام از مجموعه‌های A ، B ، C ، D پیشامدهایی برای فضای نمونه S می‌باشند:

$A=\{\phi\} \quad D=\{۱,۲,۳\} \quad C=\{۲\} \quad D=\{۱,۲,۳,۴,۵,۶\}$

۲-۲-۱ انواع پیشامدها و اعمال روی پیشامدها

با توجه به تعریف پیشامد حالت‌های مختلفی برای پیشامدها خواهیم داشت که عبارتند از:

پیشامد ساده: پیشامدی که تنها دارای یک عضو باشد را پیشامد ساده می‌نامیم مثل

$E=\{(\text{خط و شیر})\}$

که در نتیجه حاصل از پرتاب دو سکه با هم می‌باشد.

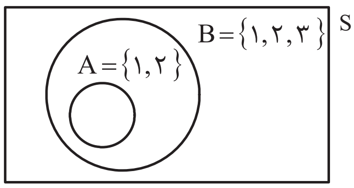
پیشامد مرکب: پیشامدی دارای دو عضو یا بیشتر را مرکب می‌نامیم مثل: $E=\{۱,۲\}$

پیشامد تهی یا محال: اگر پیشامدی دارای هیچ عضوی نباشد به آن پیشامد محال گویند.

پیشامد حتمی: پیشامدی که برابر با فضای نمونه S باشد پیشامد حتمی نامیده می‌شود مثلاً در پرتاب یک عدد تاس $E=\{۱,۲,۳,۴,۵,۶\}$ پیشامد حتمی است.

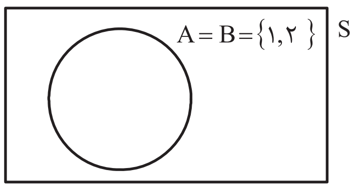
از آنجا که پیشامدها خود یک مجموعه می‌باشند می‌توان اعمالی که روی مجموعه‌ها تعریف می‌شود را روی پیشامدها نیز اعمال نمود. دو پیشامد را A و B و فضای نمونه را S در نظر می‌گیریم.

۱- هر گاه وقوع پیشامد A وقوع پیشامد B را نتیجه دهد گوییم پیشامد A زیر پیشامد B می‌باشد و آنرا بصورت $A\subset B$ نمایش می‌دهیم. این مطلب را در شکل نیز مشاهده می‌کنید.



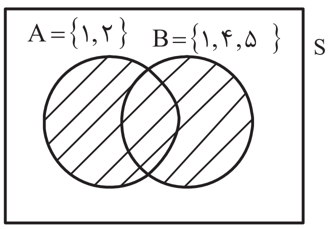
۲- هر گاه وقوع پیشامد A وقوع پیشامد B را نتیجه دهد و بالعکس دو پیشامد را مساوی گویند.

$$A=B \Leftrightarrow (A\subset B \quad , \quad B\subset A)$$



۳- اجتماع دو پیشامد: در صورتی که حداقل یکی از دو پیشامد A و B رخ دهند گویند اجتماع دو پیشامد یا $A \cup B$ رخ داده است و آنرا بصورت زیر می‌نویسیم:

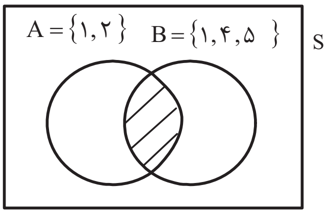
$$A \cup B = \{x \mid x \in A \quad \text{یا} \quad x \in B\}$$



$A\cup B=\{۱,۲,۴,۵\}$

۴- اشتراک دو پیشامد: در صورتی که هر دو پیشامد A و B بصورت همزمان رخ دهند گویند اشتراک آندو یا $A \cap B$ رخ داده است.

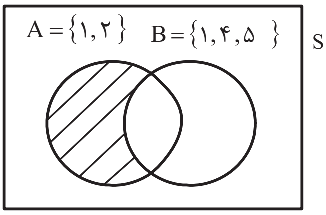
$$A \cap B = \{x \mid x \in A \ , \ \ x \in B\}$$



$A \cap B = \{۱\}$

۵- تفاضل دو پیشامد: اگر فقط A رخ دهد در حالی که B رخ نداده باشد گویند تفاضل دو پیشامد یا A-B رخ داده است.

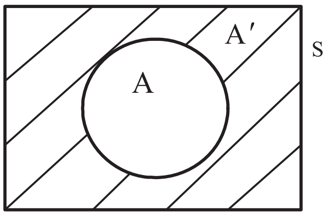
$$A - B = \{x \mid x \in A \ \ , \ x \in B\}$$



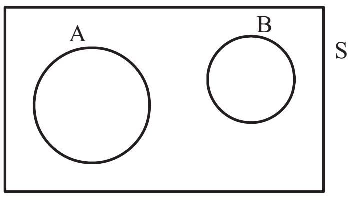
$A - B = \{۲\}$

۶- متمم یک پیشامد: A' را متمم پیشامد A در نظر می‌گیریم وقتیکه رخ دادن A' به معنی عدم وقوع A می‌باشد.

$$A' = \{x \mid x \in S \ \ , \ x \notin A\}$$



۷- در صورتی که دو پیشامد همزمان نتواند رخ دهند یا به عبارتی اشتراکی نداشته باشند آندو را دو پیشامد ناسازگار گوییم در این حالت $A \cap B = \phi$



۲-۳ احتمال

به هر یک از پیشامدها می‌توان عددی نسبت داد که معرف میزان احتمال وقوع آن پیشامد باشد. برای این منظور می‌بایستی از قواعدی پیروی کنیم که برای تمام پیشامدها یکی باشد با استفاده از تعریف فراوانی نسبی می‌توان به محاسبه احتمال پرداخت به این ترتیب که یک آزمایش را N مرتبه انجام می‌دهیم و تعداد دفعاتی که پیشامد A مشاهده شده را $n(A)$ در نظر می‌گیریم در این صورت فراوانی نسبی وقوع پیشامد A عبارتست از $\frac{n(A)}{N}$ که آنرا احتمال وقوع پیشامد A نامیده و با $P(A)$ نمایش می‌دهیم.

فراوانی نسبی همواره عددی بین صفر و یک می‌باشد بنابراین مقدار احتمال نیز در این بازه می‌باشد از طرفی احتمال وقوع تمام اعضای فضای نمونه برابر ۱ می‌باشد یعنی $p(S)=۱$ از طرفی تعداد دفعات مشاهده یکی از دو پیشامد A یا B در صورتی که آندو با یکدیگر اشتراکی نداشته باشند برابر است با تعداد دفعات مشاهده $A \cap B$ بنابراین $p(A \cup B) = p(A) + p(B)$ به شرطی که $A \cap B = \phi$

سه قلعده بدست آمده در بالا، اصول موضوعه احتمال می‌گوییم و بطور کلی احتمالات باید تابع قوانین زیر باشند:

$۱ - p(S) = ۱$

$۲ - p(A) \geq ۰ \qquad \forall \ A \subset S$

$۳ - p(A_۱ \cup A_۲ \cup \cdots) = p(A_۱) + p(A_۲) + \cdots$

هر گاه $A_i \cap A_j = \phi$ ، $\forall \ i \neq j$

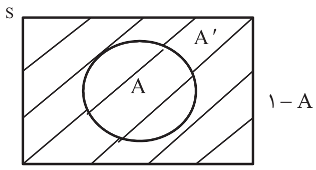
۲ . ۳ . ۱ قوانین احتمال

قبلاً با استفاده از قوانین مجموعه‌ها اعمالی را روی پیشامدها تعریف کردیم حالا با استفاده از آنها و اصول موضوعه احتمال، چندین قانون برای احتمالات بدست می‌آوریم.

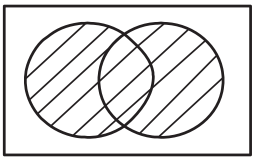
قضیه: $p(\phi) = ۰$

اثبات: $S \cup \phi = \phi$ پس بنابر اصل ۳

$$p(S\cup\phi) = p(s) + p(\phi) = ۱ + p(\phi) \quad \Rightarrow \quad ۱ + p(\phi) = ۱ \Rightarrow p(\phi) = ۰$$



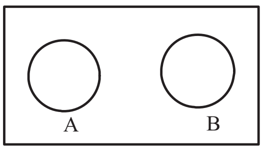
با استدلالی مشابه می‌توان قضایای زیر را اثبات نمود:



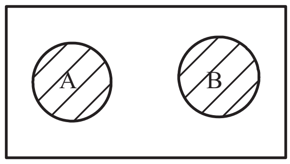
۱- $p(A') = ۱ - p(A)$

۲- $p(A\cup B) = p(A) + p(B) - p(A\cap B)$

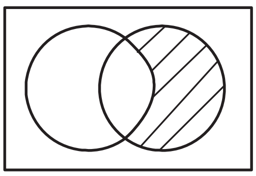
۳- اگر A و B ناسازگار باشند داریم $A\cap B = \phi$ و در نتیجه



$p(A\cup B) = ۰$

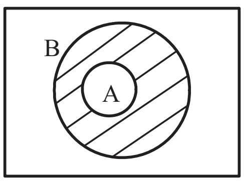


$p(A\cup B) = p(A) + p(B)$



۴- $p(A\cup B) = p(B\cap A') = p(B) - p(A\cap B)$

۵- اگر $A \subset B$ انگاه:



$$\begin{cases} p(B-A) = p(B) - p(A) \\ p(A) \le p(B) \end{cases}$$

مثال ۱۶: سه سکه را همزمان پرتاب می‌کنیم احتمال بدست آوردن سه شیر یا سه خط چقدر است؟

حل: پیشامد A را بدست آوردن سه شیر در نظر می‌گیریم و به همین ترتیب پیشامد B را بدست آمدن سه خط

$A = \{$ش و ش و ش$\}$ $\qquad$ $B = \{$خ و خ و خ$\}$

از آنجا که $A\cap B = \phi$ بنابراین A و b ناسازگار هستند و در نتیجه پیشامد بدست آمدن سه شیر یا سه خط برابر است با:

$p(A\cup B) = p(A) + p(B)$

با توجه به قواعد شمارش تعداد اعضای فضای نمونه برابر است با: $n(A) = ۱ ,\ n(S) = ۲^{۳} = ۸$

بنابراین:

$$p(A)=\frac{n(A)}{n(S)}=\frac{۱}{۸}$$

$$p(B)=\frac{n(B)}{n(S)}=\frac{۱}{۸}$$

پس:

$$p(A\cup B)=p(A)+p(B)=\frac{۱}{۸}+\frac{۱}{۸}=\frac{۲}{۸}$$

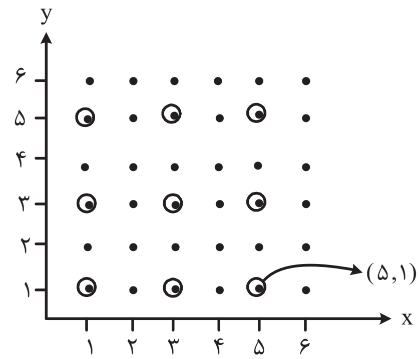
مثال ۱۷: یک جفت تاس را می‌ریزیم احتمال بدست آمدن اعدادی فرد روی هر دو تاس چقدر است؟

A={آمدن اعداد فرد بر روی دو تاس}

$$n(A)=۳\times۳=۹$$

سه عدد فرد بر روی تاس اول $\{۱,۳,۵\}$

در شکل روبرو اعضای مجموعه A در دوایری نشان داده شده‌اند:



تعداد کل حالات از $n(S)=۶\times۶=۳۶$

$$p(A)=\frac{n(A)}{n(S)}=\frac{۹}{۳۶}=\frac{۱}{۴}$$

مثال ۱۸: یک سکه را ۵ مرتبه پرتاب می‌کنیم احتمال اینکه حداقل دوبار شیر بیاید چقدر است؟

$A=$ پیشامد اینکه حداقل ۲ بار شیر بیاید

$A'=$ پیشامد یا اصلاً شیر نیاید یا دقیقاً یکبار شیر بیاید $=\left\{\begin{array}{l}\text{و (خ و خ و ش و خ) و (خ و خ و خ و ش) و (خ و خ و خ و خ)}\\ \text{(ش و خ و خ و خ) و (خ و ش و خ و خ و خ) و (خ و خ و ش و خ و خ)}\end{array}\right\}$

$$n(S)=۲^۵=۳۲$$

$$p(A')=\frac{۶}{۳۲}=\frac{۳}{۱۶}$$

$$p(A)=۱-p(A')=۱-\frac{۳}{۱۶}=\frac{۱۳}{۱۶}$$

مثال ۱۹: الف) تعداد جوابهای صحیح و غیر منفی معادله $x_۱+x_۲+\cdots+x_n=r$ را در صورتیکه $x_i\geq ۰\ ,\ i=۱,۲,\ldots,n$ باشد را بدست آورید؟

ب) می‌خواهیم ۷ عدد بلیط را بین سه نفر تقسیم کنیم بطوریکه به هر نفر حداقل یک بلیط برسد مطلوبست احتمال اینکه به نفر دوم حداقل ۲ بلیط برسد؟

حل: الف) این مساله مشابه حالتیست که بخواهیم r شئی را بین n نفر تقسیم کنیم بطوریکه در این حالت به هر شخص می‌توان بیشتر از یک شئی داد و ترتیب اهمیتی ندارد در قسمت قوانین شمارش نشان دادیم که جواب چنین مساله‌ای با استفاده از رابطه $\binom{n+r-۱}{r}$ بدست می‌آید.

ب) تعداد کل حالات فضای نمونه در این حالت از حل معادله $x_i\geq ۱$ $\quad x_۱+x_۲+x_۳=۷$ در نظر گرفته شود اما برای استفاده از رابطه $\binom{n+r-۱}{r}$ باید $x_i\geq ۰$ باشد بنابراین از تغییر متغیر زیر استفاده می‌کنیم.

$$x_i \geq 1 \;\rightarrow\; x_i - 1 \geq 0 \quad y_i = x_i - 1 \;\Rightarrow\; y_i \geq 0$$
$$\Rightarrow\; x_i = y_i + 1 \qquad y_1 + 1 + y_2 + 1 + y_3 + 1 = 7$$
$$\Rightarrow y_1 + y_2 + y_3 = 7 - 3 = 4$$

بنابراین مساله به صورت زیر خلاصه می‌شود:

$$\begin{cases} y_1 + y_2 + y_3 = 4 \\ y_i \geq 0 \end{cases}$$

که جواب آن عبارتست از $n(s) = \binom{3+4-1}{4} = 15$

برای بدست آوردن تعداد حالات مطلوب باید مساله زیر را حل کنیم:

$$\begin{cases} x_1 + x_2 + x_3 = 7 \\ x_1 \;,\; x_3 \geq 1 \\ x_2 \geq 2 \end{cases}$$

دوباره با استفاده از تغییر متغیر داریم:

$$y_1 = x_1 - 1 \quad , \quad y_3 = x_3 - 1 \quad , \quad y_2 = x_2 - 2$$
$$\Rightarrow y_i \geq 0$$

که در نتیجه مساله به فرم زیر تبدیل می‌شود:

$$A: \begin{cases} y_1 + y_2 + y_3 = 3 \\ y_i \geq 0 \end{cases}$$

که جواب آن عبارتست از:

$$n(A) = \binom{3+3-1}{3} = 10$$

بنابراین احتمال فوق عبارتست از:

$$p(A) = \frac{n(A)}{n(S)} = \frac{10}{15} = \frac{2}{3}$$

۲ . ۴ احتمال روی فضای نمونه نامتناهی

با توجه به نوع مسایل فضای نمونه می‌تواند نامتناهی باشد در این حالت مجموعه پیشامدها نیز نامتناهی خواهند بود البته احتمال پیشامدها می‌بایستی در اصول موضوعه صدق کند بطوریکه $0 \leq p(A_i) \leq 1 \;,\; \sum_{i=1}^{+\infty} p(A_i) = 1$.

مثال ۲۰: A و B به ترتیب به سوی هدفی شلیک می‌کنند A با احتمال $\frac{1}{2}$ و B با احتمال $\frac{3}{4}$ هدف را مورد اصابت قرار می‌دهند مطلوبست احتمال اینکه B زودتر از A هدف را بزند؟ (در صورتی که A اول شروع کند).

پیشامد اینکه A هدف را بزند : A'

$$p(A') = 1 - \frac{1}{2} = \frac{1}{2}$$

احتمال اینکه B به هدف نزند : B'

برای اینکه B زودتر هدف را مورد اصابت قرار دهد حالات زیر را خواهیم داشت:

پیشامد	$A'B$	$A'B'A'B$	$A'B'A'B'A'B'$	$\cdots$
احتمال	$\frac{1}{2} \times \frac{3}{4}$	$(\frac{1}{2})^2 \frac{3}{4} \times \frac{1}{4}$	$(\frac{1}{2})^3 (\frac{1}{4})^2 \frac{3}{4}$	$\cdots$

$$p(B')=1-\frac{3}{4}=\frac{1}{4}$$

از آنجا که پیشامدها ناسازگار هستند باید احتمال اجتماع تمام پیشامدها محاسبه شود که برابرست با مجموع احتمالات هر یک از پیشامدها یعنی:

$$p(A'B\cup A'B'A'B\cup\cdots)=p(A'B)+p(A'B'A'B)+\cdots=\frac{1}{2}\times\frac{3}{4}+(\frac{1}{2})^2\frac{1}{4}\times\frac{3}{4}+\cdots$$

$$=\frac{3}{4}\times\frac{1}{2}(1+\frac{1}{8}+(\frac{1}{8})^2+\cdots)=\frac{3}{8}\frac{1}{1-\frac{1}{8}}=\frac{3}{7}$$

مثال ۲۱: سکه‌ای را آنقدر پرتاب می‌کنیم تا برای اولین بار شیر مشاهده شود مطلوبست:

الف) احتمال آنکه تعداد زوجی پرتاب لازم باشد.

ب) احتمال آنکه حداکثر ۱۰ پرتاب لازم باشد.

حل: برای اینکه برای بار اول شیر مشاهده کنیم می‌بایستی در تمام پرتاب‌های قبلی خط مشاهده شده باشد در این صورت احتمال‌ها را می‌توان در جدول زیر خلاصه کنیم:

S	e_1 ش	e_2 ش خ	e_3 ش خ خ	e_4 ش خ خ خ	$\cdots$
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\cdots$

پیشامد اینکه تعداد زوجی پرتاب لازم باشد: A

الف) برای اینکه تعداد زوجی پرتاب لازم باشد باید پیشامدهای $e_2, e_4, e_6, \ldots$ رخ دهند بنابراین:

$$p(A)=\frac{1}{4}+\frac{1}{16}+\frac{1}{24}+\cdots=\frac{\frac{1}{4}}{1-\frac{1}{4}}=\frac{1}{3}$$

همچنین می‌توان احتمال اینکه تعداد فدی پرتاب لازم باشد را از روی $p(A)$ محاسبه نمود. چون پیشامد تعداد فردی پرتاب متمم پیشامد A می‌باشد بنابراین:

$$p(A')=1-p(A)=1-\frac{1}{2}=\frac{2}{3}$$

ب) پیشامد انکه حداکثر ۱۰ پرتاب لازم باشد $=B$

برای محاسبه احتمال پیشامد B می‌بایستی مجموع ۱۰ جمله اول از جدول را بدست بیاوریم اما با استفاده از متمم این پیشامد می‌توان تعداد آنرا به راحتی محاسبه نمود:

پیشامد اینکه حداقل ۱۱ پرتاب لازم باشد $=B'$

$$p(B')=(\frac{1}{2})^{11}+(\frac{1}{2})^{12}+\cdots=\frac{(\frac{1}{2})^{11}}{1-\frac{1}{2}}=(\frac{1}{2})^{10}$$

$$p(B)=1-p(B')=1-(\frac{1}{2})^{10}=\frac{1023}{1024}\approx 0/999$$

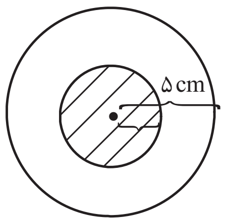
۲ . ۵ فضای نمونه پیوسته

در شرایطی می‌توان فضای نمونه را طوری تعریف نمود که بصورت یک بازه یا یک سطح در نظر گرفته شود. مثلاً $S=[1,2]$ در این حالت هر زیر بازه مثل $E=[1/5,1/75]$ می‌تواند به عنوان یک پیشامد در نظر گرفته شود. برای محاسبه احتمال پیشامد باز هم حالت مطلوب را به کل فضای نمونه در نظر می‌گیریم یعنی:

$$p(A)=\frac{\text{طول بازه } A}{\text{طول فاصله } S} \quad \text{یا} \quad \frac{\text{مساحت ناحیه } A}{\text{مساحت فضای نمونه } S}$$

مثال ۲۲: تیراندازی به هدفی شلیک می‌کند که قطر آن ۱۰ سانتیمتر و قطر دایره مرکزی هدف ۲ سانتیمتر است احتمال اینکه تیر به مرکز هدف اصابت کند چقدر است؟

حل:



با توجه به شکل مساحت فضای نمونه عبارتست از: 25π

و مساحت ناحیه مطلوب عبارتست از: π

بنابراین احتمال مورد نظر عبارتست از:

$$p(A)=\frac{\pi}{25\pi}=\frac{1}{25}$$

مثال ۲۳: در مثال قبل احتمال اینکه تیرانداز دقیقاً مرکز هدف را بزند چقدر است؟

از آنجا که مرکز هدف یک نقطه محسوب می‌شود بنابراین مساحت آن صفر واحد می‌باشد و در نتیجه احتمال مورد نظر برابر صفر خواهد بود.

۲ . ۶ احتمال شرطی

در بسیاری از مواقع می‌دانیم که پیشامد A رخ داده و می‌خواهیم احتمال رخ دادن پیشامد B را مشروط بر اینکه A رخ داده است بدست بیاوریم در این حالت می‌بایستی از تعاریف احتمال شرطی استفاده کنیم. به مثال زیر توجه کنید:

مثال ۲۴: یک تاس طوری طراحی ششده است که احتمال آمدن هر عدد متناسب با آن عدد می‌باشد در این صورت:

الف) احتمال رخ دادن یک عدد زوج را بیابید.

ب) اگر بدانیم عدد حاصل از پرتاب کوچکتر یا مساوی با عدد ۴ می‌باشد احتمال آمدن یک عدد فرد را بیابید.

ابتدا احتمال هر عدد را محاسبه می‌کنیم با توجه به اینکه احتمال آمدن هر عدد باید متناسب با آن عدد باشد معادله زیر را داریم:

$$x+2x+3x+4x+5x+6x=1$$

که در آن x احتمال آمدن عدد ۱ می‌باشد.

$$21x=1 \quad \rightarrow \quad x=\frac{1}{21}$$

بنابراین:

S	۱	۲	۳	۴	۵	۶
احتمال	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

توجه کنید که مجموع احتمالات برابر ۱ می‌باشد.

الف) پیشامد رخ دادن عدد زوج $= A$

$$A=\{2,4,6\}$$

$$p(A)=\frac{2}{21}+\frac{4}{21}+\frac{6}{21}=\frac{12}{21}$$

ب) در این حالت مجموعه فضای نمونه محدود می‌باشد به $B=\{1,2,3,4\}$ زیرا می‌دانیم عدد بدست آمده کوچکتر یا مساوی با ۴ می‌باشد. از طرفی چون احتمال آمدن هر عدد متناسب با همان عدد است پس مدل احتمال روی این فضای نمونه جدید عبارتست از:

$$x+2x+3x+4x=1 \quad \rightarrow \quad 10x=1 \quad \rightarrow \quad x=\frac{1}{10}$$

S	۱	۲	۳	۴
احتمال	$\frac{۱}{۱۰}$	$\frac{۲}{۱۰}$	$\frac{۳}{۱۰}$	$\frac{۴}{۱۰}$

$C=$ (پیشامد رخ دادن عدد فرد به شرط وقوع B)

بنابراین احتمال وقوع یک عدد فرد به شرطی که پیشامد B رخ داده باشد عبارتست از:

$$p(C)=\frac{۱}{۱۰}+\frac{۳}{۱۰}=\frac{۴}{۱۰}=\frac{۲}{۵}$$

در حالت کلی تعریف زیر را خواهیم داشت:

تعریف: احتمال وقوع پیشامد B به شرط وقوع پیشامد A که آنرا با نماد $P(B|A)$ نمایش می‌دهیم به صورت زیر تعریف می‌شود:

$$P(B|A)=\frac{P(A\cap B)}{p(A)} \quad , \quad P(A)\neq ۰$$

مثال ۲۵: با استفاده از تعریف احتمال مثال ۲۴ (ب) را محاسبه می‌کنیم.

$A=\{۱,۲,۳,۴\}$ پیشامد وقوع عددی کوچکتر از ۴

$B=$ پیشامد رخ دادن عددی فرد $=\{۱,۳,۵\}$

$$P(A)=\frac{۱}{۲۱}+\frac{۲}{۲۱}+\frac{۳}{۲۱}+\frac{۴}{۲۱}=\frac{۱۰}{۲۱}$$

$$A\cap B=\{۱,۳\} \rightarrow P(A\cap B)=\frac{۱}{۲۱}+\frac{۳}{۲۱}=\frac{۴}{۲۱}$$

$$\Rightarrow P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{۴}{۲۱}}{\frac{۱۰}{۲۱}}=\frac{۴}{۱۰}=\frac{۲}{۵}$$

ملاحظه می‌کنید که در این حالت دیگر نیازی به محاسبه جدول احتمالات میانی نمی‌باشد.

مثال ۲۶: در جعبه‌ای ۵ مهره به رنگ آبی و ۴ مهره به رنگ قرمز موجود می‌باشند از این جعبه ۳ مهره به تصادف خارج می‌کنیم اگر بدانیم دو مهره از سه مهره به رنگ آبی می‌باشند احتمال اینکه مهره سوم به رنگ قرمز باشد چقدر است؟

$A=$ پیشامد اینکه دو مهره از سه مهره آبی باشند

$B=$ پیشامد اینکه یک مهره از سه مهره قرمز باشد

$A\cap B=$ پیشامد اینکه دو مهره آبی و یک مهره قرمز باشد

$$P(A\cap B)=\frac{\binom{۴}{۱}\binom{۵}{۲}}{\binom{۹}{۳}} \qquad P(A)=\frac{\binom{۵}{۲}\binom{۴}{۱}+\binom{۵}{۳}}{\binom{۹}{۳}}$$

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{\binom{۵}{۳}\binom{۴}{۱}}{\binom{۹}{۳}}}{\frac{\binom{۵}{۲}\binom{۴}{۱}+\binom{۵}{۳}}{\binom{۹}{۳}}}$$

۲ . ۷ قانون ضرب احتمال

در صورتی که فرمول $P(A|B)=\dfrac{P(A\cap B)}{P(B)}$ را ساده کنیم خواهیم داشت:

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B)=P(A|B)\,P(B)$$

$$P(B|A)=\frac{P(A\cap B)}{P(A)} \Rightarrow P(A\cap B)=P(B|A)\,P(A)$$

در نتیجه فرمول کلی زیر را بدست می‌آوریم:

$$P(A\cap B) = P(A|B)P(B) = P(B|A)\,P(A)$$

این رابطه در حالتی که دو پیشامد A وb بتوانند بصورت همزمان رخ دهند به کار می‌رود و به آن قانون ضرب احتمال می‌گوییم.

مثال ۲۸: جدول زیر احتمال شاغل بودن مردان و زنان را در یک جامعه آماری نشان می‌دهد:

مثلاً در جدول احتمال اینکه شخص مرد باشد ۰/۵۷ و احتمال اینکه مرد باشد و شاغل هم باشد ۰/۵۲ می‌باشد.

		M مرد	F زن	جمع
E	شاغل	۰/۵۲	۰/۴۱	۰/۹۳
U	بیکار	۰/۰۵	۰/۰۲	۰/۰۷
	جمع	۰/۵۷	۰/۴۳	۱/۰۰

با توجه به جدول به سوالات زیر پاسخ دهید:

الف) اگر از جامعه فوق یک نفر انتخاب کنیم و بدانیم شاغل است احتمال اینکه مرد باشد چقدر است؟

ب) احتمال اینکه نفر انتخابی از جامعه فوق شاغل باشد؟

ج) اگر نفر انتخابی مرد باشد احتمال اینکه شاغل باشد چقدر است؟

حل:

الف) تعریف می‌کنیم:

$U=$ پیشامد اینکه نفر انتخابی بیکار باشد

$E=$ پیشامد اینکه نفر انتخابی شاغل باشد

$M=$ پیشامد اینکه نفر انتخابی مرد باشد

$F=$ پیشامد اینکه نفر انتخابی زن باشد

بنابراین مقدار $P(E|M)$ را می‌خواهیم:

با توجه به جدول داریم: $P(E\cap M)=۰/۵۲$

$$P(M)=P(M\cap E)+P(M\cap U)=۰/۵۲+۰/۰۵=۰/۵۷$$

$$P(E|M)=\frac{P(E\cap M)}{P(M)}=\frac{۰/۵۲}{۰/۵۷}=۰/۹۱$$

ب) پس:

$$P(E)=P(E\cap M)+P(E\cap F)=۰/۵۲+۰/۴۱=۰/۹۳$$

ج) $P(M|E)=\dfrac{P(M\cap E)}{P(E)}=\dfrac{۰/۵۲}{۰/۹۳}=۰/۵۶$

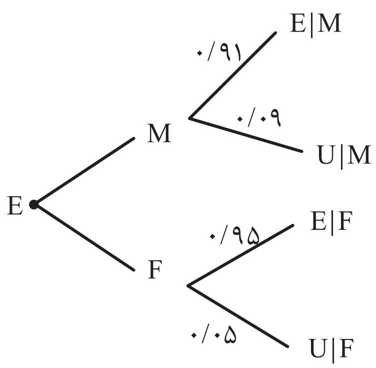
مثال ۲۹: فرض کنید در مثال قبل تنها اطلاعات زیر را در اختیار داشته باشیم:

$P(M)=۰/۵۷ \qquad P(E|M)=۰/۹۱ \qquad P(U|M)=۰/۰۹$

$P(F)=۰/۴۳ \qquad P(E|F)=۰/۹۵ \qquad P(U|F)=۰/۰۵$

در اینصورت مقادیر $P(E)$ و $P(M|E)$ را محاسبه کنید.

حل: احتمالات داده شده را می‌توان به صورت درختی در نظر گرفت در این صورت داریم:



که طبق قانون ضرب احتمال:

$$P(E)=P(E\cap M)+P(E\cap F)=P(E|M)\,P(M)+P(E|F)\,P(F)$$
$$=۰/۹۱\times۰/۵۷+۰/۹۵\times۰/۴۲=۰/۹۳$$
$$P(M|E)=\frac{P(M\cap E)}{P(E)}=\frac{P(E|M)\times P(M)}{P(E\cap M)+P(E\cap F)}$$
$$=\frac{P(E|M)\,P(M)}{P(E|M)\,P(M)+P(E|F)\,P(F)}=\frac{۰/۹۱\times۰/۵۷}{۰/۹۱\times۰/۵۷+۰/۹۵\times۰/۴۳}=\frac{۰/۵۲}{۰/۹۳}=۰/۵۶$$

۲ . ۸ پیشامدهای مستقل

رابطه $P(A|B)=\dfrac{P(A\cap B)}{P(B)}$ را در نظر بگیرید، اگر تحت شرایطی مقدار $P(A\cap B)$ برابر $P(A)\times P(B)$ شود خواهیم داشت:

$$P(A|B)=\frac{P(A)\times P(B)}{P(B)}=P(A)$$

به عبارتی احتمال وقوع A به شرط B برابر با احتمال وقوع A می‌باشد و این یعنی دانستن اینکه پیشامد B رخ داده است هیچ اطلاعاتی در مورد رخ دادن پیشامد A بدست نمی‌دهد بنابراین می‌توان گفت دو پیشامد از یکدیگر مستقل هستند و تعریف زیر را خواهیم داشت»

تعریف: دو پیشامد A و B را از یکدیگر مستقل می‌نامیم اگر داشته باشیم:

$$P(A\cap B)=P(A)\times P(B)$$

مثال ۳۰: در یک ایستگاه مترو احتمال اینکه قطار به موقع در ایستگاه قرار داشته باشد ۰/۹ می‌باشد مطلوبست:

الف) احتمال اینکه قطار سه روز متوالی به موقع در ایستگاه باشد.

ب) احتمال اینکه قطار در روز سوم برای بار دوم دیر کند.

حل:

الف) ابتدا پیشامد زیر را تعریف می‌کنیم:

احتمال اینکه قطار در روز iام به موقع در ایستگاه قرار داشته باشد $= A$

در این صورت می‌توان پیشامدهای $A_۱ , A_۲ , A_۳$ را زا یکدیگر مستقل در نظر گرفت زیرا پیشامد اینکه قطار امروز دیر کند به پیشامد اینکه قطار فردا هم دیر کند ارتباطی ندارد بنابراین با توجه به فرمول استقلال داریم:

$$P(A_۱\cap A_۲\cap A_۳)=P(A_۱)\times P(A_۲)\times P(A_۳)=۰/۹\times ۰/۹\times ۰/۹=۰/۷۲۹$$

ب) پیشامد زیر را تعریف می‌کنیم:

پیشامد اینکه قطار در روز iام دیر کند $= A'_i$

$$P(A'_i)=۱-P(A_i)=۱-۰/۹=۰/۱$$

برای اینکه قطار در روز سوم برای بار دوم دیر کند می‌بایستی در دو روز قبل حداقل یکبار دیر کرده باشد. احتمال زیر را محاسبه می‌کنیم:

$$P\left[(A'_۱\cap A'_۲\cap A'_۳)\cup(A_۱\cap A'_۲\cap A'_۳)\right]=$$
$$=P(A'_۱\cap A'_۲\cap A'_۳)+P(A_۱\cap A'_۲\cap A'_۳)$$
$$=P(A'_۱)\ P(A_۲)\ P(A'_۳)+P(A_۱)\ P(A'_۲)\ P(A'_۳)$$
$$=۰/۱\times ۰/۹\times ۰/۱+۰/۹\times ۰/۱\times ۰/۱=۰/۰۱۸$$

استقلال سه پیشامد: سه پیشامد A ، B و C را مستقل از یکدیگر می‌نامیم اگر داشته باشیم:

۱- $P(A\cap B)=P(A)\ P(B)$

۲- $P(A\cap C)=P(A)\ P(C)$

۳- $P(B\cap C)=P(B)\ P(C)$

۴- $P(A\cap B\cap C)=P(A)\ P(B)\ P(C)$

پیشامدهای ناسازگار و مستقل: به تفاوت‌های پیشامدهای ناسازگار و مستقل توجه کنید:

- اگر دو پیشامد ناسازگار باشند: نمی‌توانند همزمان رخ دهند، احتمال اشتراک آنها صفر می‌باشد، احتمال اجتماع آنها برابر مجموع احتمالات هر یک می‌باشد.

- اگر دو پیشامد مستقل از یکدیگر باشند: هر دو می‌توانند همزمان رخ دهند، احتمال اشتراک آنها برابر با حاصل ضرب احتمالات آنهاست، احتمال اجتماع آنها کوچکتر یا مساوی با مجموع احتمالات هر یک می‌باشد.

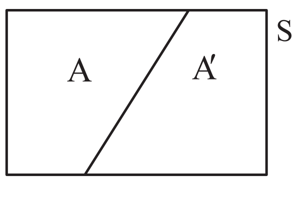
* توجه کنید که اگر دو پیشامد بخواهند بصورت همزمان هم ناسازگار باشند و هم مستقل در نتیجه باید داشته باشیم:

$$\begin{cases}P(A\cap B)=۰\\ P(A\cap B)=P(A)\ P(B)\end{cases}\Rightarrow\quad P(A(\ P(B)=۰\quad\Rightarrow\quad P(A)=۰\quad \text{یا}\quad P(B)=۰$$

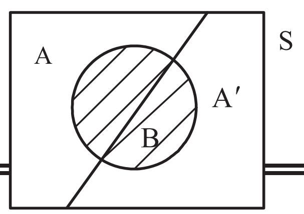
۲ . ۹ فرمول احتمال بینر و فرمول تفکیک احتمال

فرمول احتمال بینر روش ساده‌تری برای محاسبه احتمالات شرطی در حالتی که اطلاعات کمی در مورد مساله داریم ارایه می‌کند که در ذیل نحوه بدست آوردن آنرا شرح می‌دهیم:

فرض کنید فضای نمونه S را به دو پیشامد A و A' که متمم آن می‌باشد تقسیم کنیم به صورت شکل زیر:



حال برای حل مسئله در حالت کلی پیشامد B را در این فضا طوری در نظر می‌گیریم که با A و A' اشتراک داشته باشد:



می‌خواهیم احتمال A به شرط وقوع B را محاسبه کنیم:

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \qquad ۱-۱$$

از روی شکل می‌توان به راحتی احتمال وقوع B را محاسبه نمود.

$$B=(A\cap B)\cup(A'\cap B)$$

$$P(B)=P(A\cap B)+P(A'\cap B)=P(A)\,P(B|A)+P(A')\,P(B\,|\,A') \qquad ۱-۲$$

به فرمول فوق فرمول تفکیک احتمال گویند که حالت کلی‌تر آنرا در ادامه بدست می‌آوریم.

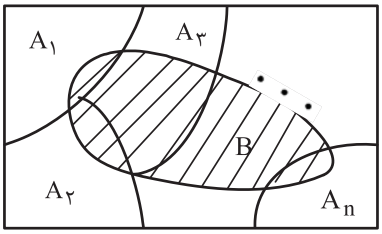
$$P(A\cap B)=P(A)\,P(B\,|\,A) \qquad ۱-۳$$

$$۱-۱,۱-۲,۱-۳ \quad\Rightarrow\quad P(A|B)=\frac{P(A)\,P(B|A)}{P(A)\,P(B\,|A)+P(A)\,P(B|A')} \qquad ۱-۴$$

رابطه ۴-۱ به فرمول بنیر معروف است.

برای بدست آوردن حالت کلی‌تر روابط ۲-۱ و ۴-۱ فرض کنید فضای نمونه به پیشامدهای $A_۱, A_۲, \ldots, A_n$ طوری تقسیم شده باشد که به ازای هر $i\neq j$ داشته باشیم $A_i\cap A_j=\phi$ و $S=\bigcup_{i=۱}^{n} A_i$.

مطابق شکل زیر:



در این حالت اصطلاحاً می‌گوییم فضای نمونه S به n پیشامد $A_۱$ تا A_n افراز شده است. و پیشامد B نیز پیشامد دلخواه در فضای نمونه S باشد در این صورت داریم:

$$B=B(B\cap S)=B\cap\left(\bigcup_{۱}^{n}A_i\right)=(B\cap A_۱)\cup(B\cap A_۲)\cup\cdots\cup(B\cap A_n)$$

که در این رابطه پیشامدهای $(B\cap A_۱)\,,\,(B\cap A_۲)\,,\,\cdots(B\cap A_n)$ دو به دو ناسازگارند.

بنابراین داریم:

$$P(B)=P(B\cap A_۱)+P(B\cap A_۲)+\cdots+P(B\cap A_n)=P(A_۱)\,P(B|A_۱)+P(A_۲)\,P(B|A_۲)$$

$$+\cdots+P(A_n)\,P(B|\,A_n)=\sum_{i=۱}^{n}P(A_i)\,P(B|A_i)$$

این رابطه فرمول تفکیک احتمال یا فرمول احتمال کل نامیده می‌شود. به همین ترتیب برای $P(A_i\,|\,B)$ داریم:

$$P(A_i|B)=\frac{P(A_i\cap B)}{P(B)}=\frac{P(A_i)\,P(B\cap A_i)}{\sum_{i=۱}^{n}P(A_i)\,P(B\cap A_i)}$$

به این رابطه فرمول احتمال بینر می‌گوییم.

مثال ۳۱: احتمال افزایش قیمت سهام یک شرکت خصوصی در بورس در طول ۲ ماه مهر و آبان برابر ۰/۶ و احتمال سقوط قیمت سهام آن در طول این ۲ ماه برابر ۰/۳ است. همچنین اگر قیمت سهام شرکت در طول یکی از این دو ماه افزایش و در ماه دیگر کاهش پیدا کند در این صورت احتمال

اینکه در ماه اول قیمت سهام افزایش پیدا کند برابر $\frac{۱}{۲}$ می‌شود. مطلوبست احتمال اینکه قیمت سهام شرکت در ماه دوم هم افزایش پیدا کند به شرطی که در ماه اول افزایش پیدا کرده باشد.

حل: برای فضای نمونه ۴ حالت زیر را داریم:

$$S\{(I , I) , (D , D) , (I , D) , (I , D)\}$$

افزایش قیمت سهام: I

کاهش قیمت سهام: D

توجه: برای سادگی حالت ثابت ماندن قیمت سهام در نظر نمی‌گیریم.

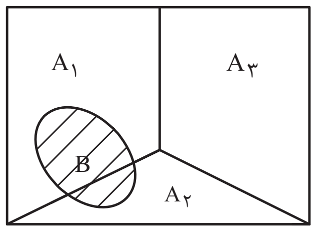
مثلاً $(I-D)$ یعنی قیمت سهام در ماه اول افزایش و در ماه دوم کاهش یافته است. حال پیشامدهای زیر را تعریف می‌نیم:

$A_۱ =$ پیشامد اینکه قیمت سهام در هر ماه افزایش پیدا می‌کند $=\{(I , I)\}$

$A_۲ =$ پیشامد اینکه قیمت سهام در یک ماه افزایش و در یک ماه کاهش داشته باشد $=\{(I , D) , (D , I)\}$

$A_۳ =$ پیشامد اینکه قیمت سهام در هر دو ماه کاهش داشته باشد $=\{(D , D)\}$

$B=$ پیشامد اینکه قیمت سهام در هر ماه اول افزایش داشته باشد



$P(A_۱|B)=$ (افزایش در ماه اول | افزایش در ماه دوم) $P=$ (افزایش در ماه دوم) P. بنابراین می‌بایستی مقدار احتمال $P(A_۱|B)$ را محاسبه کنیم با توجه به فرمول بینر داریم:

$$P(A_۱|B) = \frac{P(A_۱)\, P(B|A_۱)}{P(A_۱)\, P(B|A_۱)+ P(A_۲)\, P(B|A_۲) + P(A_۳)\, P(B|A_۳)}$$

$$P(A_۱) = ۰/۶ \quad , \quad P(B|A_۱)=۱ \quad , \quad P(B|A_۲) = \frac{۱}{۲}, \quad P(B|A_۳)= \frac{P(B\cap A_۳)}{P(A_۳)} P(A_۳) = \frac{۰}{P(A_۳)} = ۰$$

بنابراین داریم:

$$P(A_۱|B) = \frac{۰/۶ \times ۱}{۰/۶ \times ۱+۰/۱ \times \frac{۱}{۲}+۰} = ۰/۹۲۳$$

به نام خدا

فصل سوم
متغیر های تصادفی و توابع توزیع

Jalase 1 – sco 1

تعریف متغیر تصادفی

نتایج حاصل از یک آزمایش تصادفی را به صورت های مختلفی می توان بیان نمود . مثلا شیر یا خط را با نمواد 1 و -1 نیز نمایش داد . از آنجا که این اعداد در آزمایش به صورت تصادفی ظاهر می شوند ، می توان آنها را با یک متغیر تصادفی مثل x نمایش داد . با استفاده از متغیر تصافی تفسیر نتایج حاصل از آزمایش های تصادفی ساده تر می شود .

تعریف :

یک متغیر تصادفی ، تابعی است از فضای نمونه به زیر مجموعه ای ناتهی از اعداد حقیقی که آن را با حروف بزرگ مانند X,Y,... نمایش می دهیم .

برای نمایش برد یک متغیر تصادفی یا مجموعه مقادیری که متغیر اختیار می کند ، از حروف کوچک مثل x,y,... استفاده می کنیم .

مثال 1 :

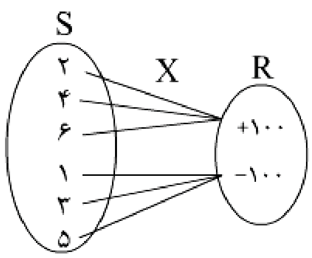
در یک بازی شخصی تاسی را می ریزد ، اگر عدد زوج بیاورد 100 تومان به او می دهیم و اگر عدد فرد بیاورد 100 تومان از او می گیریم . اگرمتغیر تصادفی x در آمد شخص در نظر گرفته شود ، مقادیر آن را بدست آورید .

حل :

ابتدا فضای نمونه را بدست می آوریم .

$$S = \{ ۱ ، ۲ ، ۳ ، ۴ ، ۵ ، ۶ \}$$

همان طور ی می بینید با توجه به نمودار زیر فضای نمونه S تحت تابع x به یک زیر مجموعه از R متشکل از دو عضو 100+ و 100- برده شده است .



عدد 100+ را برای دریافت و 100- را برای پرداخت پول در نظر می گیریم .

$$X : S \rightarrow R$$

$X(۲) = ۱۰۰ \qquad X(۱) = -۱۰۰$

$X(۴) = ۱۰۰ \qquad X(۳) = -۱۰۰$

$X(۶) = ۱۰۰ \qquad X(۵) = -۱۰۰$

از آنجا کع احتمال هر یک از اعداد 1 تا 6 برابر $\frac{1}{6}$ می باشد داریم :

$$A = \{ ۲ ، ۴ ، ۶ \}$$

$$P(X = ۱۰۰) = P(A) = \frac{n(A)}{n(S)} = \frac{۳}{۶} = \frac{۱}{۲}$$

$$B = \{ ۱ ، ۳ ، ۵ \}$$

$$P(X = -۱۰۰) = P(B) = \frac{n(B)}{n(S)} = \frac{۳}{۶} = \frac{۱}{۲}$$

حال می توانیم میانگین درآمد شخص را پس از تکرار بازی به دفعات محاسبه کنیم

X	۱۰۰	-۱۰۰
P	$\frac{۱}{۲}$	$\frac{۱}{۲}$

$$\Rightarrow \bar{x} = \frac{۱۰۰ \times \frac{۱}{۲} + -۱۰۰ \times \frac{۱}{۲}}{\frac{۱}{۲} + \frac{۱}{۲}} = ۰$$

با توجه به $\bar{X} = 0$ می توان نتیجه گرفت که شخص با تکرار بازی پولی به دست نمی آورد .

Jalase 1 – sco 2

انواع متغیر های تصادفی

متغیر های تصادفی عموما بر 2 دسته می باشند :

1. گسسته : متغیر هایی که برد آن یا مجموعه مقادیری که اختیار می کند ، به صورت از هم جدا و یا شمارش پذیر از اعداد حقیقی باشند .
2. مقادیری اختیار می کند ، که در مجموعه ای پیوسته یا شمارش ناپذیر از اعداد حقیقی قرار دارند .

توجه :

همان طور که در فصل قبل اشاره کردیم احتمال وقوع هر عضو منفرد از مجموعه ای پیوسته برابر صفر می باشد . این مطلب برای متغیر های تصادفی پیوسته نیز برقرار می باشد . یعنی احتمال برابر بودن مقدار یک متغیر تصادفی پیوسته x بت هر عنصر منفرد از فضای برد آن برابر صفر می باشد .

مثال 2 :

در یک جعبه 22 مهره موجود می باشد که هر مهره دارای یکی از شماره های 1و 2و 3و4 و5 و6 می باشد . در ضمن به تعداد عدد نوشته شده بر روی هر مهره از همان مهره درون جعبه موجود می باشد . یک مهره از جعبه خارج می کنیم . اگر متغیر تصادفی x شماره مهره خارج شده باشد ، مطلوبست مقادیر احتکالی که متغیر x به خود می گیرد .

حل :

با توجه به صورت مثال یک مهره با شماره یک خواهیم داشت ، 2مهره با شماره 2 و به همین ترتیب 6 مهره با شماره 6 داریم . فضای نمونه مقادیری که x اختیار می کند :

$S_X = \{ ۱ ، ۲ ، ۳ ، ۴ ، ۵ ، ۶ \}$

در نتیجه تابع احتمال به صورت زیر بدست می آید :

X	۱	۲	۳	۴	۵	۶
$P_X(x)$	$\frac{۱}{۲۱}$	$\frac{۲}{۲۱}$	$\frac{۳}{۲۱}$	$\frac{۴}{۲۱}$	$\frac{۵}{۲۱}$	$\frac{۶}{۲۱}$

مثلا $P(x=6)=\frac{6}{21}$ است .

می توانیم به ازای هر مقدار x احتمال آن را به فرم $P_X(x)$ نشان داد که به آن تابع احتمال متغیر تصادفی x گوییم .

تعریف : تابع احتمال متغیر تصادفی x ، تابعی است که از یک متغیر حقیقی x که با $P_X(x)$ نمایش داده می شود:

$$P_X(x) = P(X = x)$$

هر تابع احتمال برای متغیر تصادفی گسسته x باید در شرایط زیر صدق کند .

۱) به ازای هر X , $P_X(x) \geq ۰$

۲) $\sum_{\text{برد } X} P_X(x) = ۱$

Jalase 1 – sco 3

توابع توزیع و چگالی برای متغیر تصادفی پیوسته

اگر $P_X(x)$ تابع احتمال متغیر تصادفی گسسته x باشد ، تابع توزیع $F_X(x)$ به صورت زیر تعریف می شود :

$$\forall x \in R \quad ; \quad F_X(x) = p\,(X \leq x) = \sum_{t \leq x} P_X(t)$$

مثال 4 :

تابع توزیع را برای متغیر تصادفی که در مثال قبل حل شد را بدست آورید .

حل :

با توجه به جدول احتمالات متغیر تصادفی x داریم :

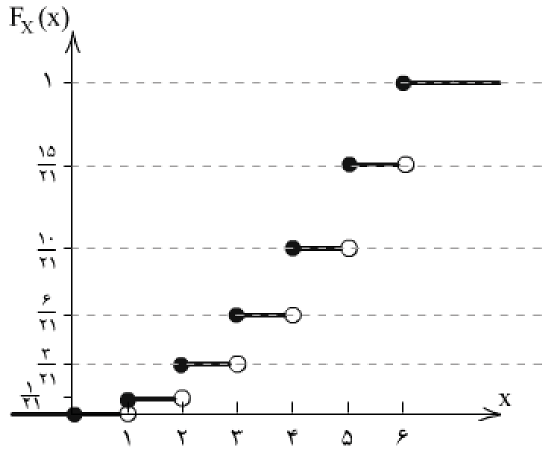
X	۱	۲	۳	۴	۵	۶
$P_X(x)$	$\frac{۱}{۲۱}$	$\frac{۲}{۲۱}$	$\frac{۳}{۲۱}$	$\frac{۴}{۲۱}$	$\frac{۵}{۲۱}$	$\frac{۶}{۲۱}$

$$F_X(۱) = P(X \leq ۱) = \frac{۱}{۲۱}$$

$$F_X(۲/۵) = P(X \leq ۲/۵) = \sum_{t \leq ۲/۵} P_X(t) = P_X(۱) + P_X(۲) = \frac{۳}{۲۱}$$

به ازای هر $x \prec 0$ مقدار تابع توزیع برابر صفر می باشد . داریم :

$$F_X(x) = \begin{cases} ۰ & x < ۰ \\ ۰ & ۰ \leq x < ۱ \\ \frac{۱}{۲۱} & ۱ \leq x < ۲ \\ \frac{۳}{۲۱} & ۲ \leq x < ۳ \\ \frac{۶}{۲۱} & ۳ \leq x < ۴ \\ \frac{۱۰}{۲۱} & ۴ \leq x < ۵ \\ \frac{۱۵}{۲۱} & ۵ \leq x < ۶ \\ ۱ & ۶ \leq x \end{cases}$$



همان طور که مشاهده می شود تعریف تابع توزیع متغیر تصادفی x مشابه تعریف فراوانی تجمعی نسبی تعریف شده در فصل 1 است . مقدار فراوانی تجمعی قبل از 1 برابر 0 از 1 تا نزدیک 2 ، $\frac{1}{21}$ الی آخر است .

Jalase 2 – sco 1

<u>خواص توابع توزیع</u>

توابع توزیع دارای خواص مشترکی می باشند . که با توجه به تعریف آنها بدست می آید . این خصوصیات در مسائل قبل نیز مشاهده می شوند که عبارتند از :

۱)

$$۰ \le F_X(x) \le ۱ \quad ; \quad \forall x \in R$$

۲)

$$\lim_{x \to +\infty} F_X(x) = ۱ \quad ; \quad \lim_{x \to -\infty} F_X(x) = ۰$$

$$F_X(\infty) = ۱ \qquad F_X(-\infty) = ۰$$

۳) توابع توزیع همواره صعودی (غیر اکید) می باشد .

$$\forall a \le b \rightarrow F_X(a) \le F_X(b)$$

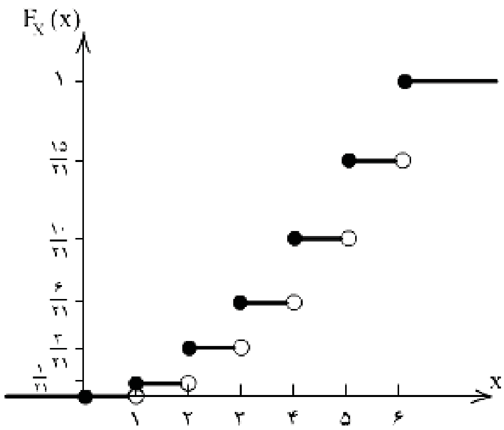
۴) تابع توزیع $F(x)$ در تمام نقاط حداقل از سمت راست ، پیوسته می باشد .

$$\lim_{h \to ۰^+} F_X(x+h) = F_X(x)$$

Jalase 2 – sco 2

همواره از روی یک تابع توزیع متغیر تصادفی می توان مقادیر احتمال را محاسبه نمود .
به عنوان مثال اگر بخواهیم در مثال قبل مقدار $P(x=4)$ را محاسبه کنیم ، با توجه به گسسته بودن متغیر تصادفی x داریم :

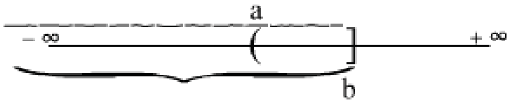
$$F_X(x) = \begin{cases} ۰ & x < ۰ \\ ۰ & ۰ \le x < ۱ \\ \frac{۱}{۲۱} & ۱ \le x < ۲ \\ \frac{۳}{۲۱} & ۲ \le x < ۳ \\ \frac{۶}{۲۱} & ۳ \le x < ۴ \\ \frac{۱۰}{۲۱} & ۴ \le x < ۵ \\ \frac{۱۵}{۲۱} & ۵ \le x < ۶ \\ ۱ & ۶ \le x \end{cases}$$



$$P(X=۴) = \lim_{h \to ۰} P(۴ - h < x \le ۴) \qquad ۱-۳$$

در این صورت $P(x=4)$ را خواهیم داشت . برای محاسبه $P(a \prec x \leq b)$ به صورت زیر عمل می کنیم .

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$$



توجه : برای سادگی در نوشتن می توان اندیس x را از تابع $F_X(x)$ حذف نمود و نوشت $F(x)$. البته اگر به شرطی که بدانیم با متغیر تصادفی x کار می کنیم . بنابراین :

$$P(a < X \leq b) = F(b) - F(a)$$

و داریم :

$$P(X=۴) = \lim_{h \to ۰} P(۴ - h < x \leq ۴) \qquad ۱-۳$$

$$\lim_{h \to ۰} [\, F(۴) - F(۴-h) \,] = F(۴) - \lim_{h \to ۰} F(۴-h)$$

که $\lim\limits_{h\to 0} F(4-h)$ برابر حد چپ تابع F_X در نقطه 4 می باشد و با $F(4^-)$ نمایش می دهیم .

حال با توجه به ضابطه تابع F_X :

$$F(۴^-) = \frac{۶}{۲۱}$$

و داریم :

$$P(X=۴) = F(۴) - F(۴^-) = \frac{۱۰}{۲۱} - \frac{۶}{۲۱} = \frac{۴}{۲۱}$$

درستی حاصل عبارت بالا از روی تابع احتمال $P_X(x)$ نیز آشکار است .

Jalase 2 – sco 3

در حالت کلی می توان رابطه زیر را نوشت :

$$P(X=b) = F(b) - F(b^-)$$

مقدار $P(x=b)$ برابر با میزان جهش نمودار تابع F_X در نقطه b می باشد . همچنین تمامی حالات دیگر را با توجه به فرمول بالا و تعریف تابع توزیع می توان بدست آورد.

۱) $P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$

۲) $P(a \le X < b) = P(X < b) - P(X < a) = F(b^-) - F(a^-)$

۳) $P(a < X < b) = P(X < b) - P(X \le a) = F(b^-) - F(a)$

۴) $P(a \le X \le b) = P(X \le b) - P(X < a) = F(b) - F(a^-)$

۵) $P(X > a) = ۱ - P(X \le a) = ۱ - F(a)$

۶) $P(X < a) = F(a^-)$

حد چپ تابع توزیع در نقطه a

Jalase 2 – sco 4

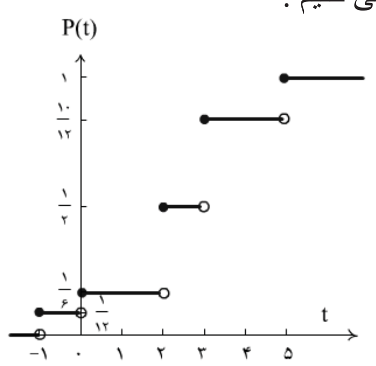
مثال 5 :

تابع توزیع برای متغیر تصادفی گسسته T به صورت زیر است . تابع احتمال آن را بدست بیاورید و نمودار آن را رسم کنید.

$$F(t) = \begin{cases} ۰ & t < -۱ \\ \frac{۱}{۱۲} & -۱ \le t < ۰ \\ \frac{۱}{۶} & ۰ \le t < ۲ \\ \frac{۱}{۲} & ۲ \le t < ۳ \\ \frac{۱۰}{۱۲} & ۳ \le t < ۵ \\ ۱ & ۵ \le t \end{cases}$$

حل:

ابتدا نمودار تابع توزیع را رسم می کنیم .



مقادیری که تابع احتمال $P(T=t)$ می پذیرد ، در نقاط انفصال تابع $F(t)$ رخ می دهد . زیرا در نقاط دیگر مثل T=1 داریم :

$$P(T=١)=P(١\leq T\leq ١)=F(١)-F(١^-)=\frac{١}{۶}-\frac{١}{۶}=٠$$

مقدار P(t) را در نقطه انفصال تابع توزیع F(t) به دست می آوریم :

$$P(T=-١)=F(-١)-F(-١^-)=\frac{١}{١٢}-٠=\frac{١}{١٢}$$

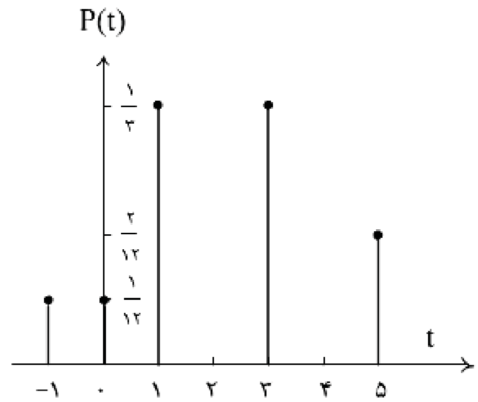
$$P(T=٠)=F(٠)-F(٠^-)=\frac{١}{۶}-\frac{١}{١٢}=\frac{١}{١٢}$$

$$P(T=٢)=F(٢)-F(٢^-)=\frac{١}{٢}-\frac{١}{۶}=\frac{١}{٣}$$

$$P(T=٣)=F(٣)-F(٣^-)=\frac{١٠}{١٢}-\frac{١}{٢}=\frac{۴}{١٢}=\frac{١}{٣}$$

$$P(T=۵)=F(۵)-F(۵^-)=١-\frac{١٠}{١٢}=\frac{٢}{١٢}=\frac{١}{۶}$$

نمودار احتمال P(t) شبیه به نمودار میله ای است .



Jalase 3 – sco 1

توابع چگالی

برای متغیر تصادفی گسسته X تابع چگالی که به فرم $f_X(x)$ نمایش داده می شود به صورت زیر تعریف می شود :

$$f_X(x)=P(X=x) \qquad \text{X عدد حقیقی}$$

به عبارت دیگر تابع چگالی همان تابع احتمال متغیر تصادفی گسسته X می باشد که در رابطه زیر صدق کند :

$$F_X(x)=\sum_{\text{X برد}} f_X(x)$$

در برخی از کتاب ها به جای تابع احتمال برای متغیر های تصادفی گسسته مفهوم تابع چگالی را که دقیقا همان تعریف را دارد به کار می برند .

Jalase 3 sco 2

توابع توزیع و چگالی برای متغیر تصادفی پیوسته

قبل از هر چیز به مثال زیر توجه کنید :

مثال :

فرض کنید مدت زمان توقف اتوبوس در ایستگاه و قبل از حرکت به صورت تصادفی 0 الی 15 دقیقه باشد . در این صورت متغیر تصادفی X را مدت زمان تأخیر حرکت اتوبوس پس از 5 دقیقه تعریف می کنیم . مطلوب است احتمال این که اتوبوس به مدت 5 الی 6 دقیقه تأخیر داشته باشد .

حل :

در این مثال با یک فضای نمونه پیوسته مواجه ایم که می توان آن را به صورت زیر نمایش داد :

$$S = \{ ۵ \leq x \leq ۱۵ \}$$

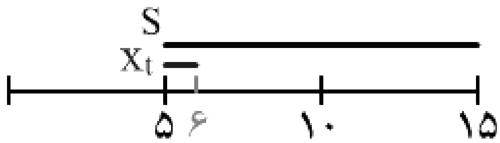
نمایش تابعی متغیر تصادفی X عبارت است از $x_t = t - 5$ و $(t \in S)$ که میزان تأخیر حرکت اتوبوس را نشان می دهد .

به این ترتیب می توان احتمال این که اتوبوس به مدت 5 الی 6 دقیقه تأخیر داشته باشد را محاسبه نمود .

احتمال این که $5 \leq x_t \leq 6$ باشد برابر است با طول بازه ی $[5,6]$ تقسیم بر طول بازه $[5,15]$.

$$\frac{[۵ ، ۶]}{[۵ ، ۱۵]} = ۰/۱$$

لازم به تذکر است که در این جا تابع احتمال یک فاصله پیشامد مورد نظر به طول فاصله فضای نمونه ، تعریف شده است .



حال با توجه به این مثال به تعریف تابع چگالی ، تابع توزیع احتمال برای متغیر تصادفی پیوسته می پردازیم .

فرض کنید x = متغیر تصادفی پیوسته باشد در این حالت نیز تعاریف تابع توزیع و چگالی در مورد متغیر x صدق می کند . با این تفاوت که در این حالت برای بدست آوردن تابع توزیع با توجه به پیوسته بودن از انتگرال گیری به جای جمع بستن روی تابع چگالی استفاده می کنیم . داریم :

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)\,dx$$

که $f_X(x)$ تابع چگالی متغیر تصادفی پیوسته x است .
از این رابطه می توان تابع چگالی را بر حسب تابع توزیع بدست آورد ، به کمک مشتق گیری یعنی

$$f_X(x) = \frac{d}{dx} F_X(x)$$

به عبارتی با مشتق گرفتن از تابع توزیع مقدار چگالی به دست می آید . با توجه به تعاریف می توان خواص زیر را توابع توزیع و چگالی متغیر تصادفی پیوسته x بدست آورد.

Jalase 3 – sco 3

1- برای هر $x \in R$ ، داریم $f_X(x) \geq 0$.

توجه کنید که در حالت پیوسته $f_X(x)$ الزاما کوچکتر یا مساوی یک نمی باشد زیرا همان طور که قبلا اشاره کردیم مقدار تابع احتمال برای متغیر تصادفی در یک نقطه صفر می باشد و احتمالات می بایستی در این حالت در یک بازه محاسبه شوند .

-2

$$\int_{-\infty}^{+\infty} f_X(x)\, dx = ۱$$

برای محاسبه احتمال قرار گرفتن متغیر تصادفی x در یک بازه به این ترتیب عمل می کنیم :

$$P(a < X \leq b) = F(b) - F(a) = \int_{-\infty}^{b} f_X(x)\, dx - \int_{-\infty}^{a} f_X(x)\, dx$$

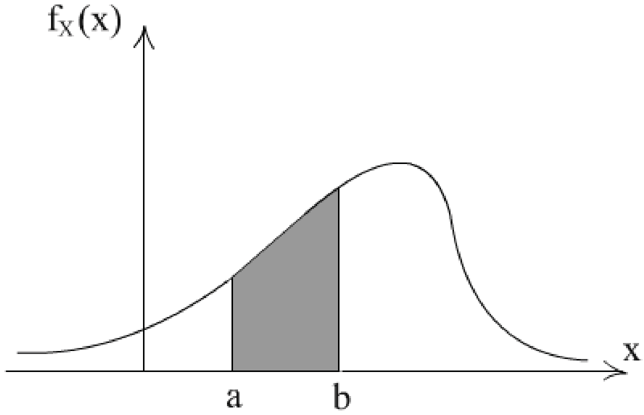
$$= \int_{a}^{b} f_X(x)\, dx$$

از طرفی در حالت پیوسته احتمال این که x برابرعدد دیگری باشد صفر می باشد . بنابراین می توان نوشت:

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$

$$= \int_{a}^{b} f_X(x)\, dx$$

از لحاظ هندسی احتمال قرار گرفتن متغیر تصادفی پیوسته x در بازه a تاb برابر است با سطح زیر منحنی تابع چگالی از a تا b . مطابق شکل زیر :



Jalase 4 – sco 1

مثال :

تابع چگالی احتمال متغیر تصادفی پیوسته x به صورت زیر می باشد :

$$f_X(x) = \begin{cases} x & ۰ \le x \le ۱ \\ c - x & ۱ \le x \le ۲ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

الف) مقدار متغیر c را بدست آورید .

ب) تابع توزیع متغیر تصادفی x را بدست آورید .

ج) احتمال $P(\frac{1}{2} \le x \le \frac{3}{2})$ را یکبار با استفاده از تابع چگالی $f_X(x)$ و یکبار با استفاده از تابع توزیع $F_X(x)$ در نقطه x محاسبه کنید .

حل :

الف) برای بدست آوردن متغیر c می دانیم مساحت زیر منحنی f_X یعنی تابع چگالی می بایستی برابر واحد باشد . بنابراین :

$$\int_x f_X(x)dx = ۱ \quad \Rightarrow \quad \int_۰^۲ f_X(x)dx = \int_۰^۱ x\,dx + \int_۱^۲ (c - x)dx =$$

$$= \frac{x^۲}{۲}\Big|_۰^۱ + (cx - \frac{x^۲}{۲})\Big|_۱^۲ = \frac{۱}{۲} + (۲c - ۲) - c + \frac{۱}{۲}$$

$$= \frac{۱}{۲} + ۲c - ۲ - c + \frac{۱}{۲} = c - ۱$$

$$\Rightarrow \quad c - ۱ = ۱ \quad \Rightarrow \quad c = ۲$$

Jalase 4 – sco 2

ب) مقدار تابع توزیع f_X را محاسبه می کنیم . می دانیم :

$$F_X(x)=\int_{-\infty}^{x} f_X(t)dt$$

برای حل می بایستی این مقدار را در چهار حالت $x \prec 0$ و $0 \le x \prec 1$ و $1 \le x \prec 2$ و $2 \le x$.

اگر $x < ۰ \quad \Rightarrow \quad F_X(x)=\int_{-\infty}^{۰} f_X(t)dt = ۰$

اگر $۰ \le x < ۱ \quad \Rightarrow \quad F_X(x)=\int_{-\infty}^{۰} f_X(t)dt + \int_{۰}^{x} t\,dt$

$$= ۰ + \frac{t^۲}{۲}\Big|_{۰}^{x} = \frac{x^۲}{۲}$$

اگر $۱ \le x < ۲ \quad \Rightarrow F_X(x)=\int_{-\infty}^{۰} f_X(t)dt + \int_{۰}^{۱} f_X(t)dt + \int_{۱}^{x} f_X(t)dt$

$$= ۰ + \int_{۰}^{۱} t\,dt + \int_{۱}^{x} (۲-t)dt =$$

$$= ۰ + \frac{t^۲}{۲}\Big|_{۰}^{۱} + (۲t - \frac{t^۲}{۲})\Big|_{۱}^{x}$$

$$= ۰ + \frac{۱}{۲} + (۲x - \frac{x^۲}{۲}) - (۲ - \frac{۱}{۲}) =$$

$$= ۰ + \frac{۱}{۲} + ۲x - \frac{x^۲}{۲} - ۲ + \frac{۱}{۲} = ۲x - \frac{x^۲}{۲} - ۱$$

اگر $۲ \le x \quad \Rightarrow F_X(x)=\int_{-\infty}^{۰} f_X(t)dt + \int_{۰}^{۱} f_X(t)dt + \int_{۱}^{۲} f_X(t)dt + \int_{۲}^{x} f_X(t)dt =$

$$\Rightarrow F_X(x) = ۰ + \int_{۰}^{۱} t\,dt + \int_{۱}^{۲} (۲-t)dt + \Big|_{۲}^{x} f_X(t)dt$$

$$= ۰ + \frac{t^۲}{۲}\Big|_{۰}^{۱} + (۲t - \frac{t^۲}{۲})\Big|_{۱}^{۲} + ۰$$

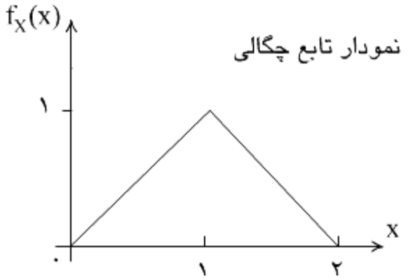
$$= \frac{۱}{۲} + (۴-۲) - (۲ - \frac{۱}{۲}) = \frac{۱}{۲} + \frac{۱}{۲} = ۱$$

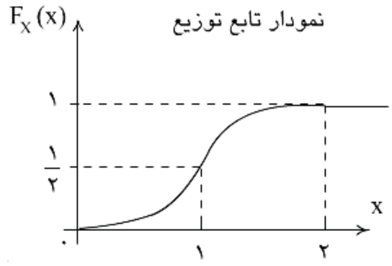
بنابراین ضابطه تابع توزیع $F_X(x)$ برابر است با :

$$F_X(x)=\begin{cases} 0 & x<0 \\ \frac{x^2}{2} & 0\le x<1 \\ 2x-\frac{x^2}{2}-1 & 1\le x<2 \\ 1 & 2\le x \end{cases}$$

حال به نمودار دو تابع چگالی و توزیع توجه کنید :

نمودار تابع چگالی به صورت یک مثلث که قاعده پایینی آن بر روی محور x ها حدفاصل 0 تا 2 است ، می باشد .





Jalase 4 – sco 3

ج) ابتدا مقدار $P(\frac{1}{2}\le x\le\frac{3}{2})$ را با استفاده از تابع چگالی $f_X(x)$ بدست می آوریم .

$$P(\frac{1}{2}\le X\le\frac{3}{2})=\int_{\frac{1}{2}}^{\frac{3}{2}} f_X(t)dt=\int_{\frac{1}{2}}^{1} t\,dt+\int_{1}^{\frac{3}{2}}(2-t)dt$$

و این به خاطر این است که ضابطه تعریف تابع چگالی از 1 تا $\frac{3}{2}$ با $\frac{1}{2}$ تا 1 متفاوت است .

$$=\frac{t^2}{2}\Big|_{\frac{1}{2}}^{1}+(2t-\frac{t^2}{2})\Big|_{1}^{\frac{3}{2}}=\frac{1}{2}-\frac{1}{8}+(3-\frac{9}{8})-(2-\frac{1}{2})$$

$$=2-\frac{10}{8}=\frac{6}{8}=\frac{3}{4}=0/75$$

مجدادا مقدار $P(\frac{1}{2}\le x\le\frac{3}{2})$ را با استفاده از تابع توزیع $F_X(x)$ که در بند ب محاسبه شده است بدست می آوریم.

$$P\left(\frac{۱}{۲} \le X \le \frac{۳}{۲}\right) = F_X\left(\frac{۳}{۲}\right) - F_X\left(\frac{۱}{۲}\right)$$

$$= ۳ - \frac{۹}{۸} - ۱ - \frac{۱}{۸} = ۲ - \frac{۱۰}{۸} = \frac{۳}{۴}$$ تابع توزیع

ملاحظه ی کنید که مقدار احتمال از هر دو راه حل برابر است:

$$= ۲ - \frac{۱۰}{۸} = \frac{۶}{۸} = \frac{۳}{۴} = ۰/۷۵$$ **تابع چگالی**

Jalase4 - sco 4

مثال :

متغیر تصادفی پیوسته x دارای تابع توزیع $F_X(x)$ با ضابطه زیر می باشد :

$$F_X(x) = \begin{cases} ۰ & x < ۰ \\ c - e^{-ax} & x \ge ۰ \quad , \quad a > ۰ \end{cases}$$

الف) مقدار متغیر c را بدست آورید ؟

ب) تعیین کنید متغیر a چه مفادیری می تواند بگیرد ، به شرطی که $F_X(x)$ همچنان تابع توزیع باقی بماند .

ج) تابع چگالی احتمال متغیر تصادفی x را بدست بیاورید .

د) مقادیر احتمال $P(x \le 1)$ و $P(0 \le x \le 2)$ و $P(x \ge \ln 10)$ را بدست آورید .

ه) به ازای چع مقداری از متغیر y مقدار $P(x \le y) = \dfrac{1}{2}$ می باشد ؟

حل :

می دانیم شرط این که تابعی مثل $F_X(x)$ تابع توزیع باشد برابر است با :

۱ – $۰ \le F_X(x) \le ۱ \quad \forall x$

۲ – $\lim\limits_{x \to -\infty} F_X(x) = ۰ \quad , \quad \lim\limits_{x \to +\infty} F_X(x) = ۱$

۳ – $F_X(a) \le F_X(b) \quad \forall a \le b$

۴ – $\lim\limits_{h \to ۰} F_X(x+h) = F_X(x) \quad \forall x$

پس از آن باقی شروط را بررسی می کنیم .

۱ – $\lim\limits_{x \to +\infty} c - e^{-ax} = c - \lim\limits_{x \to +\infty} e^{-ax} = c \rightarrow c = ۱$

اگر $x < ۰$ آنگاه $F_X(x) = ۰$ می باشد ، که در نتیجه شرط برقرار است .

اگر $x \geq ۰$ آنگاه $F_X(x) = ۱ - e^{-ax}$ داریم :

اگر $x \geq ۰ \Rightarrow ۰ \leq e^{-ax} \leq ۱ \rightarrow -۱ \leq e^{-ax} < ۰ \rightarrow ۱ - ۱ \leq e^{-ax} < ۱$

$$\rightarrow ۰ \leq ۱ - e^{-ax} < ۱ \rightarrow ۰ \leq F_X(x) \leq ۱$$

$$\forall x \leq y \Rightarrow e^{-ax} \geq e^{-ay} \rightarrow -e^{-ax} \leq -e^{-ay}$$

$$\Rightarrow ۱ - e^{-ax} \leq ۱ - e^{-ay} \rightarrow F_X(x) \leq F_X(y)$$

Jalase 4 – sco 5

همچنین به ازای هر $x \geq 0$ تابع e^{-ax} پیوسته می باشد . بنابراین تابع $1-e^{-ax}$ نیز حداقل از سمت راست پیوستگی دارد و در نقطه x=0 داریم :

$$\lim_{x \to ۰^+} ۱ - e^{-ax} = F_X(۰) = ۰$$

بنابراین تابع در نقطه 0 پیوسته می باشد . با توجه به برقراری هر 4 شرط می توانیم بگوییم $F_X(x)$ با ضابطه زیر یک تابع توزیع احتمال می باشد .

$$F_X(x) = \begin{cases} ۰ & x < ۰ \\ ۱ - e^{-ax} & x \geq ۰ \end{cases}$$

ب) از آنجایی که تابع $F_X(x)$ مستقل از مقدار a در شرایط تابع توزیع صدق می کند ، بنابراین به ازای همه مقادیر $a \succ 0$ تابع $F_X(x)$ نیز تابع توزیع می باشد .

ج) برای بدست آوردن تابع چگالی $f_X(x)$ کافی است از تابع توزیع مشتق بگیریم به این ترتیب که :

اگر $x < ۰ \qquad F'_X(x) = ۰$

اگر $x \geq ۰ \qquad F'_X(x) = ۰ - (-a)e^{-ax} = ae^{-ax}$

توجه کنید که تابع توزیع در نقطه x=0 پیوسته می باشد . اما تابع در این نقطه مشتق ندارد زیرا مشتق چپ و راست در x=0 با هم برابر نمی باشند .

مشتق چپ در صفر $\Rightarrow \quad x < ۰ \rightarrow F'_X(۰) = ۰$

مشتق راست در صفر $\Rightarrow \quad x \geq ۰ \rightarrow F'_X(۰) = ae^{-a(۰)} = a$

بنابراین تابع چگالی احتمال برابر است با :

$$f_X(x) = \begin{cases} ae^{-ax} & x > ۰ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

Jalase 4 – sco 6

د)

$$P(x \leq ۱) = F_X(۱) = \int_{-\infty}^{۱} f_X(x)dx = ۱ - e^{-a}$$

$$P(۰ \leq x \leq ۲) = F_X(۲) - F_X(۰) = ۱ - e^{-۲a} - (۱ - ۱) = ۱ - e^{-۲a}$$

$$P(x \geq Ln۱۰) = ۱ - P(x < Ln۱۰) = ۱ - F_X(Ln۱۰) =$$

$$= ۱ - ۱ + e^{-aLn۱۰} = e^{Ln۱۰^{-a}} = ۱۰^{-a} = \frac{۱}{۱۰^{a}}$$

ه) به فرم زیر عمل می کنیم :

$$P(X \leq y) = \frac{۱}{۲} \Rightarrow ۱ - e^{-ay} = \frac{۱}{۲}$$

$$\Rightarrow \frac{۱}{۲} = e^{-ay} \rightarrow Ln\frac{۱}{۲} = -ay \rightarrow y = \frac{-۱}{a} Ln\frac{۱}{۲}$$

فصل چهارم

امید ریاضی

۱-۴ با توجه به مطالب فصل اول می‌دانیم که اگر مقادیر $x_۱ \cdots x_n$ را از یک جامعه آماری داشته باشیم میانگین آنها را با $\overline{x}=\frac{\sum_{۱}^{n} f_i x_i}{\sum_{۱}^{n} f_i}$ نمایش می‌دهیم حال فرض کنید مقادیر $x_۱ \cdots x_n$ مقادیری باشند که متغیر تصادفی گسسته X با مقدار احتمال $f(xi)$ بخود می‌گیرد. در این صورت امید ریاضی متغیر تصادفی X را همان میانگین مقادیر $x_۱ \cdots x_n$ با فراوانی $f(xi)$ در نظر می‌گیریم و با $E[X]$ یا μ_x به صورت زیر نمایش می‌دهیم:

$$E[X]=\sum_{\text{برد } X} f(xi)\, x_i$$

توجه کنید که در این حالت $\sum_{\text{برد } X} f_i =۱$ زیرا $f(xi)$ خود تابع احتمال متغیر تصادفی X می‌باشد. با توجه به مطالب فوق می‌توان نتیجه گرفت که امید ریاضی در واقع همان میانگین یا مقدار مورد انتظار متغیر تصادفی X است یعنی اگر تحت شرایط یکسان یک آزمایش را با توجه به مقادیر احتمال متغیر تصادفی X تکرار کنیم انتظار داریم چه مقداری از متغیر تصادفی X را مشاهده کنیم. برای روشن شدن مطلب به مثال زیر توجه کنید:

۲-۴ مثال ۸: یک شرکت بیمه انواع اتومبیل‌ها را بیمه بدنه می‌کند. فرض کنید در طول یک سال ۲۰ درصد افراد هیچگاه تصادف نکنند، ۳۰ درصد افراد تصادفاتی با مجموع هزینه ۱۰۰ هزار تومان برای بیمه، ۴۰ درصد افراد تصادفاتی با مجموع هزینه ۱ میلیون تومان برای بیمه و ۱۰ درصد افراد تصادفاتی با مجموع هزینه ۱۰ میلیون تومان برای بیمه داشته باشند. در این صورت انتظار داریم شرکت بیمه بطور متوسط در طول یک سال چه هزینه‌ای را برای بیمه بدنه اتومبیل‌ها پرداخت کند؟

حل: ابتدا متغیر تصادفی X را احتمال تصادف اتومبیل‌ها و مقادیر آنرا برابر با هزینه تقبل شده توسط شرکت بیمه در نظر می‌گیریم بنابر این داریم:

$$f(x)=\begin{cases} ۰/۲ & x=x_۱=۰ \\ ۰/۲ & x=x_۲=۱۰۰/۰۰۰ \\ ۰/۴ & x=x_۳=۱/۰۰۰/۰۰۰ \\ ۰/۱ & x=x_۴=۱۰/۰۰۰/۰۰۰ \end{cases}$$

حال امید ریاضی متغیر تصادفی X را که برابر است با مقدار متوسط هزینه پرداخت شده توسط شرکت بیمه، محاسبه می‌کنیم:

$$E[X]=\sum_{۱}^{۴} f(xi)\, x_i =۰/۲\times۰+(۱۰۰/۰۰۰)\,۰/۳+(۱/۰۰۰/۰۰۰)\,۰/۴+(۱۰/۰۰۰/۰۰۰)\,۰/۱$$

$$=۳۰/۰۰۰+۴۰۰/۰۰۰+۱/۰۰۰/۰۰۰=۱/۴۳۰/۰۰۰$$

به این ترتیب شرکت بیمه می‌بایستی سالانه مبلغ ۱/۴۳۰/۰۰۰ تومان را به ازای بیمه بدنه اتومبیل‌ها پرداخت کند.

۳-۴ مثال ۹: فرض کنید در مثال قبل شرکت بیمه تعداد ۱۰۰ اتومبیل را بیمه بدنه کرده باشد در این صورت حداقل قیمت پیشنهادی شرکت برای هر اتومبیل چقدر باشد تا شرکت ضرر کند؟

حل: توجه کنید که با توجه به مقدار امید ریاضی متغیر تصادفی X که در مثال قبل درست آمد می‌دانیم شرکت سالانه ۱/۴۳۰/۰۰۰ تومان هزینه می‌کند. با تقسیم این مبلغ بر تعداد اتومبیل‌ها میزان هزینه به ازای هر اتومبیل بدست می‌آید که برابر است با:

میزان هزینه سالانه هر اتومبیل $۱۴/۳۰۰ = \frac{۱/۴۳۰/۰۰۰}{۱۰۰}$

حال برای آنکه شرکت ضرر ندهد می‌بایستی به ازای بیمه بدنه هر اتومبیل حداقل مبلغ ۱۴/۳۰۰ تومان را دریافت کند.

۴-۴ برخی از خواص امیر ریاضی

از آنجا که امید ریاضی بر پایه میانگین $\overline{X}$ تعریف شده است خواص بدست آمده برای میانگین در مورد امید ریاضی نیز برقرار است که عبارتند از:

۱- $E[ax] = aE[X]$ مقدار ثابت A

۲- $E[x+b] = E[X]+b$ مقدار ثابت B

و در حالت کلی: $E[ax+b] = aE[X]+b$

۴-۵ امید ریاضی تابعی از یک متغیر تصادفی

تا بحال امید ریاضی را بر حسب متغیر تصادفی X محاسبه کردیم اما می‌توان آنرا بر حسب تابعی مثل $X^۲$، $۳X-۲$، محاسبه نمود به این ترتیب می‌توان طیف وسیعتری از مسائل را حل نمود. برای محاسبه امید ریاضی بر حسب تابعی مثل $g(X)$ از رابطه زیر استفاده می‌کنیم:

$$E[g(x)] = \sum_{\text{برد } x} g(x)\, f_X(x)$$

به عنوان مثال $E[X]$ در واقع همان $E[g(x)=x]$ می‌باشد.

واریانس و انحراف معیار یک متغیر تصادفی X نیز با قرار دادن $g(X) = (X-\mu_x)^۲$ بدست می‌آید و داریم:

$$\delta_X^۲ = E\left[(X-\mu_x)^۲\right]$$

و انحراف معیار متغیر تصادفی X با جذر از $\delta_X^۲$ بدست می‌آید $\delta_X = \sqrt{\delta_X^۲}$ می‌توان نشان داد که $\delta_X^۲$ از رابطه زیر بدست می‌آید:

$$\delta_X^۲ = E[X^۲] - E^۲[X]$$

اثبات:

$$\delta_X^۲ = E[(X-E[X])^۲] = E\left[X^۲ - ۲XE[X] + E^۲[X]\right]$$
$$= E[X^۲] - E[۲XE[X]] + E^۲[X]$$
$$= E[X^۲] - E[۲E[X]]E[X] + E^۲[X]$$
$$E[X^۲] - E^۲[X]$$

۴-۶ گشتاورها

اگر امید ریاضی را بر حسب تابع $g(X) = (X-a)^k$ محاسبه کنیم به آن گشتاور مرتبه kام حول a گویند به عبارتی:

گشتاور مرتبه k ام حول a $= E[(X-a)^k]$

همچنین اگر $a=\overline{x}$ به آن گشتاور مرتبه kام مرکزی گویند و اگر $a=۰$ باشد به آن گشتاور مرتبه kام حول صفر گویند. که آنرا با m_k نمایش می‌دهند.

گشتاور مرتبه kام مرکزی $= E[(X-\overline{X})^k]$

m_k = حول صفر = kام گشتاور مرتبه $= E[(X^k]$

۴-۷ مثال ۱۰: مقادیر $m_۱$ و $m_۲$ را بر حسب μ_X و $\delta_X^۲$ بدست بیاورید:

حل:

$$m_1 = E\left[X^1\right] = \mu_X$$

$$m_2 = E\left[X^2\right] = \delta_X^2 + \mu_X^2$$

بنابراین با داشتن مقادیر گشتاورهای مرتبه اول و دوم می‌توان مقادیر امید ریاضی و واریانس را بدست آورد. همچنین با داشتن مقادیر سایر گشتاورهای مرتبه سوم به بعد نیز می‌توان اطلاعاتی در مورد متغیر تصادفی X بدست آورد.

۴-۸ تابع مولد گشتاورها

تابع مولد گشتاورها برای متغیر تصادفی X بصورت زیر تعریف می‌شود و با $m_X(t)$ نمایش داده می‌شود:

$$m_X(t) = E\left[e^{tX}\right]$$

$m_X(t)$ را به این دلیل تابع مولد گشتاورها می‌نامیم که با k بار مشتق گیری از آن نسبت به t و قرار دادن $t=0$ مقدار گشتاور مرتبه kام متغیر تصادفی X بدست می‌آید.

به عنوان مثال با یکبار مشتق گیری نسبت به t داریم: (با فرض اینکه بتوان ترتیب اعمال مشتق‌گیری و امیدگیری را جابجا نمود)

$$m_X^{(1)}(t) = \frac{d}{dt} E\left[e^{tX}\right] = E\left[\frac{d}{dt} e^{tX}\right] = E\left[X e^{tX}\right]$$

با قرار دادن $t=0$ داریم:

$$m_X^{(1)}(0) = E\left[X e^{0X}\right] = E[X] = m_1$$

که همان گشتاور مرتبه اول یا μ_X می‌باشد. به همین ترتیب با K بار مشتق‌گیری داریم:

$$m_X^{(k)}(t) = E\left[\frac{d^k}{dt^k} e^{tX}\right] = E\left[X^k e^{tX}\right]$$

و در $t=0$ خواهیم داشت:

$$m_X^{(k)}(0) = E\left[X^k\right]$$

که همان گشتاور مرتبه kام متغیر X می‌باشد.

۴-۹ امید ریاضی برای متغیر تصادفی پیوسته

برای متغیر تصادفی پیوسته X امید ریاضی به جای جمع بستن روی مقادیر $x f_x(x)$ با انتگرال‌گیری روی این مقادیر بدست‌می‌آید بصورت زیر:

$$E[X] = \int_{-\infty}^{+\infty} x f_x(x)\, dx$$

همینطور داریم:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_x(x)\, dx$$

به همین ترتیب برای محاسبه گشتاورها و توابع مولد برای متغیر تصافی پیوسته X از رابطه فوق در فرمول استفاده می‌شود.

۴-۱۰ مثال: فرض کنید برای متغیر تصادفی X مقدار تابع مولد گشتاور برابر با $m_X(t) = \frac{e^t - 1}{t}$ باشد در این صورت مقدار گشتاور مرتبه kام حول مبدأ یا $E[X^K]$ را محاسبه کنید.

حل: می‌دانیم مشتق مرتبه Kام تابع مولد گشتاور نسبت به t در نقطه صفر برابر با $E\left[X^K\right]$ می‌باشد بنابراین:

$$E\left[X^K\right] = \frac{d^k}{dt^k}\left(\frac{e^t - 1}{t}\right)\Big|_{t=0} \Rightarrow \frac{d}{dt}\left(\frac{e^t - 1}{t}\right) = \frac{t e^t - e^t + 1}{t^2}$$

ملاحظه می‌کنید که مشتق مرتبه اول در نقطه $t=0$ موجود نمی‌باشد به همین ترتیب سایر مشتقات مراتب بالاتر نیز در نقطه $t=0$ موجود نمی‌باشند اما این به معنی عدم وجود $E[X^K]$ نمی‌باشد.

در مواقعی که مشتق در نقطه $t=0$ موجود نباشد می‌توان از قضیه زیر برای بدست آوردن گشتاور kام حول مبدأ استفاده کرد.

قضیه: گشتاور مرتبه kام متغیر تصادفی X حول مبدأ برابر است با ضریب $\frac{t^k}{k!}$ در بسط تیلور تابع مولد گشتاور متغیر تصادفی X:

$$m_X(t)=\sum_{k=0}^{\infty} E\left[X^K\right]\frac{t^k}{k!}$$

حال با استفاده از قضیه فوق مجدداً مقدار $E[X^K]$ را در مثال قبل محاسبه می‌کنیم. می‌دانیم: $e^t=\sum_{k=0}^{\infty}\frac{t^k}{k!}$ بنابراین:

$$m_X(t)=\frac{e^t-1}{t}=\frac{\sum_{k=0}^{\infty}\frac{t^k}{k!}-1}{t}=\sum_{k=0}^{\infty}\frac{t^{k-1}}{k!}=\sum_{k=0}^{\infty}\frac{t^k}{(k+1)!}$$

$$=\sum_{k=0}^{\infty}\underbrace{\frac{1}{k+1}}_{E\left[X^K\right]}\frac{t^k}{k!}$$

بنابراین ملاحظه می‌کنید که ضریب $\frac{t^k}{k!}$ در بسط غوق برابر $\frac{1}{k+1}$ می‌باشد یعنی: $E[X^K]=\frac{1}{1+k}$.

۴-۱۱ مثال: تابع مولد گشتاور متغیر تصافی X برابر است با $m_X(t)=\frac{1}{1-t}$ مطلوبست محاسبه μ_X ، δ_X^2 و $E[X^K]$.

حل: در این مثال مقدار مشتق مرتبه Kام $\frac{1}{1-t}$ در نقطه $t=0$ موجود می‌باشد اما محاسبه آن مشکل می‌باشد بنابر این برای راحتی از بسط تیلور استفاده می‌کنیم:

$$m_X(t)=\frac{1}{1-t}=1+t+t^2+t^3+\cdots=\sum_{k=0}^{\infty}t^k=\sum_{k=0}^{\infty}k!\frac{t^k}{k!}$$

بنابراین $E[X^K]=k!$.

حال مقادیر μ_X و δ_X^2 برابرند با:

$$\mu_X=E[X]=1!=1$$
$$\delta_X^2=E\left[X^2\right]-E^2[X]=2!-1=1$$

۴-۱۲ مثال: متغیر تصادفی X دارای تابع چگالی زیر می‌باشد:

$$f_x(x)=\begin{cases}\frac{1}{2}e^{-\frac{x}{2}} & x>0\\ 0 & x\le 0\end{cases}$$

مطلوبست :

الف) محاسبه تابع توزیع متغیر تصادفی X.

ب) محاسبه μ_X ، δ^2_X و گشتاور مرتبه kام حول مبدأ و حول میانگین

حل: الف:

$$F_X(x) = \int_{-\infty}^{x} f_x(t)\,dt = \int_{0}^{x} \frac{1}{2} e^{-\frac{1}{2}}\,dt = -e^{-\frac{1}{2}}\Big|_{0}^{x} = -e^{-\frac{x}{2}} + 1$$

بنابراین:

$$F_X(x) = \begin{cases} 1-e^{-\frac{x}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

ب) ابتدا تابع مولد گشتاور را محاسبه می‌کنیم:

$$m_X(t) = E\left[e^{tx}\right] = \int_{-\infty}^{\infty} e^{tx}\,\frac{1}{2}\,e^{-\frac{x}{2}}\,dx = \int_{0}^{\infty} \frac{1}{2}\,e^{x(t-\frac{1}{2})}\,dx$$

$$= \frac{1}{2(t-\frac{1}{2})}\,e^{x(t-\frac{1}{2})}\Big|_{0}^{\infty} = \frac{1}{2(t-\frac{1}{2})} = \frac{1}{1-2t}$$

به شباهت بین تابع مولد گشتاور در ایم مثال و مثال قبل توجه کنید.

$$\mu_x = E[X] = \frac{d}{dt}\left(\frac{1}{1-2t}\right)\Big|_{t=0} = 2$$

$$\delta^2_X = E[X^2] - E^2[X] = \frac{d^2}{dt^2}\left(\frac{1}{1-2t}\right)\Big|_{t=0}^{-4} = 8-4 = 4$$

برای محاسبه گشتاور مرتبه kام حول میانگین داریم:

$$E[(X-\mu_X)^K] = \text{گشتاور مرتبه kام حول میانگین}$$

برای محاسبه $E[(X-\mu_X)^K]$ از تابع مولد گشتاور حول میانگین که به صورت زیر تعریف می‌شود استفاده می‌کنیم:

$$m_{X-\mu_X}(t) = E[e^{t(X-\mu_X)}]$$

حال توجه کنید که:

$$m_{X-\mu_X}(t) = E\left[e^{tX-t\mu_x}\right] = E\left[e^{tX}e^{-t\mu_x}\right] = e^{-t\mu_x}\,E\left[e^{tX}\right] = e^{-t\mu_x}\,m_X(t)$$

بنابراین رابطه کلی زیر بین گشتاور مرتبه kام حول مبدأ و میانگین موجود است:

$$m_{X-\mu_X}(t) = e^{-t\mu_x}\,m_X(t)$$

حال به راحتی می‌توان $m_{X-\mu_X}(t)$ را با توجه به مقدار $m_X(t)$ بدست آورد داریم:

$$m_{X-\mu_X}(t) = e^{-2t}\,\frac{1}{1-2t}$$

بنابراین مقدار گشتاور مرتبه kام حول میانگین برابر است با مشتق مرتبه kام $m_{X-\mu_X}(t)$ نسبت به t در نقطه $t = 0$.

مقدار $E[X^K]$ برابر است با:

$$m_X(t)=\frac{1}{1-2t}=1+2t+(2t)^2+(2t)^2=\sum_{k=0}^{\infty}(2t)^k$$

$$=\sum_{k=0}^{\infty}2^k k!\left(\frac{t^k}{k!}\right)\quad\Rightarrow\quad E[X^K]=2^K K!$$

۴-۱۳ مثال: تابع توزیع متغیر تصادفی X بصورت زیر داده شده است:

$$f(x)=\begin{cases}0 & x<0\\ \frac{1}{2}x^2 & 0\le x<1\\ -1+2x-\frac{1}{2}x^2 & 1\le x<2\\ 1 & x\ge 2\end{cases}$$

مطلوبست:

الف) تابع چگالی متغیر تصادفی X.

ب) μ_X و δ_X^2

حل: الف)

$$f_x(x)=\frac{d}{dx}f(x)\quad\Rightarrow\quad \text{اگر}\ \ x<0\quad f(x)=0$$

$$\text{اگر}\ \ 0\le x<1\quad\Rightarrow\quad f(x)=2-x$$

$$x\ge 2\quad\Rightarrow\quad f(x)=0$$

بنابراین:

$$f(x)=\begin{cases}x & 0\le x<1\\ 2-x & 1\le x<2\\ 0 & \text{سایر مقادیر}\end{cases}$$

ب)

$$E[X]=\int_{-\infty}^{\infty}f(x)\,dx=\int_0^1 x^2\,dx+\int_1^2 x(2-x)\,dx=$$

$$\frac{x^3}{3}\Big|_0^1+\left(x^2-\frac{x^3}{3}\right)\Big|_1^2=\frac{1}{3}+\left(4-\frac{8}{4}\right)-\left(1-\frac{1}{3}\right)=1$$

$$E\left[X^2\right]=\int_{-\infty}^{\infty}x^2 f(x)\,dx=\int_0^1 x^3\,dx+\int_1^2 x^2(2-x)\,dx=$$

$$\frac{x^4}{4}\Big|_0^1+\left(\frac{2x^3}{3}-\frac{x^4}{4}\right)\Big|_1^2=\frac{1}{4}+\left(\frac{16}{3}-\frac{16}{4}\right)-\left(\frac{2}{3}-\frac{1}{4}\right)=\frac{7}{6}$$

$$\delta_X^2=E\left[X^2\right]-E^2[X]=\frac{7}{6}-1=\frac{1}{6}$$

فصل پنجم

۵-۱ متغیرهای تصادفی دو و چند بعدی

فرض کنید دو متغیر تصادفی X و Y را داشته باشیم می‌دانیم هر کدام از دو متغیر به هر عضو فضای نمونه S مقداری حقیقی و منحصر به فرد نسبت می‌دهند همینطور به ازای هر کدام از X و Y یک تابع چگالی احتمال $f_X(x)$ و $f_Y(y)$ خواهیم داشت.

اگر بخواهیم احتمال وقوع همزمان مقادیر برد هر یک از متغیرهای X و y را بصورت یک تابع احتمال نشان دهیم، متغیر تصادفی $Z=(X,Y)$ را متغیر تصادفی دو بعدی با تابع احتمال $f_{X,Y}(x,y)$ معرفی می‌کنیم.

متغیر تصادفی دو بعدی تمامی خواص متغیرهای یک بعدی را دارا می‌باشد همینطور خواص زیر برای تابع چگالی احتمال آن برقرار است:

$$f_{X,Y}(x,y) = P(X=x \ , \ Y=y)$$

برای متغیر تصادفی دو بعدی گسسته:

$$۱- \quad ۰ \le f_{X,Y}(x,y) \le ۱ \quad \forall x,y$$

$$۲- \quad \sum_{\text{برد } Y} \ \sum_{\text{برد } X} f_{X,Y}(x,y) = ۱$$

برای متغیر تصادفی دو بعدی پیوسته:

$$۱- \quad f_{X,Y}(x,y) \ge ۰ \quad \forall x,y$$

$$۲- \quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y) = ۱$$

$$۳- \quad P(a \le x \le b \ , \ c \le y \le d) = \int_c^d \int_a^b f_{X,Y}(x,y)\, dx\, dy$$

برای روشن شدن مطلب به مثال زیر توجه کنید:

مثال ۱: یک عدد تاس را که بر روی سه وجه آن عدد ۱ و بر روی سه وجه دیگر عدد ۲ حک شده است را دو بار پرتاب می‌کنیم و متغیرهای X و y را بصورت زیر تعریف می‌کنیم:

$X =$ مجموع دو عدد ظاهر شده $= \{۲,۳,۴\}$

$Y =$ تفاضل دو عدد ظاهر شده $= \{-۱,۰,۱\}$

بنابراین متغیر تصادفی دو بعدی (X,Y) را می‌توان به این ترتیب تعریف نمود:

زوج مرتب نمایش دهنده مجموع و تفاضل دو عدد ظاهر $=(X,Y)$ شده در دو بار پرتاب تاس.

چند عضو فضای نمونه S عبارتند از: $= \{(۲,-۱) \ , (۲,۰) \ , (۲,۱) \ , (۲,-۱) \ , \ \dots\}$

به این ترتیب تابع احتمال دو بعدی $f_{X,Y}(x,y)$ بصورت زیر می‌باشد:

۵-۳

Y \ X	۲	۳	۴
-۱	۰	$\frac{۱}{۴}$	۰
۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
۱	۰	$\frac{۱}{۴}$	۰

به عنوان مثال مقادیر احتمال بصورت زیر محاسبه می‌شود:

$$f_{X,Y}(x=۲ \ , \ y=-۱) = ۰$$

برای اینکه مجموع دو عدد ظاهر شده برابر با عدد ۲ باشد باید بر روی تاس در پرتاب اول عدد ۱ و در پرتاب دوم عدد ۱ ظاهر شده باشد. بنابراین احتمال این حالت برابر صفر می‌باشد.

اما برای حالت $(۲,۰)$ این احتمال برابر است با:

$$f_{X,Y}(x=۲\ ,\ y=۰) = \frac{۱}{۲} \times \frac{۱}{۲} = \frac{۱}{۴}$$

برای آنکه مجموع ۲ باشد می‌بایستی در پرتاب اول عدد ۱ و در پرتاب دوم عدد ۱ ظاهر شده باشد بنابراین احتمال اینکه در پرتاب اول عدد ظاهر شود برابر $\frac{۱}{۲}$ می‌باشد و چون ظاهر شدن عدد ۱ در پرتاب دوم مستقل از پرتاب اول است بنابراین احتمال ظاهر شدن ۱ در هر پرتاب برابر است با:

$$\frac{۱}{۲} \times \frac{۱}{۲} = \frac{۱}{۴}$$

از روی تابع چگالی احتمال دو بعدی در این مثال می‌توان مقدار تابع چگالی احتمال متغیر تصادفی X را مستقل از Y محاسبه کنیم که آنرا تابع احتمال حاشیه‌ای یا کناری X می‌نامیم. که جمع بستن روی مقادیر Y در هر سطر بدست می‌آید به عبارتی:

$$f_X(x) \sum_{y} f_{X,Y}(x,y)$$

Y \ X	۲	۳	۴	$f_Y(y)$
−۱	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$	$\frac{۱}{۲}$
۱	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
$f_X(x)$	$\frac{۱}{۴}$	$\frac{۱}{۲}$	$\frac{۱}{۴}$	

به همین ترتیب تابع چگالی احتمال Y نیز با جمع بستن روی مقادیر X در جدول بدست می‌آید که به آن تابع احتمال حاشیه‌ای برای Y گویند.

$$f_Y(y) \sum_{x} f_{X,Y}(x,y)$$

۴–۵ برای نمونه $f_X(x)$ بصورت زیر بدست می‌آید:

$$f_X(۲) = f(۲,-۱) + f(۲,۰) + f(۲,۱) = ۰ + \frac{۱}{۴} + ۰ = \frac{۱}{۴}$$

$$f_X(۳) = f(۳,-۱) + f(۳,۰) + f(۳,۱) = \frac{۱}{۴} + ۰ + \frac{۱}{۴} = \frac{۱}{۲}$$

$$f_X(۴) = f(۴,-۱) + f(۴,۰) + f(۴,۱) = ۰ + \frac{۱}{۴} + ۰ = \frac{۱}{۴}$$

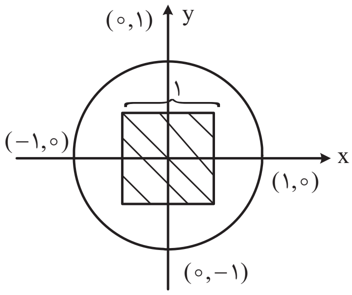
بنابراین:

$$f_X(x) = \begin{cases} \frac{۱}{۴} & x=۲ \\ \frac{۱}{۲} & x=۳ \\ \frac{۱}{۴} & x=۴ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

صحت مقادیر $f_X(x)$ را با استفاده از تعریف متغیر تصادفی X نیز می‌توان بررسی نمود مثلاً $f_X(۳)$ یعنی عدد ظاهر شده در پرتاب اول تاس ۱ و در پرتاب دوم ۲ بوده است یا عدد ظاهر شده در پرتاب اول ۲ و در پرتاب دوم ۱ بوده است که احتمال آن برابر است با:

$$f_X(۲) = \frac{۱}{۲} \times \frac{۱}{۲} + \frac{۱}{۲} \times \frac{۱}{۲} = \frac{۱}{۲}$$

۵–۵ مثال ۲: از درون دایره واحد یک نقطه به تصادف انتخاب می‌کنیم متغیر تصادفی دو بعدی (X, Y) را طول و عرض نقطه انتخاب شده در نظر می‌گیریم مطلوبست محاسبه تابع چگالی احتمال دو بعدی $f_{X,Y}(x, y)$ و با استفاده از آن احتمال اینکه نقطه انتخاب شده درون مربع واحد قرار بگیرد را محاسبه کنید. همچنین تابع چگالی احتمال حاشیه‌ای برای X و Y را محاسبه کنید.



حل: با توجه به شکل روبرو مقدار تابع چگالی را بصورت زیر در نظر می‌گیریم:

$$f_{X,Y}(x,y) = \begin{cases} c & x^{2}+y^{2} \le 1 \\ 0 & \text{سایر مقادیر} \end{cases}$$

حال مقدار مجهول C را محاسبه می‌کنیم:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx\, dy = 1$$

$$\Rightarrow \int_{0}^{1} \int_{0}^{2\pi} c\, r\, d\theta\, dr \quad \Rightarrow c\pi = 1 \quad \rightarrow \quad c = \frac{1}{\pi}$$

بنابراین تابع چگالی احتمال دو متغیر $f_{X,Y}(x, y)$ برابر است با:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & x^{2}+y^{2} \le 1 \\ 0 & \text{سایر مقادیر} \end{cases}$$

برای آنکه نقطه انتخابی درون مربع واحد باشد باید داشته باشیم:

$$-\frac{1}{2} \le +x \le \frac{1}{2} \quad , \quad -\frac{1}{2} \le y \le \frac{1}{2}$$

مقدار احتمال با انتگرال‌گیری روی $f_{X,Y}(x, y)$ در بازه فوق بدست می‌آید:

$$p(-\frac{1}{2} \le x \le \frac{1}{2} \ , \ \frac{1}{2} \le y \le \frac{1}{2}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\pi}\, dx\, dy = \frac{1}{\pi}$$

توجه کنید که مقدار $\frac{1}{\pi}$ برابر با مساحت مربع تقسیم بر مساحت کل دایره می‌باشد که در فصل دوم نیز به این روش محاسبه می‌شد.

۶–۵ برای محاسبه تابع چگالی احتمال حاشیه‌ای برای متغیر X در حالت پیوسته به جای جمع بستن روی مقادیر Y از انتگرال‌گیری استفاده می‌کنیم به عبارتی:

$$f_{X,Y}(x,y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy = \int_{-\sqrt{1-x^{2}}}^{\sqrt{1-x^{2}}} \frac{1}{\pi}(y) \Bigg|_{-\sqrt{1-x^{2}}}^{\sqrt{1-x^{2}}}$$

$$= \frac{1}{\pi}\left(\sqrt{1-x^{2}} - (-\sqrt{1-x^{2}})\right) = \frac{2\sqrt{1-x^{2}}}{\pi}$$

به همین ترتیب مقدار $f_Y(y)$ نیز بدست می‌آید:

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}(x)\ \Big|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}$$

$$= \frac{1}{\pi}\left(\sqrt{1-y^2} - (-\sqrt{1-y^2})\right) = \frac{2\sqrt{1-y^2}}{\pi}$$

بنابراین خواهیم داشت:

$$f_X(x) = \begin{cases} \dfrac{2\sqrt{1-x^2}}{\pi} & -1 \le x \le 1 \\ 0 & \text{سایر مقادیر} \end{cases}$$

$$f_Y(y) = \begin{cases} \dfrac{2\sqrt{1-y^2}}{\pi} & -1 \le y \le 1 \\ 0 & \text{سایر مقادیر} \end{cases}$$

۵-۷ ۱.۱.۴ تابع توزیع دو متغیره

تابع توزیع دو متغیره نیز کاملاً مشابه حالت یک متغیه بدست می‌آید به این ترتیب که:

$$F_{X,Y}(t_1, t_2) = p(x \le t_1,\ y \le t_2) \qquad \forall\ (t_1, t_2) \in R^2$$

حال اگر X و Y گسسته باشند تابع توزیع با جمع بستن روی X و Y تا (t_1, t_2) بدست می‌آید و اگر X و Y پیوسته باشند با انتگرال‌گیری تا (t_1, t_2).

حل: با توجه به تعریف داریم:

Y \ X	۲	۳	۴
-1	0	$\frac{1}{4}$	$\frac{1}{4}$
0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$	1

با توجه به تابع توزیع می‌توان مقدار

احتمال اینکه مجموع دو عدد ظاهر شده کمتر از ۴ و تفاضل دو عدد ظاهر شده کمتر از ۰ باشد را بدست آورد که برابر است با $f_{X,Y}(4,0)$:

$$p(x \le 4,\ y \le 0) = F_{X,Y}(4,0) = \frac{3}{4}$$

۵-۸ ۲.۱.۴ امید ریاضی و گشتاورها برای توابع چگالی دو متغیره

همانند توابع چگالی یک متغیره امید ریاضی و توابع مولد گشتاورهای توام X و Y به صورت زیر محاسبه می‌شوند:

برای متغیرهای گسسته X و Y:

$$E\left[g(X,Y)\right] = \sum_{\text{برد } Y} \sum_{\text{برد } X} g(X,Y)\, f_{X,Y}(x,y)$$

برای متغیرهای پیوسته X و Y:

$$E\left[g(X,Y)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(X,Y)\, f_{X,Y}(x,y)\,dx\,dy$$

تابع مولد گشتاورهای متغیر تصادفی (X,Y):

$$m_{X,Y}(t_1, t_2) = E\left[e^{t_1 X + t_2 Y}\right]$$

توجه کنید که در این حالت نیز با گرفتن مشتق‌های پاره‌ای از تابع مولد گشتاور (X,Y) نسبت به t_1, t_2 و قرار دادن $(t_1, t_2) = (0,0)$ مقدار i jامین گشتاور توام (X,Y) بدست می‌آید. بصورت زیر:

$$\frac{\partial}{\partial t_1} m_{X,Y}(t_1,t_2) = \frac{\partial}{\partial t_1} E\left[e^{t_1 X + t_2 Y}\right] = E\left[x e^{t_1 X + t_2 Y}\right]$$

به همین ترتیب:

$$\frac{\partial}{\partial t_2} m_{X,Y}(t_1,t_2) = E\left[Y e^{t_1 X + t_2 Y}\right]$$

با ادامه دادن روند فوق در نهایت داریم:

$$\frac{\partial^{i+j}}{\partial_{t_1}^{i}\ \partial_{t_2}^{j}} m_{X,Y}(t_1,t_2) = E\left[X^i\ Y^j\ e^{t_1 X + t_2 Y}\right]$$

با قرار دادن $(t_1,t_2)=(0,0)$ گشتاورهای (X,Y) بدست می‌آیند که عبارتند از:

$$m_{ij} = E\left[X^i\ Y^J\right] \qquad i,j=0,1,2,\ldots$$

(i و j هر دو با هم صفر نمی‌باشند)

توجه کنید که m_{0j} و m_{i0} همان گشتاورهای متغیرهای تصادفی X و Y به تنهایی می‌باشند یعنی:

$$m_{i0} = E[X^i]$$

$$m_{0j} = E[y^j]$$

همینطور:

$$m_{11} = E[X,Y]$$

$$m_{10} = E[X] \qquad m_{02} = E[Y^2]$$

$$m_{01} = E[Y] \qquad m_{21} = E[X^2 Y],\ \ldots$$

۵-۹ مثال ۴: برای مثال ۲ مقدار مورد انتظار برای X ، Y و مجموع طول و عرض مشاهده شده را بدست بیاورید:

حل: طبق تعریف $E[X]$ بصورت زیر محاسبه می‌شود:

$$E[X] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\, f_{X,Y}(x,y)\, dx\, dy = \int_0^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} x\frac{1}{\pi} dy\, dx$$

$$= \int_0^1 \int_0^{2\pi} \frac{1}{\pi} r^2 \cos\theta\, d\theta\, dr = \frac{r^3}{3}\Big|_0^1\ \sin\theta \int_0^{2\pi} = \frac{1}{\pi}\frac{1}{3}(0-0)=0$$

البته از آنجا که بازه انتگرال‌گیری متقارن بوده و تابع نیز فرد می‌باشد و می‌توانستیم بدون محاسبه انتگرال نیز صفر بودن جواب را بدست بیاوریم.

مقدار $E[X]$ نیز همانند $E[X]$ بصورت زیر بدست می‌آید:

$$E[X] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\, f_{X,Y}(x,y)\, dx\, dy = \int_0^1 \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} y\frac{1}{\pi} dy\, dx = 0$$

باز هم به علت تقارن و فرد بودن تابع انتگرال فوق صفر می‌باشد.

۵-۱۰ حال مقدار $E[X+Y]$ را محاسبه می‌کنیم:

$$E[X+Y] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (X+Y)\, f_{X,Y}(x,y)\, dx\, dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x\, f_{X,Y}(x,y) + y\, f_{X,Y}(x,y)]\, dx\, dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\, f_{X,Y}\, dy = E[X]+E[Y] = 0+0=0$$

توجه کنید که از مثال فوق می‌توان به این نتیجه رسید که $E[X+Y]=E[X]+E[Y]$ و در حالت کلی اگر $G(X)$ و $H(Y)$ به ترتیب توابعی از X و Y باشند داریم:

$$E[G(X)+H(Y)]=E[G(X)]+E[H(Y)]$$

همچنین توجه کنید که برای محاسبه $E[G(X)]+H(Y)]$ نیازی به دانستن تابع چگالی توام متغیرهای تصادفی X و Y نمی‌باشد بلکه تنها با داشتن توابع چگالی کناری $f_X(x)$ و $f_Y(y)$ می‌توان آنرا محاسبه نمود.

۵-۱۱ ۲۰۴ کوواریانس یا همپراشی

برای مقایسه میزان وابستگی میان دو متغیر تصادفی X و Y از شاخصی به نام کوواریانس یا همپراشی X و Y استفاده می‌کنیم. که به صورت زیر تعریف می‌شود:

$$\delta_{X,Y}=COV(X,Y)=E[(X-\mu_x)(Y-\mu_y)]$$

توجه کنید که مقدار $COV(X,Y)$ در صورتی مثبت خواهد بود که اگر X بزرگتر از میانگینش باشد آنگاه Y نیز چنین باشد به عبارتی اگر X و Y هر دو هم جهت با یکدیگر افزایش یا کاهش داشته باشند مقدار کوواریانس مثبت خواهد بود

$$X\uparrow\ ,\quad Y\uparrow\ \Rightarrow\ \text{Sign}\quad (cov(X,Y))=+۱$$

$$X\downarrow\ ,\quad Y\downarrow\ \Rightarrow\ \text{Sign}\quad (cov(X,Y))=+۱$$

به همین ترتیب اگر $(X-\mu_x)$ و $(Y-\mu_y)$ با احتمال زیاد دارای علامت مخالف باشند یا به عبارتی X و Y در خلاف جهت هم باشند یعنی با افزایش یکی دیگری کاهش پیدا کند یا بلعکس در این صورت کوواریانس X و Y منفی خواهد بود:

$$X\uparrow\ ,\quad Y\downarrow\ \Rightarrow\ \text{sign}\quad (cov(X,Y))=-۱$$

$$X\downarrow\ ,\quad Y\uparrow\ \Rightarrow\ \text{sign}\quad (cov(X,Y))=-۱$$

مقدار کوواریانس را از رابطه زیر که ساده‌تر می‌باشد نیز می‌توان بدست آورد:

$$cov(X,Y)=E[XY]-E[X]E[Y]$$

$$cov(X,Y)=E[(X-\mu_X)(Y-\mu_Y]=E[XY-\mu_X Y-\mu_X\mu_Y]$$

اثبات:

$$=E[XY]-\mu_X E[Y]-\mu_Y E[X]+\mu_X\mu_Y$$

$$=E[XY]-\mu_X\mu_Y-\mu_Y\mu_X=E[XY]-\mu_X\mu_Y$$

۵-۱۲ با استفاده از رابطه فوق خواص متعددی را می‌توان برای کوواریانس بدست آورد که عبارتند از:

$$۱-\quad cov(X,Y)=var(X)=\delta^۲_X$$

اثبات:

$$cov(X,Y)=E[X.X]-E[X]E[X]=E[X^۲]-E^۲[X]=\delta^۲_X$$

$$۲-\quad cov(X,Y)=cov(Y,X)$$

اثبات:

$$cov(X,Y)=E[XY]-E[X]E[Y]=E[YX]-E[Y]E[X]=cov(Y,X)$$

$$۳-\quad cov(X,Y)=\circ$$ (C مقدار ثابت)

اثبات:

$$cov(X,Y)=E[X.C]-E[X]E[C]=CE[X]-CE[X]=\circ$$

$$۴-cov(ax\pm b,cy\pm d)=ac\ cov(X,Y)$$ (a,b,c,d مقادیر اثبات)

اثبات:

$$cov(ax+b,cy+d)=E[(ax+b)(cy+d)]-E[ax+b]E[cy+d]$$

$$=E[acxy+adx+bcy+bd]-(aE[X]+b)(cE[Y]+d)$$

$$=acE[XY]+ad\cancel{E}[X]+bc\cancel{E}[Y]+bd-acE[X]E[Y]-ad\cancel{E}[X]-bc\cancel{E}[Y]-bd$$

$$ac\,E[XY]-ac\,E[X]\,E[Y]=ac\,(E[XY]-E[X]\,E[Y])=ac\,\mathrm{cov}(X,Y)$$

بنابراین تغیر مبدأ تاثیری روی مقدار کوواریانس ندارد اما تغیر مقیاس بر روی مقدار کوواریانس موثر می‌باشد.

$۵-\ \mathrm{cov}\,(X,Y\pm Z)=\mathrm{cov}(X,Y)\pm\mathrm{cov}(X,Y)$ (متغیرهای تصادفی X , Y , Z)

اثبات:

$$\mathrm{cov}(X,Y\pm Z)=E[X(Y\pm Z)]-E[X]\,E[Y\pm Z]$$
$$=E[XY\pm XZ)]-E[X]\left[E[Y]\pm E[Z]\right]$$
$$=E[XY]\pm E[XZ)]-E[X]\left[E[Y]\pm E[X]\,E[Z]\right]$$
$$=E[XY]-E[X]\,E[Y]\pm(E[XZ)\pm E[X]\,E[Z])$$
$$=\mathrm{cov}(X,Y)\pm\mathrm{cov}(X,Z)$$

۵-۱۳ ۱.۲.۴ ضریب همبستگی خطی بین X و Y

مقدار کوواریانس نیز همانند واریانس به مقیاس متغیرهای تصادفی X و y وابسته است این مطلب را در خواص کوواریانس نیز نشان دادیم. به عبارتی ممکن است برای یک قانون احتمال، مقدار کوواریانس بسیار بیشتر از دیگری باشد، اما نمی‌توان گفت تمایل تغیر کردن X و Y با یکدیگر در حالت اول بیشتر از حالت دوم می‌باشد مشابه ایم مطلب را در فصل اول نیز برای واریانس دو جامعه آماری داشتیم که برای حل آن از ضریب تغیرات استفاده نمودیم. در اینجا برای بدست آوردن معیاری دقیق برای مقایسه میزان تمایل تغیرات همزمان دو متغیر C و Y از ضریب همبستگی خطی استفاده می‌کنیم که بصورت زیر تعریف می‌شود و با $f_{X,Y}$ نمایش داده می‌شود:

$$f_{X,Y}=\mathrm{cov}\left(\frac{X-\mu_X}{\delta_X},\frac{Y-\mu_Y}{\delta_Y}\right)=\frac{\delta_{XY}}{\delta_X\,\delta_Y}$$

همچنین با ساده نمودن رابطه فوق داریم:

$$\mathrm{cov}\left(\frac{X-\mu_X}{\delta_X},\frac{Y-\mu_Y}{\delta_Y}\right)=E\left[\left(\frac{X-\mu_X}{\delta_X}\right)\left(\frac{Y-\mu_Y}{\delta_Y}\right)\right]-E\left[\frac{X-\mu_X}{\delta_X}\right]\ E\left[\frac{Y-\mu_Y}{\delta_Y}\right]$$

با توجه به اینکه متغیر تصادفی $\frac{X-\mu_X}{\delta_X}$ برمال شده می‌باشد بنابراین امید ریاضی آن برابر صفر می‌باشد یعنی: $E\left[\frac{X-\mu_X}{\delta_X}\right]=E\left[\frac{Y-\mu_Y}{\delta_Y}\right]=۰$

بنابراین داریم: $f_{X,Y}=E\left[\left(\frac{X-\mu_X}{\delta_X}\right)\left(\frac{Y-\mu_Y}{\delta_Y}\right)\right]$

۵-۱۴ حال نشان می‌دهیم که مقدار $f_{X,Y}$ همواره بین ۱ و ۱- قرار دارد یعنی $|f_{X,Y}|\le ۱$.

اثبات: دو متغیر تصادفی w و z و مقدار متغیر a را در نظر بگیرید در این صورت متغیر تصادفی $T=(aw-z)^۲$ را تعریف می‌کنیم. به ازای هر a مقادیر متغیر تصادفی T همواره مثبت است بنابراین امید ریاضی آنها نیز مثبت می‌باشد یعنی داریم: $E[(aw-z)^۲]$.

$$\Rightarrow\quad E[a^۲w^۲-۲awz+z^۲]\ge ۰$$

$$\Rightarrow\quad a^۲E[w^۲]-۲aE[wz]+E[z^۲]\ge ۰\qquad ;\forall a$$

حال چون به ازای هر a نامساوی فوق برقرار می‌باشد می‌توان قرار داد $a=\frac{E[wz]}{E[w^۲]}$.

بنابراین:

$$\Rightarrow \frac{E^2[wz]}{E^2[w^2]} E[w^2] - 2\frac{E^2[wz]}{E[w^2]} + E[z^2] \geq 0$$

$$\Rightarrow -\frac{E^2[wz]}{E[w^2]} + E[z^2] \geq 0 \quad \Rightarrow \quad \frac{E^2[wz]}{E[w^2]\,E[z^2]} \leq 1$$

حال به جای متغیرهای تصادفی w و z قرار می‌دهیم: $w = X - \mu_X \ , \ Z = y - \mu_Y$.

داریم:

$$\frac{E^2\left[(X-\mu_X)(Y-\mu_Y)\right]}{E\left[(X-\mu_X)^2\right] E\left[(Y-\mu_Y)^2\right]} \leq 1 \quad \Rightarrow \frac{\delta_{XY}^2}{\delta_X^2\ \delta_Y^2} \leq 1$$

$$\Rightarrow \rho_{XY}^2 \leq 1 \quad \Rightarrow \quad |\rho_{XY}| \leq 1$$

۵-۱۵ توجه کنید که در اثبات فوق اگر حالت تساوی در نامساوی $E[(aw-z)^2]$ رخ دهد داریم: $E[T] = E[(aw-z)^2] = 0$
اما متغیر تصادفی T همواره مثبت است بنابراین میانگین آن تنها زمانی صفر می‌باشد که تمام مقادیری که T قبول می‌کند برابر صفر باشند به عبارت دقیقتر متغیر تصادفی T با احتمال ۱ برابر صفر می‌باشد یعنی:

$$\begin{aligned} p(T=0)=1 \quad &\Rightarrow \quad p((aw-z^2)=0)=1 \\ &\Rightarrow \quad p(aw-z=0)=1 \\ &\Rightarrow \quad p(aw=z) = 1 \\ &\Rightarrow \quad p(Y-\mu_Y = a(X-\mu_X)) = 1 \\ &\Rightarrow \quad p(Y = ax + \mu_Y - a\mu_X) = 1 \end{aligned}$$

۵-۱۶ بنابراین با احتمال صددرصد متغیر تصادفی Y برابر با $ax - \mu_Y - a\mu_X$ می‌باشد و این نشان می‌دهد که از مقادیر قابل استفاده به (X, Y) تنها مقادیری که روی یک خط راست به معادله $Y = ax - \mu_Y - a\mu_X$ قرار دارند می‌توانند احتمال مثبت داشته باشد. به همین دلیل است که می‌گوییم ضریب همبستگی خطی میزان وابستگی خطی بین دو متغیر X و Y را نشان می‌دهد.
هر چه مقدار $P_{X,Y}$ به عدد ۱ یا ۱- نزدیک‌تر باشد، متغیرهٔای X و Y تمایل بیشتری به متغیر مستقیم یا معکوس با یکدیگر خواهند داشت و اگر $P_{X,Y} = 0$ باشد، بدین معنی است که متغیرهای X و Y به صورت خطی به یکدیگر وابسته نمی‌باشند. مثال زیر نشان می‌دهد که ممکن است برای دو متغیر X و Y داشته باشیم $P_{X,Y} = 0$ اما در عین حال دو متغیر کاملاً به یکدیگر وابسته باشند.

۵-۱۷ مثال ۵: متغیر تصادفی X بصورت زیر تعریف شده است:

$$f_X(x) = \begin{cases} \frac{1}{4} & -2 \leq x \leq 2 \\ 0 & \text{سایر مقادیر} \end{cases}$$

متغیر تصادفی Y را برابر X^2 تعریف می‌کنیم در این صورت:

الف) $f_Y(y)$ را محاسبه کنید.

ب) ρ_{XY} را محاسبه کنید.

حل: با توجه به اینکه $Y = X^۲$ داریم:

$$F_Y(y) = p(Y \le y) = p(X^۲ \le y) = p(-\sqrt{y} \le X \le \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

حال می‌بایستی تابع توزیع متغیر تصادفی X را بدست بیاوریم:

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_{-۲}^{x} \frac{۱}{۴}\,dx = \frac{۱}{۴}(x+۲)$$

$$F_X(x) = \begin{cases} ۰ & x \le -۲ \\ \frac{۱}{۴}(x+۲) & -۲ \le x \le ۲ \\ ۱ & x > ۲ \end{cases}$$

بنابراین تابع توزیع Y برابر است با:

$$F_Y(y) = \frac{۱}{۴}(\sqrt{y}+۲) - \frac{۱}{۴}(-\sqrt{y}+۲) = \frac{۱}{۲}\sqrt{y}$$

$$F_Y(y) = \begin{cases} ۰ & y < ۰ \\ \frac{\sqrt{y}}{۲} & ۰ \le y \le ۴ \\ ۱ & ۴ < y \end{cases}$$

با مشتق‌گیری از $F_Y(y)$ تابع چگالی Y را بدست می‌آوریم.

$$F_Y(y) = \begin{cases} \frac{۱}{۴\sqrt{y}} & ۰ \le y \le ۴ \\ ۰ & \text{سایر} \end{cases}$$

۱۸-۵ ب: برای محاسبه ρ_{XY} به مقادیر $E[X]$, $E[Y]$, $E[XY]$ نیاز داریم زیرا :

$$\rho_{XY} = \frac{\delta_{XY}}{\delta_X\,\delta_Y} = \frac{E[XY]-E[X]E[Y]}{\delta_X\,\delta_Y}$$

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \int_{-۲}^{۲} \frac{۱}{۴} x\,dx = ۰$$

به دلیل فرد بودن تابع $x\frac{۱}{۴}$ در بازه $[-۲\,,\,۲]$ مقدار انتگرال صفر می‌باشد.

$$E[Y] = \int_{۰}^{۴} y\,\frac{۱}{۴}\,\frac{۱}{\sqrt{y}}\,dx = \frac{۱}{۴}\left(\frac{۲}{۳}y^{\frac{۳}{۲}}\right)\Big|_{۰}^{۴} = \frac{۸}{۶}$$

$$E[XY] = E[XX^۲] = E[X^۳] = \int_{-۲}^{۲} x^۳\,\frac{۱}{۴}\,dx = ۰$$

به این ترتیب مقدار ρ_{XY} برابر است با:

$$\underbrace{\delta_{XY}}_{\text{کوواریانس}} = \underbrace{۰}_{\text{اول}} - \underbrace{\frac{۸}{۶}\times ۰}_{\text{دوم}} = ۰ \quad \Rightarrow \quad \rho_{XY} = ۰$$

همانطور که ملاحظه می‌کنید مقدار ρ_{XY} صفر می‌باشد و این به معنی عدم وابستگی دو متغیر X و Y نمی‌باشد بلکه به معنی عدم وابستگی خطی آندو است به وضوح X و Y بصورت توانی $(Y = X^۲)$ به یکدیگر وابسته‌اند.

مثال ۶: نشان دهید ضریب همبستگی خطی با اعمال تغیر مبدأ و مقیاس روی متغیرهای تصادفی X و Y بدون تغیر باقی می‌ماند؟

حل: برای این منظور می‌بایستی ثابت کنیم:

$$\rho_{ax+b\,,\,cy+d} = \rho_{XY}$$

اثبات:

$$\rho_{ax+b\,,\,cy+d} = \frac{\operatorname{cov}(ax+b, cy+d)}{\sqrt{\operatorname{var}(ax+b)\ \operatorname{var}(cy+d)}} = \frac{ac\ \operatorname{cov}(X,Y)}{\sqrt{a^{۲}\ \operatorname{var}(X)\ c^{۲}\ \operatorname{var}(Y)}}$$

$$= \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\ \operatorname{var}(Y)}} = \rho_{X,Y}$$

۲۰–۵ مثال ۷: در مثال ۱ مقدار $\operatorname{var}(۳x - ۲y + ۱)$ را محاسبه کنید:

در مثال ۱ جدول احتمالات دو متغیره بصورت زیر بدست آمد:

Y \ X	۲	۳	۴	
−۱	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$	$\frac{۲}{۴}$
۱	۰	$\frac{۱}{۴}$	۰	$\frac{۱}{۴}$
	$\frac{۱}{۴}$	$\frac{۱}{۲}$	$\frac{۱}{۴}$	

می‌دانیم: $\operatorname{var}(X) = \operatorname{cov}(X,X)$

بنابراین می‌توان نوشت:

$$\operatorname{var}(۳X-۲Y+۱) = \operatorname{var}(۳X-۲Y) = \operatorname{cov}(۳X-۲Y\,,\,۳X-۲Y)$$
$$= \operatorname{cov}(۳X,۳X) + \operatorname{cov}(۲Y,۲Y) - ۱۲\operatorname{cov}(X,Y)$$
$$= ۹۷\ \operatorname{var}(X) + ۴۷\operatorname{var}(Y) - ۱۲\operatorname{cov}(X,Y)$$

بنابراین می‌بایستی مقادیر $\delta_X^{۲}\ ,\ \delta_Y^{۲}\ ,\ \delta_{XY}$ را از روی جدول محاسبه کنیم:

$$E[X] = \frac{۱}{۴}\times ۲ + \frac{۱}{۲}\times ۳ + \frac{۱}{۴}\times ۴ = \frac{۱}{۲} + \frac{۳}{۲} + ۱ = ۳$$

$$E[Y] = \frac{۱}{۴}(-۱) + \frac{۲}{۴}\times ۰ + \frac{۱}{۴}۱ = ۰$$

$$E[X^{۲}] = \frac{۱}{۴}\times ۴ + \frac{۱}{۲}\times ۹ + \frac{۱}{۴}\times ۱۶ = ۱ + \frac{۹}{۲} + ۴ = \frac{۱۹}{۲}$$

$$E[Y^{۲}] = \frac{۱}{۴}\times ۱ + \frac{۲}{۴}\times ۰ \times \frac{۱}{۴}\times ۱ = \frac{۱}{۲}$$

$$\operatorname{var}(X) = E[X^{۲}] - E^{۲}[X] = \frac{۱۹}{۲} - ۹ = \frac{۱}{۲}$$

$$\operatorname{var}(Y) = E[Y^{۲}] - E^{۲}[Y] = \frac{۱}{۲} - ۰ = \frac{۱}{۲}$$

$$E[XY] = (۳\times -۱)\times\frac{۱}{۴} + (۳\times ۱)\times\frac{۱}{۴}\times(۴\times ۰)\times\frac{۱}{۴}\times(۳\times ۰)\times\frac{۱}{۴} = ۰$$

$$\Rightarrow \operatorname{cov}(X,Y) = E[XY] - E[X]\,E[Y] = ۰$$

$$\Rightarrow \operatorname{var}(۳X-۲Y+۱) = ۹\times\frac{۱}{۲} + ۴\,\frac{۱}{۲} - ۰ = \frac{۱۳}{۲} = ۶/۵$$

فصل ششم

۶-۱-۱

۱.۶ برخی توابع توزیع گسسته

در فصل قبل با چگونگی بدست آوردن توابع توزیع و چگالی آشنا شدید در این فصل قصد داریم چند تابع چگالی و توزیع خاص را مورد بررسی قرار دهیم. بسیاری از آزمایش‌های تصادفی با وجود اینکه ظاهراً متفاوت می‌باشند اما از ماهیت یکسانی برخوردار می‌باشند و از یک الگو پیروی می‌کنند به همین دلیل است که بررسی توابع چگالی و توزیع آنها مهم می‌باشد.

۱.۱.۶ متغیر تصادفی برنولی و دو جمله‌ای

اگر برای یک آزمایش تصادفی تنها دو نتیجه موفقیت و شکست امکان‌پذیر باشد به آن آزمایش، آزمایش برنولی می‌گوییم و احتمال موفقیت را با p و شکست را با q نمایش می‌دهیم. مثلاً پرتاب یک سکه یک آزمایش برنولی است که دو حالت ممکن را در پی دارد. می‌توانیم آمدن شیر را موفقیت و آمدن خط را شکست در نظر بگیریم. بدیهیست که در این حالت $p=\frac{۱}{۲}$, $q=۱-p=\frac{۱}{۲}$.

۶-۱-۲

یک متغیر تصادفی برنولی عبارتست از تعداد پیروزی (۰ یا ۱ بار) در یک مرتبه انجام آزمایش برنولی. به این ترتیب تابع چگالی برای متغیر تصادفی برنولی X بصورت زیر خواهد بود:

$$f_X(x)=\begin{cases} p & x=۱ \\ q & x=۰ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

متغیر تصادفی برنولی را با نماد $B(۱,p)$ نمایش می‌دهیم که در آن عدد ۱ نمایش دهنده یکبار انجام آزمایش است. تابع چگالی متغیر تصادفی برنولی را بصورت زیر هم می‌توان نوشت:

$$f_{(x)}=p^x\ q^{۱-x} \qquad x=۰\,,۱$$
$$=۰ \qquad \text{سایر مقادیر}$$

حال امید، واریانس و تابع مولد گشتاور متغیر تصادفی برنولی را بدست می‌آوریم:

$$E[X]=۱\times p+۰\times q=p$$
$$E[X^۲]=۱^۲\times p+۰^۲\times q=p$$

پس:

$$\operatorname{var}(x)=E[x^۲]-E^۲[X]=p-p^۲=p(۱-p)=pq$$

به همین ترتیب:

$$m_X(t)=E[e^{tx}]=pe^{t\times ۱}+qe^{t\times ۰}=pe^t+q$$

۶-۲

مثال ۱: شخصی ۵ بلیط می‌فروشد که یکی از آنها قلابی است از وی یک بلیط خریداری می‌نیم، قرار می‌دهیم:

$$X=\begin{cases} ۱ & \text{اگر بلیط اصلی‌را خریداری کرده باشیم} \\ ۰ & \text{اگر بلیط قلابی‌را خریداری کرده باشیم} \end{cases}$$

در این صورت تابع چگالی متغیر تصادفی X را بدست آورید. و $\operatorname{var}[X]\,,\,E[X]$ را بدست آورید.

حل: با توجه به اینکه متغیر تصادفی X دو حالت موفقیت و شکست را نشان می‌دهد بنابراین از متغیر تصادفی برنولی پیروی می‌کند. حالت مقدار احتمال p و q را محاسبه می‌کنیم:

$p=$ احتمال اینکه بلیط اصلی خریداری شود $=\frac{۴}{۵}$

$q=۱-p=\frac{۱}{۵}$

بنابراین تابع چگالی برابر است با:

$$f(x)=\begin{cases}\frac{۴}{۵} & x=۱\\ \frac{۱}{۵} & x=۰\\ ۰ & \text{سایر مقادیر}\end{cases}$$

$E[X]=p=\frac{۴}{۵}$

$var[X]=pq=\frac{۴}{۵}\frac{۱}{۵}=\frac{۴}{۲۵}$

عدد $E[X]=\frac{۴}{۵}$ نشان می‌دهد که اگر از شخصی مورد نظر مثلاً ۱۰۰ عدد بلیط خریداری کنیم باید انتظار داشته باشیم $۸۰=\frac{۴}{۵}\times ۱۰۰$ عدد از بلیط‌ها اصلی باشند.

۳-۶ 2.6 متغیر تصادفی دو جمله‌ای

اگر یک آزمایش برنولی را n بار بطور مستقل انجام دهیم و متغیر تصادفی X را برابر با تعداد پیروزی‌ها در این n بار انجام آزمایش در نظر بگیریم، در این صورت متغیر تصادفی X را دو جمله‌ای می‌نامیم.

برای بدست آوردن تابع چگالی متغیر تصادفی دو جمله‌ای می‌بایستی مقادیری که X قبول می‌کند و مقدار احتمال آنرا بدست بیاوریم. از آنجا که متغیر تصادفی دو جمله‌ای تعداد پیروزی‌ها در n بار از انجام مستقل آزمایش برنولی می‌باشد بنابراین تعداد پیروزی‌ها می‌تواند هر یک از مقادیر ۰ تا n باشد و احتمال اینکه x بار پیروز و در $n-x$ بار آزمایش باقیمانده شکست بخوریم با توجه به استقلال آزمایشها برابر است با:

$$(p.\underbrace{p\ldots p}_{x\ \text{بار}})(q.\underbrace{q\ldots q}_{n-x\ \text{بار}})=p^x q^{n-x}$$

اما این مساله که در کدام یک از n آزمایش پیروز شویم نیز مهم است. با توجه به قواعد شمارشی به $\binom{n}{x}$ صریق می‌توان در n آزمایش پیروز شد

بنابراین تابع چگالی احتمال برای متغیر تصادفی دو جمله‌ای X برابر است با:

$$f_X(x)=\binom{n}{x}p^x q^{n-x} \qquad x=۰,۱,۲,\ldots,n$$
$$=۰ \qquad \text{سایر مقادیر}$$

اگر $X_۱,X_۲,\ldots,X_n$ مجموعه n متغیر تصادفی برنولی باشند در این صورت با توجه به تعریف متغیر تصادفی دو جمله‌ای X عبارتست از:

$$X=X_۱+X_۲+\cdots+X_n$$

متغیر تصادفی دو جمله‌ای را بصورت $X\sim B(n,p)$ نمایش می‌دهیم که شامل دو پارامتر می‌باشد:

$p=$ احتمال پیروزی

$n=$ تعداد آزمایشها

به این تریتب متغیر تصادفی برنولی حالت خاصی از متغیر تصادفی دو جمله‌ای با پارامتر $n=۱$ می‌باشد.

حال به محاسبه مقادیر امید ریاضی، واریانس و تابع مولد گشتاور برای متغیر تصادفی دو جمله‌ای می‌پردازیم.

برای محاسبه $E[X]$ توجه می‌کنیم که متغیر تصادفی دو جمله‌ای X برابر است با مجموع نتایج حاصله از n بار انجام مستقل آزمایش برنولی بنابراین:

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

$$var(X) = VAR(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} var(X_i) = \sum_{i=1}^{n} pq = npq$$

به همین ترتیب:

$$m_X(t) = E\left[e^{tx}\right] - E\left[e^{t(\sum_{i=1}^{n} X_i)}\right] = E\left[e^{tX_1 + tX_2 + \cdots + tX_n}\right]$$

$$= E\left[e^{tX_1} e^{tX_2} \ldots e^{tX_n}\right] = E\left[e^{tx_1}\right] E\left[e^{tx_2}\right] \cdots E\left[e^{tx_n}\right]$$

$$= (pe^t + q)(pe^t + q) \underbrace{\cdots}_{n \text{ بار}} (pe^t + q) = (pe^t + q)^n$$

۵-۶ مثال ۲: اگر در مثال ۱ تعداد ۱۰ بلیط از فروشنده خریداری می‌کنیم و متغیر X را تعداد بلیط‌های اصلی در نظر بگیریم. مطلوبست:

الف) تابع چگالی احتمال متغیر تصادفی X.

ب) رسم نمودار تابع چگالی X.

ج) مقدار مد.

د) امید و واریانس متغیر X.

حل: الف) در این مساله می‌بایستی یک آزمایش برنولی را ۱۰ مرتبه بصورت مستقل تکرار کنیم بنابراین X یک متغیر تصادفی دو جمله‌ای می‌باشد و تابع چگالی آن عبارتست از:

$$X \sim B(۱۰, \frac{۴}{۵})$$

$$f_X(x) = \binom{۱۰}{x} (\frac{۴}{۵})^x (\frac{۱}{۵})^{۱۰-x} \qquad x = ۰, ۱, ۲, \ldots, ۱۰$$

$$= ۰ \qquad \text{سایر مقادیر}$$

حال مقدار احتمال را به ازای $x = ۰, \ldots, ۱۰$ بدست می‌اوریم:

$$f_X(x=۰) = \binom{۱۰}{۰} (\frac{۴}{۵})^۰ (\frac{۱}{۵})^{۱۰} = \frac{۱}{۵^{۱۰}} = ۰$$

$$f_X(x=۱) = \binom{۱۰}{۱} (\frac{۴}{۵})^۱ (\frac{۱}{۵})^۹ = ۱۰ \frac{۴}{۵^{۱۰}} = ۰$$

$$f_X(x=۲) = \binom{۱۰}{۲} (\frac{۴}{۵})^۲ (\frac{۱}{۵})^۸ = ۴۵ \frac{۴^۲}{۵^{۱۰}} = ۰$$

$$f_X(x=۳) = \binom{۱۰}{۳} (\frac{۴}{۵})^۳ (\frac{۱}{۵})^۷ = ۱۲۰ \frac{۴^۳}{۵^{۱۰}} = ۰/۰۰۰۷$$

$$f_X(x=۳) = \binom{۱۰}{۳} (\frac{۴}{۵})^۳ (\frac{۱}{۵})^۷ = ۱۲۰ \frac{۴^۳}{۵^{۱۰}} = ۰/۰۰۰۷$$

$$f_X(x=۴)=\binom{۱۰}{۴}(\frac{۴}{۵})^{۴}(\frac{۱}{۵})^{۶}=۲۱۰\frac{۴^{۴}}{۵^{۱۰}}=۰/۰۰۵۵$$

$$f_X(x=۵)=\binom{۱۰}{۵}(\frac{۴}{۵})^{۵}(\frac{۱}{۵})^{۵}=۲۵۲\frac{۴^{۵}}{۵^{۱۰}}=۰/۰۲۶$$

$$f_X(x=۶)=\binom{۱۰}{۶}(\frac{۴}{۵})^{۶}(\frac{۱}{۵})^{۴}=۲۱۰\frac{۴^{۶}}{۵^{۱۰}}=۰/۰۸۸$$

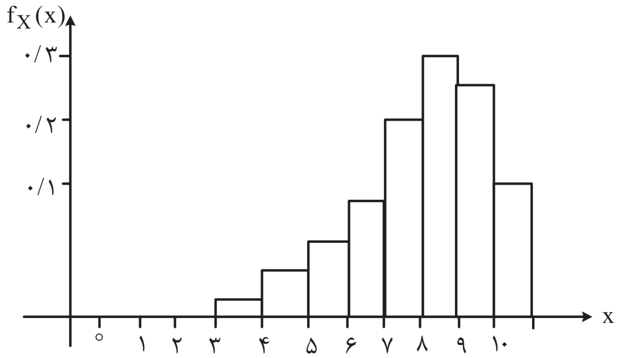
$$f_X(x=۷)=\binom{۱۰}{۷}(\frac{۴}{۵})^{۷}(\frac{۱}{۵})^{۳}=۱۲۰\frac{۴^{۷}}{۵^{۱۰}}=۰/۲$$

$$f_X(x=۸)=\binom{۱۰}{۸}(\frac{۴}{۵})^{۸}(\frac{۱}{۵})^{۲}=۴۵\frac{۴^{۸}}{۵^{۱۰}}=۰/۳$$

$$f_X(x=۹)=\binom{۱۰}{۹}(\frac{۴}{۵})^{۹}(\frac{۱}{۵})^{۱}=۱۰\frac{۴^{۹}}{۵^{۱۰}}=۰/۲۶۸$$

$$f_X(x=۱۰)=\binom{۱۰}{۱۰}(\frac{۴}{۵})^{۱۰}(\frac{۱}{۵})^{۰}=۱\frac{۴^{۱۰}}{۵^{۱۰}}=۰/۱$$

ب) حال نمودار مستطیلی تابع چگالی متغیر تصادفی X بصورت زیر بدست می‌آید:



توجه کنید که $\sum_{x=0}^{۱۰}\binom{۱۰}{x}(\frac{۴}{۵})^{x}(\frac{۱}{۵})^{۱۰-x}=۱$ می‌باشد.

ج) از فصل دوم به یاد دارید که مد مقداری از X است که به ازای آن تابع چگالی ماکزیمم می‌شود. از روی نمودار به وضوح پیداست که مد متغیر تصادفی X برابر با: $x=۸$ می‌باشد. یعنی به ازای هر بار خرید ۱۰ عدد بلیط از فروشنده به احتمال زیاد $(۰/۳)$ ۸ میانه توزیع نیز می‌باشد زیرا $x=۸$ اولین نقطه‌ایست که در آن مقدار تابع توزیع بزرگتر یا مساوی با $\frac{۱}{۲}$ می‌شود یعنی: $F_X(x=۸)\geq\frac{۱}{۲}$.

د)

$$E[X]=np=۱۰\times\frac{۴}{۵}=۸$$

$$var(X)=npq=۱۰\,\frac{۴}{۵}\,\frac{۱}{۵}=\frac{۸}{۵}$$

همانطور که ملاحظه می‌کنید مقدار امید ریاضی X برابر ۸ می‌باشد به همین دلیل است که در مثال ۱ نشان دادیم که اگر ۱۰۰ بلیط از فروشنده دریافت کنیم به طور متوسط $۸۰=\frac{۴}{۵}\times ۱۰۰$ عدد از آنها اصلی می‌باشند.

۱-۷-۶ ۱.۲.۶ مد توزیع دو جمله‌ای

در مثال قبل برای بدست آوردن مد از تابع چگالی استفاده نمودیم اما می‌توان مقدار مد را بدون استفاده از تابع چگالی و تنها با داشتن مقادیر پارامترهای توزیع دو جمله‌ای بدست آورد. می‌دانیم متغیر تصادفی دو جمله‌ای گسسته می‌باشد از طرفی چون تابع چگالی به ازای مقدار مد بیشترین مقدار را قبول می‌کند بنابراین داریم:

$$\begin{cases} f(x_M) \geq f(x_M - ۱) & (۱) \\ f(x_M) \geq f(x_M + ۱) & (۲) \end{cases}$$

$$(۱) \quad \Rightarrow \quad \binom{n}{x_M} p^{x_M} q^{n-x_M} \geq \binom{n}{x_M+۱} p^{x_M+۱} q^{n-(x_M+۱)}$$

$$\Rightarrow \frac{n!}{x_M!(n-x_M)!} \geq \frac{n!}{(x_M+۱)!\,(n-(x_M+۱))!} p q^{-۱}$$

$$x_M + ۱ \geq (n - x_M)\, p q^{-۱} \quad \Rightarrow \quad x_M + x_M\, p q^{-۱} \geq n p q^{-۱} - ۱$$

$$\Rightarrow x_M \geq n p - ۱ + p \quad \Rightarrow \quad x_M \geq p(n+۱) - ۱$$

۶-۷-۲ به همین ترتیب از رابطه (۲) می‌توان نتیجه گرفت:

$$x_M \leq (n+۱)p$$

در نتیجه داریم:

$$p(n+۱) - ۱ \leq x_M \leq p(n+۱)$$

حال اگر $p(n+۱)$ عددی صحیح باشد $p(n+۱)-۱$ نیز عددی صحیح است و هر دو مد متغیر تصادفی X می‌باشند در غیر این صورت بین دو عدد $p(n+۱)$ و $p(n+۱)-۱$ تنها یک عدد صجیح وجود دارد که آن عدد مد توزیع دو جمله‌ای می‌باشد.

مثال ۳: مد تغیر تصادفی دو جمله‌ای در مثال قبل را بدون استفاده از تابع چگالی بدست آورید.

حل: X دارای توزیع دو جمله‌ای با پارامترهای $n=۱۰$, $p=\frac{۴}{۵}$ می‌باشد.

بنابراین:

$$\frac{۴}{۵}(۱۰+۱) - ۱ \leq x_M \leq \frac{۴}{۵}(۱۰+۱)$$

$$۷ + \frac{۱}{۵} \leq x_M \leq ۸ + \frac{۱}{۵} \quad \Rightarrow \quad x_M = ۸$$

۶-۸-۱ ۳.۶ متغیر تصادفی هندسی (نوع اول)

اگر متغیر تصادفی X برابر باشد با تعداد شکستها قبل از رسیدن به اولین پیروزی در انجام آزمایشهای مستقل برنولی، در این صورت به آن متغیر تصادفی هندسی از نوع اول می‌گوییم. به عنوان مثال تعداد دفعات پرتاب یک سکه برای بدست آوردن اولین شیر یا تعداد دفعات شلیک به هدف تا قبل از اولین برخورد گلوله با هدف، نمونه‌هایی از متغیرهای تصادفی هندسی نوع اول می‌باشند.

برای بدست آوردن تابع چگالی متغیر تصادفی هندسی نوع اول توجه می‌کنیم که X مقادیر $۰ , ۱ , \cdots$ را قبول می‌کند و از آنجا که در آخرین آزمایش پیروز می‌شویم بنابراین در x آزمایش قبل شکست خورده‌ایم که به دلیل استقلال آزمایشها، احتمال آن برابر است q^x و برای اینکه در آزمایش بعدی پیروز شویم می‌بایستی احتمال p (پیروزی) را در عدد q^x ضرب کنیم بنابراین:

$$f_X(x) = \begin{cases} p q^x & x = ۰, ۱, \ldots \\ ۰ & \text{سایر مقادیر} \end{cases}$$

حال امید ریاضی، واریانس و تابع مولد گشتاور متغیر تصادفی هندسی نوع اول را محاسبه می‌کنیم:

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tX}\, p q^x = \sum_{x=0}^{\infty} e^{tX}\, p\,(qe^t)^x = \frac{p}{1-qe^t}$$

برای بدست آوردن امید ریاضی و واریانس از تابع مولد گشتاور استفاده می‌کنیم:

$$E[X] = M'_X(t)\,\big|_{t=0} = \frac{pqe^t}{(1-qe^t)^2}\,\big|_{t=0} = \frac{pq}{(1-q)^2} = \frac{q}{p}$$

$$E[X^2] = M''_X(t)\,\big|_{t=0} = \frac{pqe^t(1-qe^t)^2 + 2qe^t\,(1-qe^t)\,pqe^t}{(1-qe^t)^4}\,\big|_{t=0} = \frac{qp+q^2}{p^2}$$

پس:

$$\mathrm{var}(X) = E[X^2] - E^2[X] = \frac{pq+q^2}{p^2} - \frac{q^2}{p^2} = \frac{q(p+q)}{p^2} = \frac{q}{p^2}$$

۹-۶ مثال ۴: شخصی به هدفی شلیک می‌کند. گلوله با احتمال $\frac{3}{5}$ به هدف برخورد می‌کند اگر متغیر تصادفی X را تعداد شلیکها قبل از مورد اصابت قرار دادن هدف در نظر بگیریم. مطلوبست:

الف) تابع چگالی متغیر تصادفی X.

ب) احتمال اینکه در شلیک ۷ام هدف را بزند.

ج) به طور متوسط چند بار شلیک قبل از اینکه هدف مورد اصابت قرار بگیرد لازم است.

د) واریانس و تابع مولد گشتاور را برای متغیر تصادفی X محاسبه کنید.

حل: الف) متغیر تصادفی X هندسی از نوع اول می‌باشد و با پارامتر $p = \frac{3}{5}$ بنابراین تابع چگالی آن بصورت زیر بدست می‌آید:

$$f(x) = \frac{2}{5}\left(\frac{2}{5}\right)^x \qquad x = 0, 1, \ldots$$
$$= 0 \qquad \text{سایر مقادیر}$$

ب) برای اینکه در شلیک ۷ام هدف را بزند می‌بایستی ۶ شلیک ناموفق داشته باشیم بنابراین:

$$f(x=6)\,\frac{3}{5}\left(\frac{2}{5}\right)^6 = 0/002$$

ج) برای بدست آوردن متوسط تعداد شلیکها قبل از زدن بایستی امید ریاضی متغیر تصادفی X را بدست آوریم:

$$E[X] = \frac{q}{p} = \frac{\frac{2}{5}}{\frac{3}{5}} = \frac{2}{3}$$

بنابراین شخص می‌بایستی بیشتر اوقات پس از یک شکست هدف را بزند.

د)

$$\mathrm{var}(X) = \frac{q}{p^2} = \frac{\frac{2}{5}}{\frac{9}{25}} = \frac{10}{9}$$

$$M_X(t) = \frac{p}{1-qe^t} = \frac{\frac{3}{5}}{1-\frac{2}{5}e^t} = \frac{3}{5-2e^t}$$

۱۰-۶ ۱.۳.۶ متغیر تصادفی هندسی (نوع اول)

متغیر تصادفی X را تعداد آزمایشها برای رسیدن به اولین پیروزی در آزمایش‌های مستقل برنولی در نظر می‌گیریم. در این صورت X یک متغیر تصادفی هندسی نوع دوم است.برای بدست آوردن تابع چگالی متغیر تصادفی هندسی نوع دوم ابتدا توجه می‌کنیم که حداقل یک آزمایش برای رسیدن به اولین پیروزی مورد نیاز است، احتمال اینکه در آزمایش xام پیروز شویم برابر است با احتمال اینکه در $x-۱$ آزمایش قبلی شکست خورده و در آزمایش آخر پیروز شویم که اولی با احتمال $q^{x-۱}$ و دومی با احتمال p رخ می‌دهد. نتیجه تابع چگالی بصورت زیر بدست می‌آید:

$$f_X(x) = p\,q^{x-۱} \qquad x = ۱, ۲, \ldots$$
$$= ۰ \qquad \text{سایر مقادیر}$$

توجه کنید که بین متغیر تصادفی هندسی نوع اول و دوم رابطه مستقیمی برقرار است، از آنجا که متغیر تصادفی هندسی نوع دوم تعداد آزمایشها برای رسیدن به اولین پیروزی است بنابراین می‌بایستی در تمام آزمایشها شکست بخوریم غیر آزمایش آخر. اگر Y یک متغیر هندسی نوع دوم باشد و X یک متغیر هندسی نوع اول، رابطه $Y=X+۱$ بین ایندو برقرار می‌باشد. با کمک این رابطه می‌توان به راحتی امید، واریانس و تابع مولد گشتاور متغیر هندسی نوع دوم را از روی نوع اول بدست آورد:

$$E[Y] = E[X+۱] = E[X]+۱ = \frac{q}{P}+۱ = \frac{۱}{p}$$

$$var(Y) = var(X+۱) = var(X) = \frac{q}{p^۲}$$

$$M_Y(t) = E[e^{tY}] = E[e^{tY+t}] = e^t E[e^{tX}] = e^t \frac{p}{۱-qe^t}$$

مثال ۵: متغیر تصادفی Y را تعداد شلیکها برای رسیدن به اولین اصابت گلوله به هدف در مثال قبل در نظر می‌گیریم:

مطلوبست:

الف) تابع چگالی متغیر تصادفی X.

ب) احتمال اینکه در کمتر یا مساوی با دئ آزمایش هدف مورد اصابت گلوله قرار گیرد.

ج) متوسط تعداد شلیکها برای زدن هدف.

د) واریانس و تابع مولد گشتاور متغیر تصادفی Y.

حل: الف) Y یک متغیر تصادفی هندسی نوع دوم است بنابراین تاع چگالی آن بصورت زیر می‌باشد:

$$f_Y(y) = \frac{۳}{۵}\left(\frac{۲}{۵}\right)^{y-۱} \qquad y = ۱, ۲, \ldots$$
$$= ۰ \qquad \text{سایر مقادیر}$$

ب) برای اینکه در کمتر یا مساوی با دو آزمایش هدف مورد اصابت قرار گیرد می‌بایستی احتمال $f(y=۱)+f(y=۲)$ را محاسبه کنیم:

$$f(y=۱)+f(y=۲) = \left(\frac{۳}{۵}\right)\left(\frac{۲}{۵}\right)^۰ + \left(\frac{۳}{۵}\right)\left(\frac{۲}{۵}\right)^۱ = \frac{۳}{۵}\left(۱+\frac{۲}{۵}\right) = \frac{۲۱}{۲۵} = ۰/۸۴$$

ج)

$$E[Y] = \frac{۱}{p} = \frac{۱}{\frac{۳}{۵}} = \frac{۵}{۳} = ۱/۶۶$$

به همین ترتیب:

$$var(Y)=\frac{q}{p^{۲}}=\frac{\frac{۲}{۵}}{\frac{۹}{۲۵}}=\frac{۱۰}{۹}$$

د)

$$m_Y(t)=\frac{\frac{۳}{۵}e^t}{۱-\frac{۲}{۵}e^t}=\frac{۳e^t}{۵-۲e^t}$$

۱-۱۲-۶ ۴.۶ متغیر تصادفی دو جمله‌ای منفی (نوع اول)

اگر متغیر تصادفی X تعداد شکستها قبل از رسیدن به kامین پیروزی در تکرار مستقل آزمایشهای برنولی در نظر گرفته شود به آن متغیر تصادفی دو جمله‌ای منفی از نوع اول می‌گوییم.

توجه کنید که اگر $k=۱$ باشد متغیر تصادفی دو جمله‌ای منفی نوع اول همان متغیر تصادفی هندسی نوع اول می‌باشد. برای بدست آوردن تابع چگالی متغیر تصادفی دو جمله‌ای منفی نوع اول توجه کنید که در آزمایش آخر پیروزی kام رخ می‌دهد.بنابراین در آزمایشهای قبلی $k-۱$ بار پیروزی و x بار شکست رخ داده است یعنی از بین $x+k-۱$ آزمایش می‌بایستی $k-۱$ بار پیروز شویم که احتمال آن $p^{k-۱}$ می‌باشد و احتمال اینکه x بار شکست بخوریم هم مهم است که به $\binom{x+k-۱}{x}$ طریق ممکن است. نهایتاً در آخرین آزمایش هم با احتمال p پیروز می‌شویم که همان kامین پیروزی است. به همین ترتیب با ضرب موارد فوق تابع چگالی بصورت زیر بدست می‌آید:

$$f_X(x)=\binom{x+k-۱}{x}q^x p^k \qquad x=۰,۱,۲,\ldots$$

رابطه‌ای مشابه رابطه بین متغیر تصادفی برنولی و دو جمله‌ای ما بین متغیر تصادفی هندسی نوع اول و متغیر تصادفی دو جمله‌ای منفی نوع اول موجود است. فرض کنید یک آزمایش را که متغیر تصادفی مربوط به آن از متغیر تصادفی هندسی نوع اول می‌باشد، k بار تکرار کنیم و نتایج حاصله را با یکدیگر جمع کنیم در این صورت تعداد شکستها قبل از رسیدن به kامین پیروزی را بدست می‌آوریم که همان متغیر تصادفی دو جمله‌ای منفی نوع اول باشد و $X_۱,X_۲,\ldots X_K$ متغیرهای تصادفی هندسی نوع اول باشند داریم:

$$Y=X_۱+X_۲+\cdots+X_K$$

با استفاده از رابطه فوق امید ریاضی، واریانس و تابع مولد گشتاور متغیر تصادفی دو جمله‌ای منفی نوع اول را محاسبه می‌کنیم:

$$E[X]=E\left[\sum_{i=۱}^{k}X_i\right]=\sum_{i=۱}^{k}E[X_K]=\sum_{i=۱}^{k}\frac{q}{p}=k\frac{q}{p}$$

$$var(X)=var\left(\sum_{i=۱}^{k}X_i\right)=\sum_{i=۱}^{k}var(X_i)=\sum_{i=۱}^{k}\frac{q}{p^{۲}}=k\frac{q}{p^{۲}}$$

$$m_X(t)=E\left[e^{tX}\right]=E\left[e^{t\sum_{۱}^{k}X_i}\right]=E\left[e^{tX_۱}\ e^{tX_۲}\ldots e^{tX_K}\right]$$

$$E\left[e^{tX_۱}\right]E\left[e^{tX_۲}\right]\ldots E\left[e^{tX_K}\right]=$$

$$=\left(\frac{p}{۱-qe^t}\right)\left(\frac{p}{۱-qe^t}\right)\cdots\left(\frac{p}{۱-qe^t}\right)=\left(\frac{p}{۱-qe^t}\right)^k$$

۱۳-۶ مثال ۶: متغیر تصادفی X را تعداد دفعات به خطا رفتن گلوله قبل از اصابت ۳امین بار گلوله به هدف در مثال ۴ در نظر می‌گیریم، مطلوبست:

الف) تابع چگالی متغیر تصادفی X.
ب) احتمال اینکه پس از ۲ بار خطا رفتن گلوله برای سومین بار به هدف بزنیم.
ج) متوسط تعداد خطا رفتن گلوله قبل از اصابت سوم به هدف.
د) واریانس و تابع مولد گشتاور متغیر تصادفی X.

حل: الف) متغیر تصادفی X از تالع چگالی دو جمله‌ای منفی نوع اول پیروی می‌کند بنابراین تابع چگالی آن با $p=\frac{۳}{۵}$ و $k=۳$ برابر است با:

$$f_X(x)=\binom{x+۳-۱}{x}\left(\frac{۲}{۵}\right)^x\left(\frac{۳}{۵}\right)^۳$$

ب):

$$f(x=۲)=\binom{۴}{۲}\left(\frac{۲}{۵}\right)^۲\left(\frac{۳}{۵}\right)^۳=۰/۲$$

ج):

$$E[X]=K\frac{q}{p}=۳\,\frac{\frac{۲}{۵}}{\frac{۳}{۵}}=۲$$

یعنی بطور متوسط پس از ۲ بار خطا رفتن گلوله باید بتوان هدف را مورد اصابت قرار داد.
همچنین:

د)

$$\text{var}(X)=k\frac{q}{p^۲}=۳\,\frac{\frac{۲}{۵}}{\frac{۹}{۲۵}}=\frac{۳۰}{۹}$$

به همین ترتیب:

$$m_X(t)=\left(\frac{p}{۱-qe^t}\right)^k=\left(\frac{\frac{۳}{۵}}{۱-\frac{۲}{۵}e^t}\right)^۳$$

۱۴-۶ ۵.۶ متغیر تصادفی دو جمله‌ای منفی (نوع دوم)

متغیر تصادفی X را برابر با تعدد آزمایشهای لازم برای رسیدن به kامین پیروزی در تکرار مستقل آزمایشهای برنولی در نظر می‌گیریم. در این صورت متغیر تصادفی X را دو جمله‌ای منفی نوع دوم می‌نامین.

برای بدست آوردن تابع چگالی متغیر تصادفی دو جمله‌ای منفی نوع دوم ابتدا توجه کنید که تفاوت اصلی آن با متغیر تصادفی دو جمله‌ای نوع اول در این است که در نوع اول تعداد k بار پیروزی در محاسبه مقادیری که متغیر تصادفی قبول می‌کند در نظر گرفته نمی‌شود بنابراین اگر Y یک متغیر تصادفی دو جمله‌ای منفی نوع دوم باشد و X نوع اول رابطه زیر بین این دو متغیر تصادفی برقرار است:

$$Y=X+K \qquad x=۰,۱,۲,\ldots$$
$$y=k,\ k+۱,\ k+۲,\ldots$$

بنابراین تابع چگالی Y نیز با تغیر X به $Y-K$ در تالع چگالی دو جمله‌ای منفی نوع اول بدست می‌آید:

$$f_X(x)=\binom{x+k-۱}{x}q^x p^k \xrightarrow{x=y-k} f_Y(y)=\binom{y-۱}{y-k}q^{y-k}p^k$$

$$\Rightarrow\ f_Y(y)=\binom{y-۱}{k-۱}q^{y-k}p^k \qquad y=k,k+۱,k+۲,\ldots$$

اگر $X_۱,X_۲,\ldots X_K$ مجموعه k متغیر تصادفی هندسی نوع دوم باشند و Y یک متغیر تصادفی دو جمله‌ای منفی نوع دوم باشد رابطه زیر بین آنها برقرار است:

$$Y = X_1 + X_2 + \cdots + X_K$$

هر متغیر تصادفی X_i برابر است با تعداد آزمایشها برای رسیدن به اولین پیروزی، حال اگر نتایج حاصل از k آزمایش تصادفی هندسی نوع دوم را با یکدیگر جمع کنیم تعداد آزمایشهای لازم برای رسیدن به kامین پیروزی بدست می‌آید. که همان متغیر تصادفی دو جمله‌ای منفی می‌باشد.

حال با توجه به رابطه فوق مقادیر امید ریاضی، واریانس و تابع مولد گشتاور را برای متغیر تصادفی دو جمله‌ای منفی نوع دوم بدست می‌آوریم:

$$E[X] = E\left[\sum_{i=1}^{k} X_k\right] = \sum_{1}^{k} E[X_k] = \sum_{1}^{k} \frac{1}{p} = \frac{k}{p}$$

$$\operatorname{var}(X) = \operatorname{var}\left(\sum_{1}^{k} X_k\right) = \sum_{1}^{k} \operatorname{var}(X_k) = \sum_{1}^{k} \frac{q}{p^2} = K\frac{q}{p^2}$$

$$m_X(t) = E\left[e^{tX}\right] = E\left[e^{t\sum_{1}^{k} X_K}\right] = E\left[e^{tX_1}\, e^{tX_2} \cdots e^{tX_K}\right]$$

$$= E[e^{tX_1}]\, E[e^{tX_2}] \cdots E[e^{tX_K}]$$

$$= \left(\frac{pe^t}{1-qe^t}\right)\left(\frac{pe^t}{1-qe^t}\right) \cdots \left(\frac{pe^t}{1-qe^t}\right) = \left(\frac{pe^t}{1-qe^t}\right)^k$$

مثال ۷: اگر در مثال ۴ Y را تعداد دفعات شلیک گلوله برای اصابت ۳ امین بار گلوله به هدف در نظر بگیریم مطلوبست:

الف) تابع چگالی متغیر تصادفی Y.

ب) احتمال اینکه پس از ۵ بار شلیک برای سومین بار هدف را بزنیم.

ج) متوسط تعداد شلیک برای اصابت سه بار گلوله به هدف.

د) واریانس و تابع مولد گشتاور متغیر تصادفی Y.

حل: الف) متغیر تصادفی Y از توزیع دو جمله‌ای منفی نوع دوم پیروی می‌ند بنابراین تابع چگالی آن بصورت زیر خواهد بود:

$$f_Y(y) = \binom{y-1}{2} \left(\frac{2}{5}\right)^{y-3} \left(\frac{3}{5}\right)^3 \qquad y = 3, 4, 5, \ldots$$

ب) می‌بایستی $f(y=5)$ را بدست بیاوریم:

$$f(y=5) = \binom{5-1}{2} \left(\frac{2}{5}\right)^{5-3} \left(\frac{3}{5}\right)^3 = 0/2$$

همانطور که ملاحظه می‌کنید مقدار $f(y=5)$ با مقدار $f(x=2)$ در مثال قبل برابر می‌باشد زیرا رابطه $Y=X+3$ بین دو متغیر تصادفی X و Y در این دو مثال وجود دارد.

ج):

$$E[Y] = \frac{k}{p} = \frac{3}{\frac{3}{5}} = 5$$

یعنی بطور متوسط از هر ۵ بار شلیک به هدف می‌وان انتظار داشت که ۳ بار گلوله به هدف اصابت کند.

$$\operatorname{var}(Y) = \frac{kq}{p^2} = 3\,\frac{\frac{2}{5}}{\frac{9}{25}} = \frac{30}{9}$$

د):
$$m_X(t)=\left(\frac{pe^t}{1-qe^t}\right)^k=\left(\frac{\frac{3}{5}e^t}{1-\frac{2}{5}e^t}\right)^3$$

۶-۱۷-۱ ۶.۶ متغیر تصافی فوق هندسی

فرض کنید ظرفی حاوی N توپ می‌باشد، که M تای آنها سفید می‌باشند. می‌خواهیم یک نمونه nتایی از این ظرف به تصادف و بدون جاگذاری خارج کنیم متغیر تصادفی X را تعداد توپهای سفیدی که در نمونه nتایی موجود می‌باشند در نظر می‌گیریم. در این صورت X یک متغیر تصادفی فوق هندسی نامیده می‌شود.

توجه کنید که در متغیر تصادفی فوق هندسی هر بار یک عدد توپ از ظرف خارج می‌کنیم که در اولین مرتبه، احتمال سفید بودن توپ $p=\frac{M}{N}$ می‌باشد، اما چون توپها را بدون جاگذاری خارج می‌کنیم در دومین انتخاب احتمال سفید بودن توپ به نتیجه آزمایش اول وابسته است به عبارت دقیقتر در انتخاب m توپ از ظرف، که هر انتخاب یک آزمایش برنولی است، بدلیل انتخاب توپها برون جاگذاری آزمایشها از یکدیگر مستقل نمی‌باشند.

اگر توپها را با جایگذاری انتخاب می‌کردیم در این صورت X یک متغیر تصادفی دو جمله‌ای با پارامترهای n , $p=\frac{M}{N}$ می‌بوده.

برای بدست آوردن تابع چگالی متغیر تصادفی فوق هندسی فرض می‌نیم از n توپ انتخاب شده x تای آنها سفید باشند در این صورت x می‌تواند عددی بین ۰ تا M باشد. انتخاب n توپ از N توپ به $\binom{N}{n}$ طریق ممکن است. همینطور سفید بودن x توپ به $\binom{M}{x}$ طریق و مابقی به $\binom{N-M}{n-x}$ طریق ممکن است در نتیجه احتمال $p(X=x)$ بصورت زیر بدست می‌آید:

$$f_X(x)=p(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

۶-۱۷-۲ از $\binom{M}{x}$ نتیجه می‌شود که $0\le x\le M$ و از $\binom{N-M}{n-x}$ نتیجه می‌شود که:

$$0\le n-x\le N-M \quad\Rightarrow\quad n+M-N\le x\le n$$

نهایتاً بازده زیر برای مقادیری که متغیر تصادفی X قبول می‌کند بدست می‌آید:

$$\begin{cases}0\le x\le M\\ n+M-N\le x\le n\end{cases}\quad\Rightarrow\quad \mathrm{Max}\,(0\,,\,n+M-N)\le x\le \mathrm{Min}\,(M\,,\,n)$$

متغیر تصادفی فوق هندسی X را با نماد $X\sim HG(N,M,n)$ نمایش می‌دهیم.

امید ریاضی و واریانس متغیر تصادفی فوق هندسی برابر است با:

$$E[X]=n\frac{M}{N}$$
$$\mathrm{var}(X)=n\frac{M}{N}\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

۶-۱۸ مثال ۸: در یک چاپخانه از هر ۵۰ برگ چاپ شده ۵ برگ بدلیل کیفیت نامناسب چاپ به دور ریخته می‌شود. برای کنترل کیفیت از هر ۵۰ برگ ۳ برگ کنترل می‌شود و اگر حداقل یک برگ کیفیت نامناسبی داشته باشند آنگاه سایر برگها نیز کنترل می‌شوند. احتمال اینکه مسوول کنترل کیفیت مجبور باشد هر ۵۰ برگ را کنترل کند چقدر می‌باشد؟

حل: متغیر تصادفی X را تعداد برگهای بدون کیفیت در یک نمونه ۳ تایی در نظر می‌گیریم. در این صورت X متغیر تصادفی فوق هندسی با پارامترهای $X\sim HG(50,5,3)$ می‌باشد.

$$p(X=x)=\frac{\binom{۵}{x}\binom{۴۵}{۳-x}}{\binom{۵۰}{۳}} \qquad ۰\le x\le ۳$$

برای اینکه مجبور باشیم هر ۵۰ برگ را کنترل کنیم باید حداقل یک برگ انتخابی کیفیت نامناسبی داشته باشند که احتمال آن برابر است با: $p(x\ge ۱)$.

$$p(X\ge ۱)=۱-p(X=۰)=۱-\frac{\binom{۵}{۰}\binom{۴۵}{۳}}{\binom{۵۰}{۳}}=۱-۰/۷۲۳=۰/۲۷۶$$

متوسط تعداد برگهای با کیفیت نامناسب در نمونه عبارتست از:

$$E[X]=n\frac{M}{N}=۳\,\frac{۵}{۵۰}=۰/۳$$

همچنین:

$$\mathrm{var}(X)=۳\,\frac{۵}{۵۰}\,\left(\frac{۵۰-۵}{۵۰}\right)\left(\frac{۵۰-۳}{۵۰-۱}\right)=۰/۲۵$$

۱۹-۶ ۱.۶.۶ تقریب توزیع فوق هندسی بوسیله توزیع دو جمله‌ای

اگر در متغیر تصادفی فوق هندسی مقدار N در مقایسه با n بسیار بزرگ باشد، آنگاه دیگر انتخاب توپها بدون جاگذاری و با جاگذاری تقریباً معادل یکدیگر می‌باشند به همین ذلیل به جای محاسبه تابع احتمال فوق هندسی می‌توانیم از تابع احتمال دو جمله‌ای استفاده کنیم که در این حالت از پارامترهای n , $p=\frac{M}{N}$ استفاده می‌کنیم. در حالت کلی می‌بایستی شروط $N\to\infty$, $\frac{n}{N}\to ۰$ برقرار باشند تا بتوان از توزیع دو جمله‌ای استفاده نمود.

مثال ۹: در یک شهر از میان ۱۰/۰۰۰ خانوار ۴۰۰۰ خانوار دارای فرزند پسر می‌باشند اگر یک نمونه ۵ تایی انتخاب کنیم احتمال اینکه دارای فرزند پسر نباشند چقدر است؟

حل: چون نسبت $\frac{۵}{۱۰/۰۰۰}$ به صفر میل می‌کند و ۱۰۰۰۰ عدد بزرگی محسوب می‌شود می‌توان از تقریب توزیع فوق هندسی به دو جمله‌ای استفاده نمود:

$$p=\frac{M}{N}=\frac{۴۰۰۰}{۱۰۰۰۰}=\frac{۲}{۵}\quad,\quad q=\frac{۳}{۵}\quad,\quad n=۵\quad\Rightarrow\quad X\sim B(۵,\frac{۲}{۵})$$

پس:

$$p(X=x)=\begin{cases}\binom{۵}{x}\left(\frac{۲}{۵}\right)^x\left(\frac{۳}{۵}\right)^{۵-x} & x=۰,۱,۲,\ldots,۵\\ =۰ & \text{سایر مقادیر}\end{cases}$$

احتمال اینکه فرزند پسری نداشته باشند برابر است با: $p(X=۰)$.

$$p(X=۰)=\binom{۵}{۰}\left(\frac{۲}{۵}\right)^۰\left(\frac{۳}{۵}\right)^۵=۰/۰۷$$

۲۰-۶ ۷.۶ متغیر تصادفی پواسون

در بعضی از آزمایشها تعداد دفعات رخ دادن پیشامدی را می‌توان بدست آورد ولی تعداد دفعات عدم رخ دادن آنرا نمی‌توان بدست آورد مثلاً تعداد دفعاتی که در طول شبانه روز تلفن یک شرکت زنگ می‌زند شمرد ولی تعداد دفعاتی که تلفن زنگ نمی‌زند را نمی‌توان شمرد. همینطور تعداد خودروهای وارد شده به پمپ بنزین در مقایسه با خودروهای عبوری.

متغیر تصادفی X را به این صورت تعریف می‌کنیم:

$X =$ تعداد موفقیتها در یک فاصله پیوسته (زمان ، طول)

پارامتر متغیر تصادفی پواسون λ می‌باشد که برابر است با میانگین تعداد موفقیتها در طول یک بازه پیوسته. برای اینکه X متغیر تصادفی پواسون باشد می‌بایستی شرایط زیر برقرار باشد:

۱- در طول فاصله زمانی بسیار کوتاه احتمال رخ دادن یک موفقیت فقط متناسب با طول بازه زمانی باشد و بستگی به تعداد موفقیتها در خارج از فاصله زمانی نداشته باشد.

۲- احتمال روی دادن بیش از یک موفقیت در یک فاصله زمانی کوتاه تقریباً برابر صفر باشد.

۳- تعداد موفقیتهایی که در یک فاصله زمانی مشخص روی می‌دهد از تعداد موفقیتهایی که در یک فاصله زمانی دیگر رخ می‌دهد مستقل باشد.

توجه کنید که در شروط بالا منظور از فاصله زمانی هر بازه پیوسته مثل طول یا یک ناحیه مشخص می‌باشد.

تابع احتمال متغیر تصادفی پواسون با پارامتر λ بصورت زیر می‌باشد:

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = ۰, ۱, ۲, \ldots$$

$$= ۰ \qquad \text{سایر مقادیر}$$

مقادیر امید ریاضی، واریانس و تابع مولد گشتاور عبارتند از:

$$E[X] = \lambda$$

$$var(X) = \lambda$$

$$m_X(t) = e^{-\lambda(۱-e^t)}$$

معمولاً پارامتر λ به صورت $\lambda = rt$ در نظر گرفته می‌شود که در آن t طول بازه مورد نظر و r نرخ وقوع پیشامد در بازه مربوطه است.

مثال ۱۰: یک تایپیست به طور متوسط در هر ۲ صفحه ۵ غلط تایپی دارد مطلوبست:

الف) احتمال اینکه در یک صفحه اصلاً غلط تایپی نداشته باشد؟

ب) احتمال اینکه حداقل ۳ غلط تایپی در ۲ صفحه داشته باشد؟

حل: الف) در یک صفحه ۲/۵ غلط به طور متوسط موجود می‌باشد بنابراین $\lambda = ۲/۵ \times ۱$ و داریم:

$$f_X(x) = \frac{(۲/۵)^x\ e^{-۲/۵}}{x!}$$

برای اینکه اصلاً غلط تایپی نداشته باشیم بایستی $f(x=۰)$ را بدست بیاوریم:

$$f(x=۰) = \frac{(۲/۵)^۰\ e^{-۲/۵}}{۰!} = e^{-۲/۵} = ۰/۰۸$$

ب) در این حالت $\lambda = ۵$ می‌باشد و داریم:

$$f_X(x) = \frac{۵^x e^{-۵}}{x!}$$

$$p(X \geq ۳) = ۱-(p(X=۰)+p(X=۱)+p(X=۲))$$

$$= ۱-(۰/۰۰۶ + ۰/۰۳ + ۰/۰۷۵) = ۰/۸۸۹$$

۱.۷.۶ مد توزیع پواسون

مد توزیع پواسون همانند توزیع دو جمله‌ای بدست می‌آید یعنی:

$$\begin{cases} f(x_M) \ge f(x_M - ۱) & (۱) \\ f(x_M) \ge f(x_M + ۱) & (۲) \end{cases}$$

از (۱) نتیجه می‌شود:قرار دادن $x_M = \hat{x}$

$$\frac{\lambda^{\hat{x}} e^{-\lambda}}{\hat{x}!} \ge \frac{\lambda^{\hat{x}-۱} e^{-\lambda}}{(\hat{x}-۱)!} \Rightarrow x \le \lambda$$

به همین نسبت از رابطه (۲) بدست می‌آید $x \ge \lambda - ۱$ در نتیجه: $\lambda - ۱ \le x \le \lambda$

حال اگر λ عددی صحیح باشد λ , $\lambda - ۱$ هر دو مد می‌باشند و اگر λ عددی صحیح نباشد بین λ و $\lambda - ۱$ عددی صحیح موجود است که مد می‌باشد.

مثال ۱۱: برای مثال قبل مقدار مد را برای بند الف و ب محاسبه کنید؟

الف) $\lambda = ۲/۵$, $\lambda - ۱ = ۱/۵$ بنابراین مد برابر است با $x_M = ۲$.

ب) $\lambda = ۵$, $\lambda - ۱ = ۴$ بنابراین مد برابر است با ۴ , $x_M = ۵$.

۶ . ۷ . ۲ تقریب توزیع دو جمله‌ای بوسیله توزیع پواسون

فرض کنید در توزیع دو جمله‌ای $n \to \infty$, $p \to ۰$ می‌توان به جای محاسبه احتمال با استفاده از تابع احتمال دو جمله‌ای از تابع احتمال پواسون استفاده نمود. در این حالت توزیع دو جمله‌ای را با توزیع پواسون با پارامتر $\lambda = np$ تقریب می‌زنیم.

توجه کنید که برای بدست آوردن مقدار λ امید ریاضی تابع پواسون را با امید ریاضی تابع احتمال دو جمله‌ای برابر قرار می‌دهیم که در نتیجه $\lambda = np$ خواهد شد.

مثال ۱۲: در طول یک روز تعداد ۱۰۰/۰۰۰ خودرو از جلوی پمپ بنزین عبور می‌کند که تعداد ۱۰۰ خودرو وارد پمپ بنزین می‌شوندو مطلوبست احتمال اینکه در طول روز حداقل ۵۰ خودرو وارد پمپ بنزین شوند؟

حل: با محاسبه p داریم: $p = \frac{۱۰۰}{۱۰۰/۰۰۰} = ۰/۰۰۱$ که عددی بسیار کوچک می‌باشد و چون n مقداری بزرگ می‌باشد می‌توان از تقریب توزیع دو جمله‌ای به پواسون استفاده نمود:

$$\lambda = np = ۱۰۰/۰۰۰ \times ۰/۰۰۱ = ۱۰۰$$

$$f_X(x) = \frac{۱۰۰^x e^{-۱۰۰}}{x!}$$

$$p(X \ge ۵۰) = ۱ - p(X < ۵۰) \sim ۰/۹۹۹$$

۲۴-۶ ۶ . ۸ توزیع یکنواخت

متغیر تصادفی X دارای توزیع یکنواخت می‌باشد اگر احتمال رخ دادن هر یک از نقاط $x=۱ , x=۲ , \ldots , x=n$ با یکدیگر برابر باشد. تابع چگالی متغیر تصادفی یکنواخت برابر است با:

$$f_X(x) = \frac{۱}{n} \qquad x = ۱, ۲, \ldots, n$$

$$= ۰ \qquad \text{سایر مقادیر}$$

امید، واریانس و تابع گشتاور توزیع یکنواخت عبارتند از:

$$E[X] = \sum_{۱}^{n} x \frac{۱}{n} = \frac{۱}{n}\left(\frac{n(n+۱)}{۲}\right) = \frac{n+۱}{۲}$$

به همین ترتیب:

$$E[X^۲] = \sum_{۱}^{n} x^۲ \frac{۱}{n} = \frac{۱}{n} \left(\frac{n(n+۱)(۲n+۱)}{۶}\right) = \frac{(n+۱)(۲n+۱)}{۶}$$

پس:

$$var(X) = E[X^۲] - E^۲[X] = \frac{(n+۱)(۲n+۱)}{۶} - \left(\frac{n+۱}{۲}\right)^۲ = \frac{n^۲-۱}{۱۲}$$

$$m_X(t) = E[e^{tx}] = \sum_{x=۱}^{n} \frac{e^{tx}}{n} = \frac{۱}{n}\sum_{۱}^{n} (e^t)^x = \frac{۱}{n} e^t \left(\frac{۱-e^{nt}}{۱-e^t}\right)$$

مثال ۱۳: یک تاس را پرتاب می‌کنیم X را متغیر تصادفی نتیجه حاصل از پرتاب در نظر می‌گیریم مطلوبست:

الف) تابع چگالی احتمال X.

ب) احتمال آمدن عدد زوج.

ج) امید، واریانس و تابع مولد گشتاور X.

حل: الف) X یک متغیر تصادفی یکنواخت است. بنابراین:

$$f(x) = \frac{۱}{۶} \qquad x = ۱, ۲, \dots, ۶$$

$$= ۰ \qquad \text{سایر مقادیر}$$

ب):

$$p(X = \text{عددی زوج}) = p(X=۲) + p(X=۴) + p(X=۶) = \frac{۳}{۶} = \frac{۱}{۲}$$

$$E[X] = \frac{۶+۱}{۲} = ۳/۵$$

ج):

$$var(X) = \frac{۳۶-۱}{۱۲} = ۲/۹۱$$

$$m_X(t) = \frac{e^t(۱-e^{۶t})}{۶(۱-e^t)}$$

فصل هفتم

در فصل قبل برخی توزیع‌های معروف و کاربردی را برای متغیر تصادفی گسسته ارایه نمودیم و خواص و ویژگیهای هر یک را بررسی نمودیم. در این فصل به بررسی توزیعهای مبتنی بر متغیر تصادفی پیوسته می‌پردازیم.

۱.۷ توزیع یکنواخت پیوسته

ساده‌ترین توزیع پیوسته، توزیع یکنواخت می‌باشد. فرض کنید تمام نقاط در بازه (a,b) دارای امکان وقوع یکسان باشند، در این صورت متغیر تصادفی X را که برد آن مقادیر موجود در بازه (a,b) می‌باشد، متغیر تصادفی یکنواخت می‌نامند. که با نماد $X \sim U(a,b)$ نشان داده می‌شود.

تابع چگالی احتمال متغیر تصادفی یکنواخت برابر است با:

$$f_X(x) = \begin{cases} \dfrac{۱}{b-a} & a < x < b \\ ۰ & \text{سایر مقادیر} \end{cases}$$

و به این ترتیب بدست می‌آید که چون X یکنواخت می‌باشد در نظر می‌گیریم:

$$f_X(x) = c \qquad a < x < b$$

در نتیجه:

$$\int_{-\infty}^{+\infty} f_X(x)\,dx = ۱ \quad \Rightarrow \quad \int_a^b c\,dx = c\,x\Big|_a^b = c\,(b-a) = ۱ \quad \Rightarrow \quad c = \frac{۱}{b-a}$$

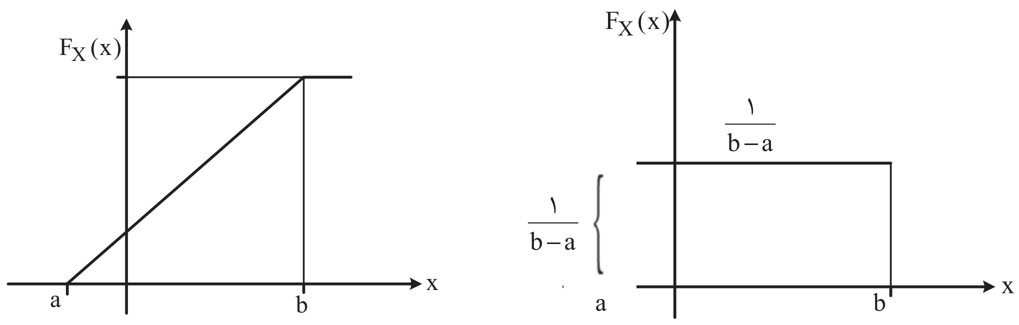
تابع توزیع متغیر تصادفی یکنواخت X برابر است با:

$$f_X(x) = \int_{-\infty}^{t} f_X(x) = \int_a^t \frac{۱}{b-a}\,dx = \frac{t-a}{b-a}$$

پس:

$$f_X(x) = \begin{cases} ۰ & t<a \\ \dfrac{t-a}{b-a} & a \le t \le b \\ ۱ & t > b \end{cases}$$

۷-۲ در شکل زیر نمودار تابع چگالی و تابع توزیع متغیر تصادفی یکنواخت نشان داده شده است:



مقادیر امید ریاضی واریانس و تابع مولد گشتاور را برای متغیر تصادفی یکنواخت X محاسبه می‌کنیم:

$$E[X] = \int_a^b x\,\frac{۱}{b-a}\,dx = \frac{۱}{b-a}\,\frac{x^۲}{۲}\Big|_a^b =$$

$$= \frac{۱}{b-a}\left(\frac{b^۲}{۲} - \frac{a^۲}{۲}\right) = \frac{(b-a)(a+b)}{۲(b-a)} = \frac{a+b}{۲}$$

$$E[X^۲] = \int_a^b x^۲ \frac{۱}{b-a} dx = \frac{۱}{b-a} \left(\frac{x^۳}{۳}\right)\Big|_a^b$$
$$= \frac{(b-a)(b^۲ - ab + a^۲)}{۳(b-a)} = \frac{a^۲ + ab + b^۲}{۳}$$

پس:

$$var(X) = E[X^۲] - E^۲[X] = \frac{a^۲ + ab + b^۲}{۳} - \frac{a^۲ + ۲ab + b^۲}{۴} = \frac{(b-a)^۲}{۱۲}$$
$$m_X(t) = E[e^{tX}] = \int_a^b e^{tX} \frac{۱}{b-a} dx = \frac{۱}{b-a} \left(\frac{e^{tX}}{t}\right)\Big|_a^b = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

با فرض $a \le c < d \le b$ مقدار احتمال $c \le x \le d$ برابر است با:

$$p(c \le x \le d) = \int_c^d \frac{۱}{b-a} dx = \frac{d-c}{b-a}$$

ملاحظه می‌کنید که مطابق با تعریف متغیر تصادفی یکنواخت مقدار احتمال تنها به طول بازه (c,d) وابسته است و نه به مقادیر c و d.

۳-۷ مثال ۱: نمرات دانشجویان یک کلاس در درس آمار به طور یکنواخت در فاصله ۱۲ الی ۲۰ توزیع شده است مطلوبست:

الف) تابع توزیع احتمال:

ب) احتمال قرارگیری نمرات در بازه $(۲۰-۱۸)$, $(۱۵-۱۳)$.

ج) میانگین نمرات کلاس، واریانس.

حل: الف) متغیر تصادفی X دارای توزیع یکنواخت است بنابراین:

$$f_X(x) = \begin{cases} \frac{۱}{۲۰-۱۲} = \frac{۱}{۸} & ۱۲ \le X \le ۲۰ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

ب):

$$P(۱۸ \le X \le ۲۰) = \int_{۱۸}^{۲۰} \frac{۱}{۸} dx \frac{۲۰-۱۸}{۸} = \frac{۲}{۸} = \frac{۱}{۴}$$
$$P(۱۳ \le X \le ۱۵) = \int_{۱۳}^{۱۵} \frac{۱}{۸} dx \frac{۱۵-۱۳}{۸} = \frac{۲}{۸} = \frac{۱}{۴}$$

ج):

$$E[X] = \frac{۲۰+۱۲}{۲} = ۱۶$$
$$var(X) = \frac{(۲۰-۱۲)^۲}{۱۲} = \frac{۶۴}{۱۲} = \frac{۱۶}{۳}$$

۴-۷ ۶. ۲ متغیر تصادفی نمایی

یک فرایند پواسون با پارامتر λ که برای یک واحد زمان نظاره می‌شود، را در نظر بگیرید اگر زمان شروع فرایند صفر $(t=۰)$ باشد و T مدت زمانی باشد که باید بگذرد تا اولین پیشامد رخ دهد در این صورت T یک متغیر تصادفی نمایی با پارامتر λ نامیده می‌شود.

تابع توزیع نمایی را با استفده از تعریف بدست می‌آوریم:

می‌دانیم $f_X(t) = p(X \le t)$ حال اگر $t < ۰$ باشد بنا به تعریف زمان منفی معنا ندارد. بنابراین اگر $t < ۰$ داریم: $p(X \le t) = ۰$ یعنی:

$$f_X(t) = ۰ \qquad t<۰$$

حال فرض می‌کنیم $t \geq 0$ باشد، احتمال اینکه هیچ پیشامدی در بازه $(0\,,\,t)$ رخ ندهد بنابه تابع چگالی احتمال پواسون برابر است با:

$$f_X(x=0) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}\Big|_{x=0} = e^{-\lambda t}$$

بنابراین $p(X > t) = e^{-\lambda t}$ اما $p(X \leq t) = 1-p(X>t)$ و داریم:

$$F_X(t) = p(X\leq t) = 1-p(X>t) = 1- e^{-\lambda t}$$

در نتیجه تابه توزیع متغیر تصادفی نمایی بصورت زیر بدست می‌آید:

$$F_X(t) = \begin{cases} 1-e^{-\lambda t} & t > 0 \\ 0 & t < 0 \end{cases}$$

با مشتق‌گیری از $F_X(t)$ تابع چگالی را بدست می‌آوریم:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

۵-۷ برای بدست آوردن مقادیر امید و واریانس از تابع مولد گشتاور استفاده می‌کنیم:

$$m_X(t) = E\left[e^{tx}\right] = \int_0^\infty e^{tx}\ \lambda e^{-\lambda x}\ dx = \lambda \int_0^\infty e^{tx}\ \lambda e^{-x(\lambda-t)}\ dx$$

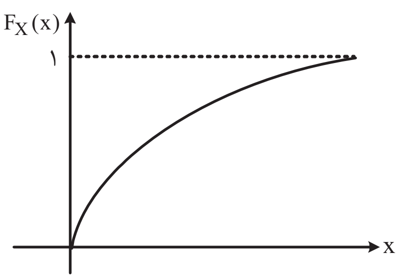
$$= \lambda\ \left(\frac{1}{t-\lambda}\ e^{-x(\lambda-t)}\right)\Big|_0^\infty = \frac{-\lambda}{t-\lambda} = \frac{\lambda}{\lambda-t}\quad (t<\lambda)$$

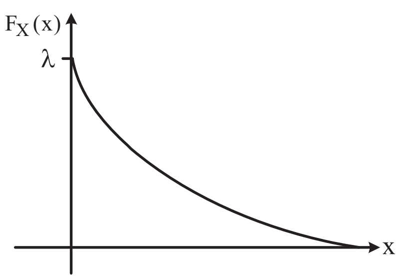
$$E[X] = m'_X(t=0) = \frac{\lambda}{(\lambda-t)^2}\Big|_{t=0} = \frac{1}{\lambda}$$

$$E[X^2] = m''_X(t=0) = \frac{2\lambda}{(\lambda-t)^3}\Big|_{t=0} = \frac{2}{\lambda^2}$$

$$\mathrm{var}(X) = E[X^2] - E^2[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

شکل زیر نمودار تابع توزیع و چگالی را برای متغیر تصادفی نمایش می‌دهد:





متغیر تصادفی نمایی با نماد $X \sim E(\lambda)$ نمایش داده می‌شود و معمولاً از آن به عنوان مدلی برای عمر قطعات و سیستم‌ها استفاده می‌شود.

۶-۷ مثال ۲: اگر لامپهای تولید شده توسط یک کارخانه به طور متوسط هر ۶ ماه یکبار بسوزند مطلوبست:

الف) تابع چگالی احتمال برای مدت زمان کارکرد لامپها.

ب) احتمال اینکه از زمان خرید، یک لامپ حداقل ۲ ماه کار کند؟

ج) احتمال اینکه یک لامپ قبل از ۴ ماه بسوزد؟

د) اگر ۱۰ عدد لامپ خریداری کنیم احتمال اینکه حداقل ۲ عدد از لامپها قبل از ۴ ماه بسوزند؟

ه) میانگین طول عمر، واریانس و تابع مولد گشتاور را برای متغیر تصادفی X که مدت زمان کارکرد لامپها می‌باشد بدست بیاورید.

۷-۷ حل: الف) متغیر تصادفی X را طول عمر لامپهای خریداری شده در نظر می‌گیریم در این صورت X یک متغیر تصادفی نمایی با پارامتر λ می‌باشد. برای بدست آوردن پارامتر λ ابتدا هر واحد زمانی را یک ماه در نظر می‌گیریم، λ برابر است با تعداد پیشامدها در یک واحد زمانی که در اینجا چون هر ۶ ماه یک لامپ می‌سوزد معادل است با اینکه $\frac{۱}{۶}$ لامپ در هر یک ماه می‌سوزد یعنی $\lambda = \frac{۱}{۶}$ و تابع چگالی برابر است با:

$$f_X(x) = \frac{۱}{۶}\, e^{-\frac{۱}{۶}x} \qquad x > ۰$$

$$= ۰ \qquad \text{سایر مقادیر}$$

ب) احتمال اینکه حداقل ۲ ماه طول بکشد تا یک لامپ بسوزد برابر است با:

$$p(X > ۲) = \int_۲^\infty \frac{۱}{۶}\, e^{-\frac{x}{۶}}\, dx = \frac{۱}{۶}\left(-۶\, e^{-\frac{x}{۶}}\right)\Big|_۲^\infty$$

$$= \frac{۱}{۶}\left(۰ - \left(-۶\, e^{-\frac{۲}{۶}}\right)\right) = e^{-\frac{۱}{۳}} = ۰/۷۱۶$$

ج) احتمال اینکه یک لامپ قبل از ۴ ماه بسوزد برابر است با:

$$p(X \le ۴) = \int_۰^۴ \frac{۱}{۶}\, e^{-\frac{x}{۶}}\, dx = \frac{۱}{۶}\left(-۶\, e^{-\frac{x}{۶}}\right)\Big|_۰^۴ = ۱ - ۶\, e^{-\frac{۴}{۶}} = ۰/۴۸۶$$

۸-۷ د) متغیر Y را بصورت زیر تعریف می‌کنیم:

$Y =$ تعداد لامپها از بین ۱۰ لامپ که در کمتر از ۴ ماه می‌سوزند.

طبق تعریف Y یک متغیر تصادفی دو جمله‌ای با پارامتر $n = ۱۰$ و $p = ۰/۴۸۶$ می‌باشد. پارامتر p از بند (ج) بدست می‌آید. احتمال اینکه هر لامپ در کمتر از ۴ ماه بسوزد برابر ۰/۴۸۶ می‌باشد، بنابراین می‌بایستی پارامتر p برابر ۰/۴۸۶ انتخاب شود. حال با احتمال اینکه حداقل ۲ عدد از لامپها قبل از ۴ ماه بسوزند برابر است با:

$$p(Y \ge ۲) = ۱ - p(Y < ۲) = ۱ - \left[p(Y=۱) + p(Y=۰)\right]$$

$$= ۱ - \left[\binom{۱۰}{۱}(۰/۴۸۶)^۱ (۰/۵۱۴)^۹ + \binom{۱۰}{۰}(۰/۴۸۶)^۰ (۰/۵۱۴)^{۱۰}\right]$$

$$= ۱ - (۰/۰۱۲ + ۰/۰۰۱۲) = ۰/۹۸۶$$

پس:

$$f_X(x) = \binom{۱۰}{y}(۰/۴۸۶)^y (۰/۵۱۴)^{۱۰-y}$$

اگر امید ریاضی $f_Y(y)$ را محاسبه کنیم داریم: $E[X] = np = ۱۰ \times ۰/۴۸۶ = ۴/۸۶$ یعنی تقریباً از هر ۱۰ عدد لامپ خریداری شده بطور متوسط ۵ لامپ در کمتر از ۴ ماه می‌سوزند.

ه) میانگین طول عمر لامپها عبارتست از: $E[X] = \frac{۱}{\lambda} = \frac{۱}{\frac{۱}{۶}} = ۶$ که با توجه به اینکه واحد زمان را هر یک ماه در نظر گرفتیم بنابراین میانگین طول عمر هر لامپ ۶ ماه می‌باشد که از صورت مثال نیز همین انتظار می‌رفت.

$$var(X)=\frac{۱}{\lambda^{۲}}=\frac{۱}{(\frac{۱}{۶})^{۲}}=۳۶$$

$$m_X(t)=\frac{\lambda}{\lambda-t}=\frac{\frac{۱}{۶}}{\frac{۱}{۶}-t}=\frac{۱}{۱-۶t}$$

۷-۹ ۶ . ۲ . ۱ رابطه توزیع هندسی و توزیع نمایی

توزیع هندسی طبق تعریف عبارتست از تعداد آزمایشها قبل از رسیدن به اولین پیروزی توزیع نمایی نیز از جهت تعریف تشابه زیادی با توزیع هندسی دارد. توزیع نمایی هم عبارتست از مدت زمانی که در یک فرایند پواسون با پارامتر λ باید سپری شود تا اولین پیشامد رخ دهد توجه کنید که میانگین توزیع هندسی $\frac{۱}{p}$ و توزیع نمایی $\frac{۱}{\lambda}$ می‌باشد. متغیرهای تصادفی هندسی و نمایی در یک خاصیت ویژه مشترک می‌باشند که هیچ متغیر تصادفی گسسته یا پیوسته دیگری این حالت را ندارد. ویژگی فوق به بی‌حافظگی معروف است و به صورت زیر می‌باشد:

برای توزیع نمایی: $p(X>a+b \mid X>a)=p(X>b)$

برای توزیع هندسی: $p(X=a+b \mid X\geq a)=p(X=b)$

اثبات برای توزیع نمایی: فرض می‌نیم A برابر با پیشامد $X>a$ و B برابر پیشامد $X>b$ و C برابر پیشامد $X>a+b$ باشد در این صورت باید ثابت کنیم: $p(C|A)=p(B)$ می‌دانیم:

$$p(C|A)=\frac{p(C\cap A)}{p(A)}$$

از آنجا که پیشامد C زیر مجموعه‌ای از پیشامد A است $(\{X>a+b\}\subset\{X>a\})$ بنابراین $C\cap A=C$ و داریم:

$$p(C|A)=\frac{p(C\cap A)}{p(A)}=\frac{p(C)}{p(A)}$$

۷-۱۰ حال مقادیر $p(A)$ و $p(B)$ و $p(C)$ را بدست می‌آوریم:

$$p(A)=p(X>a)=\int_a^{\infty}\lambda e^{-\lambda x}\,dx=e^{-\lambda a}$$

$$p(B)=p(X>b)=\int_b^{\infty}\lambda e^{-\lambda x}\,dx=e^{-\lambda b}$$

$$p(C)=p(X>a+b)=\int_{a+b}^{\infty}\lambda e^{-\lambda x}\,dx=e^{-\lambda(a+b)}$$

$$p(C|A)=\frac{p(C)}{p(A)}=\frac{e^{-\lambda(a+b)}}{e^{-\lambda a}}=e^{-\lambda b}=p(B)$$

بی‌حافظگی به این مفهوم است که اگر قطعه‌ای یا سیستمی به مدت a واحد زمان کار کرده باشد، احتمال اینکه حداقل b واحد زمان دیگر نیز کار کند $(a+b)$ برابر است با احتمال اینکه قطعه یا سیستم از لحظه صفر بخواهد حداقل b واحد زمان کار کند.

۷-۱۱ مثال ۳: در یک کارخانه ماشین‌های تولیدی هر یک ماه نیازمند سرویس تعمیرات باشند. اگر یک ماشین تولیدی ۶ ماه بدون تعمیرات کار کرده باشد احتمال اینکه در طول ماه بعد نیازمند تعمیات باشد چقدر است؟

حل: متغیر تصادفی X را مدت زمان لازم قبل از اولین تعمیر در نظر می‌گیریم در این صورت X یک متغیر تصادفی نمایی با پارامتر $\lambda=۱$ می‌باشد. می‌بایستی مقدار احتمال زیر را محاسبه کنیم:

$$p(X>۶+۱ \mid X>۶)=p(X>۷ \mid X>۶)$$
$$=p(X>۱)=\int_۱^{+\infty}e^{-x}\,dx=e^{-۱}=۰/۳۶$$

۷-۱۲ ۶ . ۳ توزیع گاما

اگر متغیرهای تصادفی $X_1, X_2, \cdots, X_n$ دارای توزیع نمایی با پارامتر λ باشند در این صورت متغیر تصادفی Y را که برابر است با مجموع متغیرهای تصادفی X_1 تا X_n، یک متغیر تصادفی گاما می‌نامیم.

$$Y = X_1 + X_2 + \cdots + X_n$$

تابع چگالی متغیر تصادفی گاما برابر است با:

$$f_X(x) = \frac{(\lambda x)^{n-1}}{T(n)} \lambda e^{-\lambda x} \qquad x \geq 0$$

$$= 0 \qquad \text{سایر مقادیر}$$

که در آن $T(n)$ تابع گاما می‌باشد و به صورت زیر تعریف می‌شود:

$$T(n) = \int_0^{\infty} x^{n-1} e^{-x} \, dx$$

$T(n)$ به ازای هر $n > 0$ موجود است و روابط زیر برای تابع گاما برقرار است:

۱− $T(n+1) = n\, T(n)$

۲− $T(n) = (n-1)!$ اگر n عدد طبیعی باشد

۳− $T(\frac{1}{2}) = \sqrt{\pi}$

از آنجا که متغیر تصادفی گاما مجموع n متغیر تصادفی نمایی است پس می‌توان مقادیر امید، واریانس و تابع مولد گشتاور را با استفاده از این خاصیت بدست آورد:

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{\lambda} = \frac{n}{\lambda}$$

$$\operatorname{var}(X) = \operatorname{var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{var}(X_i) = \sum_{1}^{n} \frac{1}{\lambda^2} = \frac{n}{\lambda^2}$$

$$m_X(t) = E\left[e^{tX}\right] = E\left[e^{t\sum_{1}^{n} X_i}\right] = E\left[e^{tX_1}\right] E\left[e^{tX_2}\right] \cdots E\left[e^{tX_n}\right]$$

$$= \left(\frac{\lambda}{\lambda - t}\right)\left(\frac{\lambda}{\lambda - t}\right) \cdots \left(\frac{\lambda}{\lambda - t}\right) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

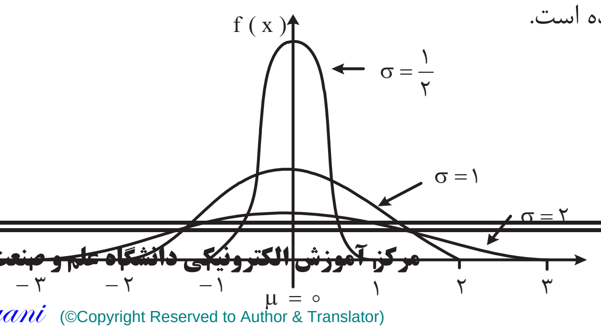
۷-۱۳ توزیع نرمال

مهمترین توزیع پیوسته که آنرا بررسی می‌کنیم توزیع نرمال می‌باشد، بسیاری از پدیده‌های طبیعی مثل قد و وزن افراد، نمرات درسی، میزان محصول در طول سال از توزیع نرمال پیروی می‌کنند.

متغیر تصادفی X دارای توزیع نرمال است و آنرا بصورت $X \sim N(\mu, \sigma^2)$ نمایش می‌دهیم اگر تابع چگالی احتمال بفرم زیر باشد:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty, \ \mu \in R, \ \sigma > 0$$

همانطور که ملاحظه می‌نید دو پارامتر توزیع نرمال همان میانگین و واریانس متغیر تصادفی X می‌باشند. مد، میانه و میانگین توزیع نرمال با یکدیگر برابر می‌باشند. در شکل زیر توزیع نرمال با مقادیر واریانس متفاوت نشان داده شده است.



تغیر میانگین در توزیع نرمال تنها نمودار تابع را به سمت راست یا چپ منتقل می‌کند. منحنی نرمال دارای خواص زیر می‌باشد:

۱- نقاط عطف منحنی عبارتند از $x_۱ = \mu + \sigma \ , \ x_۲ = \mu - \sigma$.

۲- منحنی نسبت به خط $x = \mu$ متقارن است بنابراین:

$$۱- \quad f_X(\mu - a) = f_X(\mu + a) \qquad\qquad (a > ۰)$$

$$۲- \begin{cases} p(X > a) = p(X < -a) \\ ۱ - F_X(x) = F_X(-x) \end{cases}$$

۳- مساحت سطح زیر منحنی نمودار نرمال برابر واحد می‌باشد. زیرا بوضوح داریم:

$$\int_{-\infty}^{+\infty} f(x)\, dx = ۱$$

۷-۱۴ برای بدست آوردن میانگین و واریانس از تابع مولد گشتاور استفاده می‌کنیم که بصورت زیر بدست می‌آید:

$$m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{+\infty} e^{tx} \frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x-\mu)^۲}{۲\sigma^۲}} dx$$

$$= \frac{۱}{\sigma\sqrt{۲\pi}} \int_{-\infty}^{+\infty} e^{tx - \frac{(x-\mu)^۲}{۲\sigma^۲}} dx$$

با تبدیل توان e به صورت مربع کامل داریم:

$$tx - \frac{(x-\mu)^۲}{۲\sigma^۲} = \frac{۱}{۲\sigma^۲}\left[x^۲ - ۲\mu x + \mu^۲ - ۲\sigma^۲ + x\right]$$

$$= \frac{۱}{۲\sigma^۲}\left[(x - \mu x - \sigma^۲ t)^۲ - ۲\sigma^۲ t\mu - \sigma^۴ t^۲\right]$$

$$= \frac{۱}{۲\sigma^۲}\left[(x - \mu - \sigma^۲ t)^۲\right] + t\mu + \frac{\sigma^۲ t^۲}{۲}$$

به این ترتیب:

$$m_X(t) = e^{\mu t + \frac{(\sigma t)^۲}{۲}} \int_{-\infty}^{+\infty} \underbrace{\frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x-\mu-\sigma^۲ t)^۲}{۲\sigma^۲}}}_{g_X(x)} dx$$

مقدار انتگرال $\int_{-\infty}^{+\infty} g(n)\, dx$ برابر یک می‌باشد زیرا $g_X(x)$ خود یک تابع نرمال با پارامترهای $N(\mu + \delta^۲ t \ , \ \delta^۲)$ می‌باشد. بنابراین مقدار تابع مولد گشتاور برابر است با:

$$m_X(t) = e^{\mu t + \frac{\sigma^۲ t^۲}{۲}}$$

حال داریم:

$$E[X] = m'_X(t)\Big|_{t=۰} = \left(\mu + \frac{\sigma^۲}{۲}(۲t)\right) e^{\mu t + \frac{\sigma^۲ t^۲}{۲}}\Big|_{t=۰} = \mu$$

$$E[X^۲] = m''_X(t)\Big|_{t=۰} = \sigma^۲ (e^{\mu t+\frac{\sigma^۲ t^۲}{۲}}) \;(\mu + \frac{\sigma^۲}{۲}(۲t))^۲ \; e^{\mu t+\frac{\sigma^۲ t^۲}{۲}} \Big|_{t=۰} = \sigma^۲ + \mu^۲$$

$$\Rightarrow \operatorname{var}(X) = E[X^۲] - E^۲[X] = \sigma^۲ + \mu^۲ = \sigma^۲$$

۷-۱۵ تابع توزیع متغیر تصادفی نرمال X عبارتست از:

$$F_X(t) = \int_{-\infty}^{t} \frac{۱}{\delta\sqrt{۲\pi}} \, e^{\frac{-(x-\mu)^۲}{۲\delta^۲}} \, dx$$

برای انتگرال فوق یک تابع اثلیه نمی‌توان بدست آورد به همین دلیل مجبور هستیم از روشهای عددی یک جدول برای مقادیر متفاوت t بدست بیاوریم. اما از آنجا که دو پارامتر μ و $\delta^۲$ نیز متغیر می‌باشند می‌بایستی روشی بدست بیاوریم که بتوان مقدار احتمال را بدون وابستگی به μ و $\delta^۲$ بدست آورد. متغیر تصادفی Z را بصورت زیر تعریف می‌کنیم:

$$Z = \frac{X-\mu}{\sigma}$$

بوضوح Z متغیر تصادفی نرمال با میانگین ۰ و واریانس ۱ می‌باشد. $Z \sim N(۰,۱)$ متغیر تصادفی Z را نرمال استاندارد می‌نامند. (که در فصل اول نیز برای نمونه‌های گرفته شده از یک جامعه معرفی شد).

$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{۱}{\sigma}\left[E[X]-\mu\right] = \frac{۱}{\sigma}(\mu-\mu) = ۰$$

$$\operatorname{var}(Z) = \operatorname{var}\left(\frac{X-\mu}{\sigma}\right) = \frac{۱}{\sigma^۲}\operatorname{var}(X-\mu) = \frac{\sigma^۲}{\sigma^۲} = ۱$$

$$m_Z(t) = e^{\frac{t^۲}{۲}}$$

حال با استفاده از تغیر متغیر $Z = \dfrac{X-\mu}{\sigma}$ می‌توان به سادگی مقادیر احتمال را از روی جدول توزیع متغیر تصادفی نرمال استاندارد بدست آورد.

تابع چگالی متغیر تصادفی نرمال استاندارد که آنرا با $\eta_Z(z)$ یا $\varphi(z)$ نمایش می‌دهیم برابر است با:

$$\eta_Z(z) = \varphi(z) = \frac{۱}{\sqrt{۲\pi}} \, e^{-\frac{Z^۲}{۲}} \qquad -\infty < Z < +\infty$$

تابع توزیع متغیر تصادفی نرمال استاندارد برابر است با:

$$\varphi(Z) = N_Z(z) = p_z(Z \le z) = \int_{-\infty}^{z} \frac{۱}{\sqrt{۲\pi}} \, e^{-\frac{Z^۲}{۲}} \, dx$$

جدول ضمیمه مقادیر تابع توزیع را برای $-۳/۵ \le Z \le ۳/۵$ نشان می‌دهد.

۷-۱۶ مثال ۴: با استفاده از جدول مقادیر احتمالات زیر را بدست بیاورید؟

الف) $p(Z < ۰) \;,\; p(Z > ۰)$

ب) $p(Z < ۱) \;,\; p(|Z| < \frac{۳}{۲})$

ج) $p(-۱ < Z < ۰) \;,\; p(۱ < Z < ۳)$

حل: الف)

$$p(Z < 0) = \varphi(0) = \frac{1}{2}$$

$$p(Z > 0) = 1 - \varphi(0) = 1 - \frac{1}{2} = \frac{1}{2}$$

ب)

$$p(Z < 1) = \varphi(1) = 0/8413$$

$$p(|Z| < \frac{3}{2}) = p(-\frac{3}{2} < Z < \frac{3}{2}) = \varphi(\frac{3}{2}) - \varphi(\frac{3}{2})$$

$$= 0/9332 - 0/0668 = 0/8664$$

ج)

$$p(-1 < Z < 0) = \varphi(0) - \varphi(1) = \frac{1}{2} - 0/1587 = 0/3413$$

$$p(1 < Z < 3) = \varphi(3) - \varphi(1) = 0/9987 - 0/8413 = 0/1574$$

با توجه به مثال فوق می‌وان خواص زیر را برای متغیر تصادفی نرمال استاندارد به دست آورد:

۱. $\varphi(-\infty) = 0$

۲. $\varphi(+\infty) = 1$

۳. $\varphi(a) = 1 - \varphi(-a)$

۷-۱۷ ۶ . ۴ . ۱ محاسبه مقادیر احتمال متغیر تصادفی نرمال

فرض کنید X یک متغیر تصادفی نرمال با میانگین μ و واریانس σ^2 باشد $(X \sim N(\mu, \sigma^2))$ برای بدست آوردن احتمال $p(a < x < b)$ داریم:

$$p(a < x < b) = p(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma})$$

$$= P(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma})$$

$$= \varphi(\frac{b-\mu}{\sigma}) - \varphi(\frac{a-\mu}{\sigma})$$

۷-۱۸ مثال ۵: اگر $X \sim N(5, 16)$ مطلوبست محاسبه $p(2 \le X < 7)$.

حل: داریم $\mu = 5$, $\sigma^2 = 16$ پس:

$$p(2 \le X < 7) = p(\frac{2-5}{4} \le \frac{X-5}{4} < \frac{7-5}{4})$$

$$= p(-\frac{3}{4} \le Z < \frac{1}{2})$$

$$= \varphi(\frac{1}{2}) - \varphi(-\frac{3}{4}) = 0/6915 = 0/2266 = 0/4649$$

۷-۱۹ مثال ۶: نمرات دانشجویان یک کلاس از توزیع نرمال با میانگین ۱۵ و واریانس ۴ پیروی می‌کند اگر بدانیم تمامی نمرات از ۱۰ بیشتر هستند احتمال اینکه نمرات بین ۱۲ تا ۱۶ باشد چقدر است؟

حل: متغیر تصادفی X نرمال می‌باشد $(X \sim N(15, 4))$ می‌بایستی احتمال شرطی زیر را محاسبه کنیم:

$$p(۱۲<X<۱۶)\mid X>۱۰)=\frac{p(۱۲<X<۱۶\,,\,X>۱۰)}{p(X>۱۰)}$$

$$=\frac{p(۱۲<X<۱۶)}{p(X>۱۰)}=\frac{p(\frac{۱۲-۱۵}{۲}<\frac{X-۱۵}{۲}<\frac{۱۶-۱۵}{۲})}{p(\frac{X-۱۵}{۲}>\frac{۱۰-۱۵}{۲})}$$

$$=\frac{p(-\frac{۳}{۲}<Z<\frac{۱}{۲})}{p(Z>-\frac{۵}{۲})}=\frac{\varphi(\frac{۱}{۲})-\varphi(-\frac{۳}{۲})}{\varphi(\frac{۵}{۲})}=\frac{۰/۶۹۱۵-۰/۰۶۶۸}{۰/۹۹۳۸}=۰/۶۲۸۵$$

۷-۲۰ ۶ . ۴ . ۲ تقریب توزیع دو جمله‌ای به توزیع نرمال

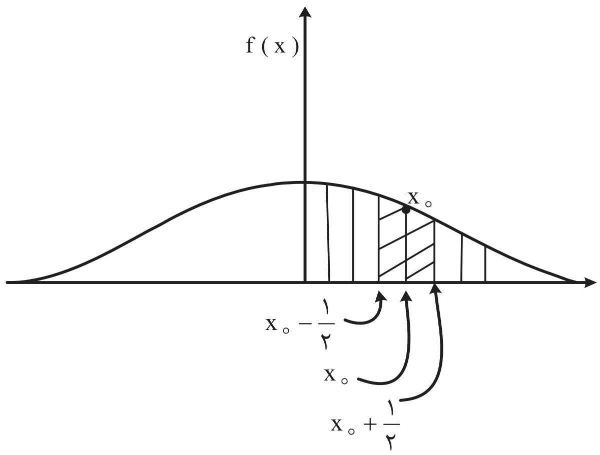
اگر در یک توزیع دو جمله‌ای تعداد آزمایشها بسیار زیاد باشد یعنی $n\to+\infty$ و در عین حال احتمال پیروزی و شکست تقریباً برابر باشند یعنی $p\approx q\approx\frac{۱}{۲}$ در این صورت می‌توان مقادیر احتمال دو جمله‌ای را با استفاده از توزیع نرمال بدست آورد.

برای تقریب توزیع دو جمله‌ای به نرمال مقادیر امید و واریانس را برابر قرار می‌دهیم یعنی:

$$\mu=np\quad,\quad\sigma^۲=npq$$

حال به شکل زیر توجه کنید:

برای بدست آوردن $p(X=x_\circ)$ در توزیع دو جمله‌ای کافیست.



در توزیع نرمال مساحت مشخص شده در شکل را بدست بیاوریم زیرا می‌دانیم $p(X=x_\circ)=f(x_\circ)$. و از طرفی مساحت حاشور خورده تقریباً با مساحت یک ذوذنقه برابر است:

$$\text{مساحت ذوذنقه حاشور خورده}=\text{ارتفاع}\times\frac{\text{مجموع دو قاعده}}{۲}=(x_\circ+\frac{۱}{۲}-(x_\circ-\frac{۱}{۲})\times\frac{f(x_\circ+\frac{۱}{۲})+f(x_\circ-\frac{۱}{۲})}{۲}$$

$$=\frac{f(x_\circ+\frac{۱}{۲})+f(x_\circ-\frac{۱}{۲})}{۲}\approx f(x_\circ)$$

حال داریم:

$$p(X\subset x_\circ)=p(x_\circ-\frac{۱}{۲}<X<x_\circ+\frac{۱}{۲})$$

$$=P\,\frac{x_\circ-\frac{۱}{۲}-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{x_\circ+\frac{۱}{۲}-\mu}{\sigma}$$

$$=\varphi(\frac{x_\circ+\frac{۱}{۲}-\mu}{\sigma})-\varphi(\frac{x_\circ-\frac{۱}{۲}-\mu}{\sigma})$$

و با جاگذاری مقادیر $\mu = np \;,\; \delta^{۲} = npq$ بدست می‌آوریم:

$$p(X=x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-np}{\sqrt{npq}}\right) - \varphi\left(\frac{x_{\circ}-\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

به همین ترتیب:

$$p(X \le x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

$$p(a < X \le b) = \varphi\left(\frac{b+\frac{۱}{۲}-np}{\sqrt{npq}}\right) - \varphi\left(\frac{a-\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

۷-۲۱ مثال ۷: احتمال بارندگی در طول یک ماه در یک شهر ۰/۳ . مطلوبست احتمال اینکه حداقل در ۵۰ شهر از ۲۰۰ شهر کشور در طول یک ماه آینده باران ببارد؟

حل: در این مثال X یک متغیر تصادفی دو جمله‌ای با پارامترهای $X \sim B(۲۰۰\,,\,۰/۳)$ می‌باشد می‌خواهیم احتمال $p(X > ۵۰)$ را محاسبه کنیم:

$$p(X > ۵۰) = ۱-p(X \le ۵۰) = ۱-\varphi\left(\frac{۵۰+\frac{۱}{۲}-(۲۰۰\times۰/۳}{\sqrt{۲۰۰\times۰/۳\times۰/۷}}\right)$$

$$=۱-\varphi\left(\frac{-۹/۵}{۶/۴}\right) = ۱-\varphi(-۱/۴۸) = ۱-۰/۰۶۹۴ = ۰/۹۳۰۶$$

۷-۲۲ ۶ . ۴ . ۳ تقریب توزیع پواسون به توزیع نرمال

برای توزیعهای پواسون با λ بزرگ (بزرگتر یا مساوی ۵) می‌توان مقادیر احتمال توزیع پواسون را با استفاده از توزیع نرمال بدست آورد در این حالت روابط دقیقاً مشابه روابط تقریب دو جمله‌ای به نرمال می‌باشند با این تفاوت که به جای $\mu \,,\, \delta^{۲}$ داریم: $\mu = \lambda \quad , \quad \sigma^{۲} = \lambda$

$$p(X=x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-\lambda}{\lambda}\right) - \varphi\left(\frac{x_{\circ}-\frac{۱}{۲}-\lambda}{\lambda}\right)$$

$$p(X \le x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-\lambda}{\lambda}\right)$$

$$p(a < X \le b) = \varphi\left(\frac{b+\frac{۱}{۲}-\lambda}{\lambda}\right) - \varphi\left(\frac{a-\frac{۱}{۲}-\lambda}{\lambda}\right)$$

۷-۲۳ مثال ۸: تلفن یک شرکت در هر ساعت تقریباً ۲۰ مرتبه زنگ می‌زند مطلوبست احتمال اینکه تلفن در طول ۲ ساعت آینده حداقل ۲۰ و حداکثر ۴۰ بار زنگ بزند؟

حل: X یک متغیر تصادفی پواسون می‌باشد واحد زمانی را هر یک ساعت در نظر می‌گیریم در این صورت $\lambda = ۱۰$ می‌باشد و می‌بایستی احتمال $p(۲۰ > X < ۴۰)$ را محاسبه کنیم:

$$p(۲۰ > X < ۴۰) = \varphi\left(\frac{۴۰+\frac{۱}{۲}-۱۰}{۱۰}\right) - \varphi\left(\frac{۲۰-\frac{۱}{۲}-۱۰}{۱۰}\right)$$

$$= \varphi(۳/۰۵) - \varphi(۰/۹۵) = ۰/۹۹۸۹ - ۰/۸۲۸۹ = ۰/۱۷$$

فصل هشتم

توزیع‌های نمونه‌ای

در فصل‌های پیش نشان دادیم که با داشتن تابع توزیع یک متغیر تصادفی چگونه می‌توان مقدار احتمالات مورد نظر را بدست آورد. اما در عمل در بسیاری از اوقات می‌دانیم که یک متغیر تصادفی (مثلاً بارش باران در طول سال) از چه توزیع پیروی می‌کند اما مشکل اینجاست که مقدار پارامترهای توزیع را نمی‌دانیم. به عنوان مثال تعداد دفعات برنده شدن حساب بانکی یک شخص در طول ۱۰ سال یک متغیر تصادفی دو جمله‌ای با پارامترهای $X \sim B(۱۰\ ,\ p)$ می‌باشد. اما p یا همان احتمال برنده شدن حساب را نمی‌دانیم بنابراین برای بدست آوردن آن نیازمند روشهایی هستیم که بتوان پارامترهای مجهول را برآورد نمود در این فصل و فصل‌های بعدی روشهایی را برای این منظور ارایه می‌کنیم.

۸ . ۱ نمونه تصادفی

یک جمعیت آماری را با متغیر تصادفی X و توزیع احتمال $f_X(x)$ در نظر بگیرید. از این جمعیت n عضو را انتخاب می‌کنیم، این n عضو را یک نمونه nتایی گوییم. اولین عضو از نمونه را با $X_۱$ و دومین عضو را با $X_۲$ و به همین ترتیب nامین عضو انتخابی را با X_n نمایش می‌دهیم. به این ترتیب نمونه nتایی به صورت $X_۱, X_۲, \ldots, X_n$ خواهد بود که در آن هر یک از X_i یا خود یک متغیر تصادفی یا همان توزیع $f_X(x)$ می‌باشند. و از آنجا که تعداد اعضای جمعیت آماری نامتناهی یا بسیار زیاد می‌باشد $(N \to \infty)$. در این صورت انتخاب X_i یا مستقل از X_j $(i \neq j)$ می‌باشد. و این یعنی توزیع توام $X_۱$ تا X_n برابر با حاصل ضرب توزیع $X_۱$ تا X_n می‌باشد. حال با توجه به مطالب فوق تعریف نمونه تصادفی را بصورت زیر ارایه می‌کنیم:

تعریف:

گوییم $X_۱, X_۲, \ldots, X_n$ یک نمونه تصادفی از متغیر تصادفی X می‌باشد اگر و فقط اگر تابع احتمال $f(X_۱, X_۲, \ldots, X_n)$ بصورت زیر باشد:

$$f(X_۱, X_۲, \ldots, X_n) = \prod_{i=۱}^{n} f_X(x_i) \qquad \forall\ (x_۱, x_۲, \ldots, x_n)$$

حال در مثال‌های بعدی نحوه بکارگیری نمونه‌گیری را مشخص می‌کنیم.

مثال ۱: مدیر یک کارخانه کنسروسازی ادعا می‌ند که از هر ۱۰۰۰۰ کنسرو تولید شده تنها ۵ قوطی دارای اشکال می‌باشد، به وضوح برای اثبات ادعای وی نمی‌توانیم تمام قوطی‌ها را بررسی کنیم بنابراین از میان تمام قوطی‌ها ۱۰ قوطی انتخاب می‌کنیم (در این مثال برای سادگی تعداد کمتری نمونه انتخاب کردیم) و آنها را مورد بررسی قرار می‌دهیم. مطلوبست:

الف) تعریف متغیر تصادفی X و نمونه‌های تصادفی x_i $\quad i = ۱, ۲, \ldots, ۱۰$

ب) تابع چگالی احتمال توام نمونه‌های تصادفی x_i .

حل: الف) متغیر تصادفی X را به این صورت تعریف می‌کنیم:

$$X = \begin{cases} ۰ & \text{اگر قوطی سالم باشد} \\ ۱ & \text{اگر قوطی ناسالم باشد} \end{cases}$$

به این ترتیب نمونه‌های تصادفی X_i به این صورت تعریف خواهد شد:

$$X = \begin{cases} ۰ & \text{اگر نمونه } i \text{ام سالم باشد} \\ ۱ & \text{اگر نمونه } i \text{ام ناسالم باشد} \end{cases}$$

بنابراین متغیر تصادفی X یک متغیر تصادفی برنولی با پارامتر p می‌باشد.

نمونه تصادفی عبارتست از $X_۱, X_۲, \ldots, X_n$ که هر یک نیز به نوبه خود دارای توزیع برنولی با همان پارامتر p می‌باشند. یعنی داریم:

$$f_{Xi}(xi) = p^{x_i}(۱-p)^{۱-x_i} \qquad i = ۱, ۲, \ldots, ۱۰$$
$$x_i = ۰ \text{ یا } ۱$$
$$= ۰ \qquad \text{برای سایر مقادیر}$$

ب) از آنجا که نمونه‌ها از بین تعداد زیادی (۱۰/۰۰۰) کنسرو انتخاب می‌شوند بنابراین انتخاب با جایگذاری و بدون جایگذاری تقریباً معادل یکدیگر می‌باشند و می‌توان گفت که X_iها از یکدیگر مستقل می‌باشند. در این صورت تابع چگالی توام X_iها عبارتست از:

$$f(x_۱, x_۲, \dots, x_{۱۰}) = \prod_{i=۱}^{۱۰} f_X(x_i) = p^{\sum_{i=۱}^{۱۰} x_i} (۱-p)^{۱۰-\sum_{i=۱}^{۱۰} x_i}$$

مثال ۲: می‌دانیم که طول قد دانشجویان یک دانشگاه یک متغیر تصادفی نرمال با میانگین μ و واریانس $\sigma^۲$ می‌باشد. اگر به تصادف ۱۰ نفر از دانشجویان انتخاب کنیم و $X_۱$ را طول قد نفر اول و $X_۲$ را طول قد نفر دوم و به همین ترتیب $X_{۱۰}$ را طول قد نفر دهم قرار دهیم. در این صورت مطلوبست:

الف) تابع چگالی احتمال توام $X_۱$ تا $X_{۱۰}$.

ب) احتمال اینکه قد هر ۱۰ نفر از ۱/۵ متر بیشتر باشد؟

ج) احتمال اینکه هیچ کس قدی بلندتر از ۲ متر نداشته باشد؟

حل: الف) از آنجا که هر یک از X_iها نمونه‌های تصادفی می‌باشند که از یکدیگر مستقل بوده و از توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ پیروی می‌کند بنابراین:

$$f(X_۱, X_۲, \dots X_n) = \prod_{i=۱}^{۱۰} f_X(x_i) = \prod_{۱}^{۱۰} \frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x_i-\mu)^۲}{۲\sigma^۲}} = \frac{۱}{\sigma^{۱۰}(۲\pi)^۵} e^{\frac{-\sum_{i=۱}^{۱۰}(x_i-\mu)^۲}{۲\sigma^۲}}$$

ب) احتمال اینکه قد هر ۱۰ نفر از ۱/۵ متر بیشتر باشد برابر است با:

$$p(X_۱ > ۱/۵) = ۱ - p(X_۱ \leq ۱/۵) = ۱ - p\left(\frac{X_۱-\mu}{\sigma} < \frac{۱/۵-\mu}{\sigma}\right)$$

$$= ۱ - N_Z\left(Z < \frac{۱/۵-\mu}{\sigma}\right)$$

برای $X_۱$ تا $X_{۱۰}$ داریم:

$$p(X_۱, \dots, X_{۱۰} > ۱/۵) = \left(۱ - N_Z\left(\frac{۱/۵-\mu}{\sigma}\right)\right)^{۱۰}$$

ج) مطابق بند (ب) مقدار این احتمال هم برابر است با:

$$p(X_۱, \dots, X_{۱۰} < ۲) = \left(N_Z\left(\frac{۲-\mu}{\sigma}\right)\right)^{۱۰}$$

همانطور که ملاحظه می‌کنید تابع چگالی توام و نتایج احتمالات به مقادیر μ و $\sigma^۲$ وابسته می‌باشد. در فصل برآورد نشان می‌دهیم که با برآورد این پارامترها می‌توان مقادیر احتمالات را محاسبه نمود.

۸ . ۲ آماره‌ها

اگر با استفاده از عضوهای یک نمونه تصادفی (مثل $X_۱$, $X_۲$, ...) یک تابع بسازیم که به مقادیر مجهول پارمتر وابستگی نداشته باشد به آن تابع یک آماره می‌گوییم.

به این ترتیب اگر $X_۱, X_۲, \dots, X_n$ یک نمونه تصادفی باشند توابع زیر همگی آماره می‌باشند:

$$X_۲ + X_۴ + X_۸ \quad , \quad \frac{X_n}{X_۱} \quad , \quad \frac{۱}{n}\sum_{i=۱}^{n}(X_i)^۲$$

$$\prod_{i=۱}^{n} X_i \quad , \quad \sum_{i=۱}^{n}(X_i - ۲)^۲$$

اما توابع زیر بشرطی آماره می‌باشند که مقادیر μ و $\sigma^۲$ مجهول نباشند:

$$X_n-\mu \quad , \quad \frac{۱}{n-۱}\sum_{۱}^{n}(X_i-\mu)^۲ \quad , \quad \frac{X_۲-X_۱}{۲\sigma^۲} \quad , \quad \frac{X_۱-\mu}{\sigma}$$

به عنوان مثال $\overline{X}=\frac{۱}{n}\sum_{i=۱}^{n}X_i$ میانگین نمونه تصادفی می‌باشد که یک آماره هم می‌باشد. اما این میانگین با پارامتر μ در جامعه تفاوت می‌کند زیرا میانگین نمونه از نمونه‌ای به نمونه دیگر تغیر می‌ند اما پارامتر جامعه μ همواره ثابت است.

در استنباط آماری با گرفتن نمونه‌هایی از جامعه و بدست آوردن آماره‌هایی مثل $\overline{X}=\frac{۱}{n}\sum_{i=۱}^{n}X_i$ یا $S^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\overline{X})^۲$ می‌خواهیم نتیجه‌گیری‌هایی در مورد پارامترهای جامعه مثا میانگین μ و واریانس $\sigma^۲$ داشته باشیم.

حال همانطور که برای یک متغیر تصادفی X مقادیری مثل میانگین، واریانس، kامین گشتاور را معرفی نمودیم برای متغیر تصادفی X نیز این مقادیر را می‌توانیم تعریف کنیم.

مقدار $\overline{X}=\frac{۱}{n}\sum_{i=۱}^{n}X_i$ میانگین نمونه گفته می‌شود. به همین ترتیب مقدار $S^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\overline{X})^۲$ نیز مقدار واریانس نمونه و $S=\sqrt{S^۲}$ نیز مقدار انحراف از معیار نمونه نامیده می‌شود.

Kامین گشتاور نمونه عبارتست از:

$$M_K=\frac{۱}{n}\sum_{i=۱}^{n}X_i^k \qquad k=۱,۲,\ldots$$

توجه کنید که مقادیر $\overline{X}$ و $S^۲$ یک آماره می‌باشد و در نتیجه خود یک متغیر تصادفی می‌باشد این مطلب برای مقادیر مختلف K در M_K (مثل $M_۱$ ، $M_۲\ldots$) نیز برقرار می‌باشد.

مثال ۳: نشان دهید روابط زیر دارای $S^۲$ برقرار است:

الف)
$$S^۲=\frac{۱}{n-۱}\left(\sum X_i^۲-n\overline{X}^۲\right)$$

ب)
$$S^۲=\frac{n}{n-۱}\left(M_۲-\overline{X}^۲\right)$$

حل: الف)

$$S^۲=\frac{۱}{n-۱}\sum_{۱}^{n}(X_i-\overline{X})^۲=\frac{۱}{n-۱}\sum_{۱}^{n}(X_i^۲-۲X_i\overline{X}+\overline{X}^۲)$$

$$=\frac{۱}{n-۱}\left(\sum_{۱}^{n}X_i^۲-۲\overline{X}\sum_{۱}^{n}X_i+n\overline{X}^۲\right)$$

$$=\frac{۱}{n-۱}\left(\sum_{۱}^{n}X_i^۲-۲\overline{X}(n\overline{X})+n\overline{X}^۲\right)=\frac{۱}{n-۱}\left(\sum_{۱}^{n}X_i^۲-n\overline{X}^۲\right)$$

ب) بوضوح از بند الف داریم:

$$S^۲=\frac{۱}{n-۱}\left(\sum_{۱}^{n}X_i^۲-n\overline{X}^۲\right)=\frac{۱}{n-۱}\times n\left(\frac{۱}{n}\sum_{۱}^{n}X_i^۲-\overline{X}^۲\right)$$

$$=\frac{n}{n-۱}\left(M_۲-\overline{X}^۲\right)$$

در بخش توزیع‌های نمونه‌ای علت بکار بردن $\frac{۱}{n-۱}$ به جای $\frac{۱}{n}$ را در محاسبه $S^۲$ توضیح می‌دهیم.

مثال ۴: در مثال ۲ نتیجه، حاصل از نمونه‌گیری از قد ۱۰ نفر از دانشجویان بصورت زیر می‌باشد:

۱۸۴ , ۱۷۳ , ۱۶۶ , ۱۹۲ , ۱۷۵ , ۱۷۰ , ۱۵۰ , ۱۶۱ , ۱۸۷ , ۱۷۸

مطلوبست:

الف) میانگین و واریانس نمونه.

ب) گشتاور اول و دوم نمونه.

حل: الف)

$$\overline{X} = \frac{۱}{۱۰}\sum_{۱}^{۱۰} X_i = \frac{۱}{۱۰}(۱۷۸+۱۸۷+۱۶۱+۱۵۰+۱۷۰+۱۷۵+۱۹۲+۱۶۶+۱۷۳+۱۸۴) = ۱۷۳/۶$$

برای محاسبه واریانس نمونه از گشتاور اول و دوم استفاده می‌نیم:

$$M_۱ = M_۱ = \frac{۱}{n}\sum_{۱}^{n} X_i = \overline{X} = ۱۷۳/۶$$

$$M_۲ = \frac{۱}{n}\sum_{۱}^{n} X_i^۲ = \frac{۱}{۱۰}\sum_{۱}^{۱۰}(X_i)^۲ = \frac{۱}{۱۰}(۱۷۸^۲+۱۸۷^۲+۱۶۱^۲+۱۵۰^۲+۱۷۰^۲$$
$$+۱۷۵^۲+۱۹۲^۲+۱۶۶^۲+۱۷۳^۲+۱۸۴^۲ = ۳۰۲۸۰/۴$$

بنابراین $S^۲$ برابراست با:

$$S^۲ = \frac{n}{n-۱}(M_۲ - \overline{X}^۲) = \frac{۱۰}{۹}(۳۰۲۸۰/۴ - ۱۷۳/۶^۲)$$
$$= \frac{۱۰}{۹}(۱۴۳/۴۴) = ۱۵۹/۳۷$$

انحراف از معیار نمونه:

$$\Rightarrow S = \sqrt{۱۵۹/۳۷} = ۱۲/۶۲$$

۸ . ۳ آماره مرتب، میانه و برد نمونه

نمونه تصادفی $X_۱, X_۲, \ldots, X_n$ از متغیر تصادفی X را در نظر بگیرید اگر مقادیر مشاهده شده نمونه تصادفی را از کوچک به بزرگ مرتب کنیم بطوریکه $X_{(۱)} \le X_{(۲)} \le \ldots \le X_{(n)}$ را داشته باشیم در این صورت $X_{(i)}$ را iامین آماره مرتب می‌نامیم. $(i = ۱, ۲, \ldots, n)$ همچنین به $X_{(۱)}$ حداقل مقدار نمونه و به $X_{(n)}$ حداکثر مقدار نمونه گفته می‌شود. میانه نمونه عبارتست از مقدار وسط آماره مرتب به عبارت دقیقتر $m_\circ$ را میانه نمونه در نظر می‌گیریم که بصورت زیر بدست می‌آید:

اگر n فرد باشد $m_\circ = X_{(\frac{n+۱}{۲})}$.

اگر n زوج باشد $= \frac{X(\frac{n}{۲}) + X(\frac{n}{۲}+۱)}{۲}$.

برد نمونه نیز عبارتست از تفاضل بزرگترین آماره مرتب از کوچکترین آماره مرتب که با $R_\circ$ نمایش داده می‌شود. $R_\circ = X_{(n)} - X_{(۱)}$

مثال ۵: نمونه تصادفی مشاهده شده در مثال ۴ را در نظر بگیرید، آماره مرتب، حداقل مقدار نمونه، حداکثر مقدار نمونه، میانه و برد نمونه را بدست بیاورید؟

حل: با مرتب نمودن نمونه‌های مشاهده شده داریم:

$$X_{(۱)}=۱۵۰ \;,\; X_{(۲)}=۱۶۱ \;,\; X_{(۳)}=۱۶۶ \;,\; X_{(۴)}=۱۷۰ \;,\; X_{(۵)}=۱۷۳$$
$$X_{(۶)}=۱۷۵ \;,\; X_{(۷)}=۱۷۸ \;,\; X_{(۸)}=۱۸۴ \;,\; X_{(۹)}=۱۸۷ \;,\; X_{(۱۰)}=۱۹۲$$

$X_{(۱)}=۱۵۰$ حداقل مقدار نمونه.

$X_{(۱۰)}=۱۹۲$ حداکثر مقدار نمونه.

میانه نمونه : $\dfrac{X_{(\frac{۱۰}{۲})}+X_{(\frac{۱۰}{۲}+۱)}}{۲}=\dfrac{X_{(۵)}+X_{(۶)}}{۲}=\dfrac{۱۷۳+۱۷۵}{۲}=۱۷۴$

برد نمونه : $X_{(۱۰)}-X_{(۱)}=۱۹۲-۱۵۰=۴۲$

۸–۱۱

۸ . ۴ تابع توزیع نمونه‌ای

برای نمونه تصادفی $X_۱, X_۲, \ldots, X_n$ و آماره مرتب $X_{(۱)} \le X_{(۲)} \le \ldots \le X_{(n)}$ تابع توزیع نمونه‌ای بصورت زیر تعریف می‌شود:

$$F^*(t)=\begin{cases} ۰ & t < X_{(۱)} \quad \text{برای} \\ \dfrac{r}{n} & X_{(r)} \le t < X_{(r+۱)} \qquad (r=۱, \ldots, n) \\ ۱ & t > X_{(n)} \end{cases}$$

۸–۱۲

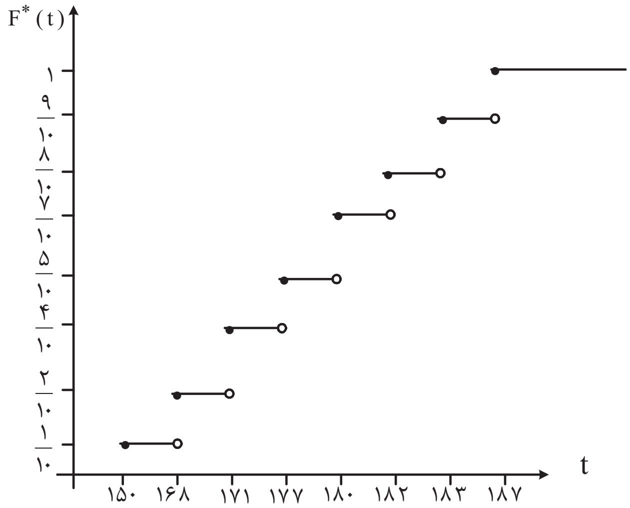
مثال ۶: اگر نمونه‌های مشاهده شد از قد دانشجویان در مثال ۴ بصورت زیر باشد مطلوبست تابع توزیع نمونه‌ای برای نمونه تصادفی مشاهده شده؟

۱۵۰ , ۱۶۸ , ۱۷۱ , ۱۷۱ , ۱۷۷ , ۱۸۰ , ۱۸۰ , ۱۸۲ , ۱۸۳ , ۱۸۷

حل:

$$F^*(t)=\begin{cases} ۰ & t < ۱۵۰ \\ \dfrac{۱}{۱۰} & ۱۵۰ \le t < ۱۶۸ \\ \dfrac{۲}{۱۰} & ۱۶۸ \le t < ۱۷۱ \\ \dfrac{۴}{۱۰} & ۱۷۱ \le t < ۱۷۷ \\ \dfrac{۵}{۱۰} & ۱۷۷ \le t < ۱۸۰ \\ \dfrac{۷}{۱۰} & ۱۸۰ \le t < ۱۸۲ \\ \dfrac{۸}{۱۰} & ۱۸۲ \le t < ۱۸۳ \\ \dfrac{۹}{۱۰} & ۱۸۳ \le t < ۱۸۷ \\ ۱ & ۱۸۷ \le t \end{cases}$$

نمودار تابع توزیع نمونه‌ای فوق نیز بصورت زیر می‌باشد.



÷

۸ . ۵ نامساوی چبی چف

فرض کنید X یک متغیر تصادفی باشد که نحوه توزیع و تابع چگالی آنرا نمی‌دانیم در این صورت برای اظهار نظر در مورد مقادیر احتمال که متغیر تصادفی X قبول می‌کند می‌توانیم از قضیه چبی چف استفاده کنیم. این قضیه بصورت زیر می‌باشد:

اگر X یک متغیر تصادفی باشد و $g(X)$ یک تابع حقیقی غیر منفی از X باشد در این صورت:

$$p\left[g(X) \geq K\right] \leq \frac{۱}{K} E\left[g(X)\right] \qquad g(X) \geq ۰ \ , \ K > ۰$$

با انتخاب $g(X) = (X-\mu_X)^۲$ و $k^۲ \sigma_X^۲$ بجای k داریم:

$$p(X-\mu_X)^۲ \geq k^۲ \sigma_X^۲) \leq \frac{۱}{k^۲ \sigma_X^۲} E\left[(X-\mu_X)^۲\right]$$

$$p(\left|X-\mu_X\right| \geq K\sigma_X) \leq \frac{۱}{k^۲}$$

بنابراین:

$$p(\left|X-\mu_X\right| < K\sigma_X) \geq ۱ - \frac{۱}{k^۲}$$

مقادیر احتمال که توسط قانون چبی چف بدست می‌آیند بسته به مقدار k از دقتهای مختلفی برخوردارند بطوریکه هر چه k بیشتر باشد دقت محاسبه احتمال توسط قانون چبی چف بیشتر است. این مطلب را در مثال زیر نشان می‌دهیم:

مثال ۷: اگر X یک متغیر تصادفی نرمال باشد در این صورت مطلوبست محاسبه احتمال $p(\left|X-\mu_X\right| < k\sigma_X)$ با استفاده از توزیع نرمال و قانون چبی چف به ازای مقادیر $k = ۱ , ۲ , ۳ , ۴$.

حل: ابتدا مقدار احتمال را برای کتغیر تصادفی نرمال X بدست می‌آوریم:

$$p(\left|X-\mu_X\right| < K\sigma_X) = p(-k < \frac{X-\mu_X}{\sigma_X} < k)$$

$$= N_Z(k) - N_Z(-k) = ۲N_Z(k) - ۱$$

حال در جدول زیر مقادیر مختلف k را با استفاده از رابطه فوق و قانون چبی چف محاسبه می‌کنیم:

K	$2N_Z(k)-1$	$1-\frac{1}{k^2}$
۱	۰/۶۸۲۶	۰
۲	۰/۹۵۴۶	۰/۷۵
۳	۰/۹۹۷۴	۰/۸۸۸۹
۴	۱/۰۰۰	۰/۹۳۳۴

همانطور که ملاحظه می‌کنید مقادیر بدست آمده برای $k=1,2$ از خطای بالایی برخوردار هستند اما با افزایش k مقدار بدست آمده از قانون چبی چف به مقدار واقعی نزدیک‌تر می‌شود. می‌توان تابع چگالی احتمال را طوری تعریف نمود که به ازای هر مقدار $k \geq 1$ مقدار واقعی احتمال با احتمال بدست آمده از رابطه چبی چف برابر باشد. این حالت را در مثال بعد نشان می‌دهیم:

مثال ۸: فرض کنید تابع چگالی احتمال متغیر تصادفی گسسته X بصورت زیر باشد:

$$f_X(x)=\frac{k^2-1}{k^2} \qquad x=0 \text{ برای } \qquad k\geq 1$$

$$=\frac{1}{2k^2} \qquad x=-k\,,\,k \text{ برای }$$

$$=0 \qquad \text{سایر مقادیر}$$

مطلوبست: الف) محاسبه μ_X , σ_X.

ب) محاسبه احتمال $p(|X-\mu_X|<k\sigma_X)$ به وسیله تابع احتمال و بوسیله قانون چبی چف و مقایسه آندو.

حل: الف)

$$\mu_x=\sum_{i=1}^{n} f(xi)\,x_i = 0\times\frac{k^2-1}{k^2}+k\left(\frac{1}{2k^2}\right)+(-k)\left(\frac{1}{2k^2}\right)=0+\frac{1}{2k}-\frac{1}{2k}=0$$

$$E\left[X^2\right]=\sum_{1}^{n} x_i^2\,f(x_i) = 0\times\frac{k^2-1}{k^2}+k^2\left(\frac{1}{2k^2}\right)+(-k)^2\left(\frac{1}{2k^2}\right)$$

$$=0+\frac{1}{2}+\frac{1}{2}=1 \quad \sigma_x^2=E\left[X^2\right]-E^2[X]=\sigma_x^2=1-0^2=1$$

ب) ابتدا احتمال دقیق را محاسبه می‌کنیم:

$$\mu_x=0 \qquad \sigma_x=1$$

$$\Rightarrow p(|X-\mu_x|<k\,\sigma_x)=p(|X|<K)$$

$$p(-k<X<k)=p(X=-k)+p(X=k)=\frac{1}{2k^2}+\frac{1}{2k^2}=\frac{1}{k^2}$$

اما طبق قانون چبی چف رابطه زیر برقرار می‌باشد:

$$p(|X-\mu_x|<k\,\sigma_x)\leq\frac{1}{k^2}$$

مقدار $\frac{1}{k^2}$ دقیقاً برابر با احتمال محاسبه شده می‌باشد بنابراین در این حالت مخصوص مقدار دقیق احتمال با مقدار بدست آمده از قانون چبی چف برابر می‌باشد و به عبارتی حالت مساوی در قانون چبی چف رخ می‌دهد.

۸ . ۶ قضیه حد مرکزی

همانطور که قبلاض اشاره کردیم تابع نرمال نقش مهمی در احتمالات بازی می‌کند زیرا اولاً محاسبه احتمال نرمال با استفاده از تابع نرمال استاندارد و جدول مربوط به آن بسیار ساده می‌باشد ثانیاً بسیاری از محاسبات احتمال را همانطور که در فصلهای قبل نشان دادیم تحت شرایط خاصی می‌توان با استفاده از تابع نرمال تقریب زد. حال در اینجا نشان می‌دهیم که اگر متغیرهای تصادفی X_i $(i = ۱,۲,...)$ از طکدیگر مستقل باشند و از یک تابع توزیع پیروی کنند، مجموع آنها به سمت تابع نرمال میل می‌کند. در اینجا نمی‌دانیم که متغیرهای تصادفی X_i از چه توزیعی پیروی می‌کند. صورت قضیه حد مرکزی بصورت زطر می‌باشد:

فرض کنید $X_۱, X_۲, ..., X_n$ نمونه‌ای تصادفی از جامعه نامتناهی X با میانگین μ و $\sigma^۲$ باشند در این صورت دنباله متغیرهای تصادفی $Z_۱, Z_۲, ..., Z_n$ را به صورت زیر تعریف می‌کنیم:

$$Z_n = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad n = ۱,۲, ...$$

که در آن $\overline{X}_n = \frac{۱}{n}\sum_{i=۱}^{n} X_i$ در این صورت اگر $F_{Z_n}(t)$ تابع توزیع متغیرهای تصادفی Z_n به ازای مقدار حقیقی t باشد داریم:

$$\lim_{n\to\infty} F_{Z_n}(t) = N_Z(t)$$

که در آن $N_Z(t)$ تابع توزیع نرمال استاندارد می‌باشد.

به عبارت ساده‌تر متغیر تصادفی $Z = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ وقتی n به سمت بی‌نهایت میل می‌کند دارای توزیع نرمال استاندارد $N(۰ , ۱)$ می‌باشد. یعنی برای nهای به اندازه کافی بزرگ داریم $F_{Z_n}(t) = N_Z(t)$ اما برای هر مقدار n داریم:

$$F_{\overline{X}_n}(t) = F_{Z_n}\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

در نتیجه برای nهای بزرگ:

$$F_{\overline{X}_n}(t) = N_Z\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

رابطه فوق به این معنی است که تابع توزیع میانگین تعداد زیادی از متغیرهای تصادفی مستقل با توزیع یکسان تقریباً مساوی است با توزیع نرمال استاندارد. به همین دلیل است که تعداد زیادی از قوانین احتمال را بوسیله نرمال تقریب می‌زنند.

۸-۱۷

به عنوان مثال به تقریب دو جمله‌ای به نرمال توجه کنید که در فصل قبل روابط آنرا بصورت زیر نشان دادیم:

$$p(X = x_۰) = N\left(\frac{x_۰ + \frac{۱}{۲} - np}{\sqrt{pq}}\right) - N\left(\frac{x_۰ - \frac{۱}{۲} - np}{\sqrt{pq}}\right)$$

می‌دانیم اگر متغیرهای تصادفی برنولی و مستقل $X_۱, X_۲, ..., X_n$ را هر یک با پارامتر p در نظر بگیریم در این صورت $Y = \sum_{۱}^{n} X_i$ یک متغیر تصادفی دو جمله‌ای با پارامتر n و p می‌باشد حال طبق قضیه حد مرکزی میانگین X_iها به توزیع نرمال میل می‌کند یعنی تابع توزیع $\frac{Y}{n} = \frac{۱}{n}\sum_{۱}^{n} X_i = \overline{X}_n$ وقتی n بزرگ باشد برابر است باک

$$F_{\frac{Y}{n}}(t) = N_Z\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

که در آن امید ریاضی $\frac{Y}{n}$ برابر است با $\mu = p$ و واریانس $\frac{Y}{n}$ نیز برابر با $\sigma^{2} = pq$ می‌باشد حال داریم:

$$F_{\frac{Y}{n}}(t) = N_Z\left(\frac{t-\mu}{\frac{\sqrt{pq}}{\sqrt{n}}}\right) = N_Z\left(\frac{t-p}{\sqrt{\frac{pq}{n}}}\right)$$

حال $F_{\frac{Y}{n}}(t)$ را به $F_Y(t)$ تبدیل می‌کنیم:

$$F_Y(t) = F_{\frac{Y}{n}}\left(\frac{t}{n}\right) = N_Z\left(\frac{\frac{t}{n}-p}{\sqrt{\frac{pq}{n}}}\right) = N_Z\left(\frac{t-np}{\sqrt{npq}}\right)$$

در نتیجه $F_Y(t) = N_Z\left(\frac{t-np}{\sqrt{npq}}\right)$ که با استفاده از روش ذوذنقه‌ای که در فصل قبل نشان دادیم می‌توان مقدار احتمال $p(X = x_\circ)$ را بصورت زیر بدست آورد:

$$p(X = x_\circ) = N\left(\frac{x_\circ + \frac{۱}{۲} - np}{\sqrt{npq}}\right) - N\left(\frac{x_\circ - \frac{۱}{۲} - np}{\sqrt{npq}}\right)$$

فصل نهم

توزیع‌های برخی از آماره‌ها

در فصل قبل با آماره‌ها و تعاریف و مقدمات آنها آشنا شدید، نشان دادیم که آماره‌ها خود متغیر می‌باشند و می‌توان تابع توزیع آنها را معنی نمود، حال در این فصل چندین آماره خاص را مورد بررسی قرار می‌دهیم و میانگین و واریانس آنها را محاسبه می‌کنیم.

۹.۱ توزیع نمونه‌ای میانگین نمونه $\overline{X}$

از یک جامعه با میانگین μ و واریانس δ^2 تعداد n نمونه $X_1, X_2, \dots, X_n$ انتخاب می‌کنیم، آماره $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ را در نظر بگیرید می‌خواهیم میانگین و واریانس $\overline{X}$ را به عنوان یک متغیر تصادفی بدست بیاوریم. طبق قضیه حد مرکزی اگر مقدار n به اندازه کافی بزرگ اختیار شود متغیر تصادفی $Z = \frac{\overline{X} - \mu}{\frac{\delta}{\sqrt{n}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد. از طرفی داریم:

$$\overline{X} = \frac{\delta}{\sqrt{n}} Z + \mu \quad , \quad E[Z] = 0 \quad , \quad \delta_Z^2 = 1$$

بنابراین:

$$E[\overline{X}] = \frac{\delta}{\sqrt{n}} E[Z] + \mu = \frac{\delta}{\sqrt{n}} \times 0 + \mu = \mu$$

بنابراین آماره $\overline{X}$ به عنوان یک متغیر تصادفی اگر تعداد نمونه‌ها زیاد باشد $(n > 30)$ همواره دارای میانگین μ یا همان میانگین جامعه می‌باشد.

۹-۳

مثال ۱: در نقطه‌ای از یک رودخانه در هر دقیقه بطور متوسط تعداد ۲۰ ماهی عبور می‌کند بدیهی است که ماهیگیر تنها در صورتی می‌تواند در هر دقیقه این تعداد ماهی را صید کند که تقریباً تمامی طول رودخانه را با تور ماهیگیری پوشش دهد یا به عبارتی در هر دقیقه از بیشتر نقاط در طول رودخانه نمونه برداری کند که معادل است با انتخاب تعداد زیادی نمونه.

حال واریانس $\overline{X}$ را محاسبه می‌کنیم:

$$var(\overline{X}) = var\left(\frac{\delta}{\sqrt{n}} Z + \mu\right) = var\left(\frac{\delta}{\sqrt{n}} Z\right) = \frac{\delta^2}{\sqrt{n}} var(Z) = \frac{\delta^2}{\sqrt{n}}$$

بنابراین واریانس $\overline{X}$ برابر با $\frac{\delta^2}{\sqrt{n}}$ می‌باشد.

۹-۴

مثال ۲: محصول شیر تولیدی یک کارخانه بطور متوسط ۲۰ روز پس از تاریخ تولید قابل مصرف می‌باشد که با واریانس ۴ روز رخ می‌دهد. می‌خواهیم از این کارخانه ۱۰ پاکت شیر تهیه کنیم. احتمال اینکه این پاکتهای شیر بطور متوسط حداقل ۱۸ روز قابل مصرف باشند چقدر است؟

حل: برای به دست آوردن احتمال می‌بایستی $p(\overline{X} \geq 18)$ را محاسبه کنیم که در واقع از تابع توزیع متغیر تصادفی $\overline{X}$ که در اینجا نرمال μ $\frac{\delta^2}{\sqrt{n}}$ در نظر گرفته می‌شود استفاده می‌کنیم. بنابراین:

$$\mu_{\overline{X}} = 20 \quad , \quad n = 10$$

$$\delta_{\overline{X}}^2 = \frac{\delta^2}{n} = \frac{(4)^2}{10} = \frac{16}{10} \quad \rightarrow \quad \frac{\delta}{\overline{X}} = \sqrt{\frac{16}{10}} = 1/26$$

$$p(\overline{X} \geq ۱۸) = p\left(\frac{\overline{X}-\mu_{\overline{X}}}{\delta_{\overline{X}}} \geq \frac{۱۸-\mu_{\overline{X}}}{\delta_{\overline{X}}}\right)$$
$$= p\left(\frac{\overline{X}-\mu_{\overline{X}}}{\delta_{\overline{X}}} \geq \frac{۱۸-۲۰}{۱/۲۶}\right) = p\left(\frac{\overline{X}-\mu_{\overline{X}}}{\delta_{\overline{X}}} \geq ۱/۵۸\right)$$
$$p = (Z \geq -۱/۵۸) = ۱ - p(Z \leq -۱/۵۸)$$
$$= ۱ - N_Z(-۱/۵۸) = ۱ - ۰/۰۵۷۱ = ۰/۹۴۲۹$$

۹-۵-۱

۹-۲ توزیع نمونه‌ای واریانس نمونه $S^۲$.

حال به محاسبه توزیع نمونه‌ای واریانس $S^۲$ و بدست آوردن میانگین و واریانس توزیع آن می‌پردازیم.

همانطور که از فصل‌های قبل به یاد دارید واریانس نمونه‌ها با $S^۲$ نشان داده می‌شود که بصورت زیر تعریف می‌شود:

$$S^۲ = \frac{۱}{n-۱} \sum_{i=۱}^{n} (X_i - \overline{X})^۲$$

قبل محاسبه توزیع نمونه‌ای $S^۲$ ابتدا میانگین و واریانس آنرا (از آنجا که محاسبه آن ساده می‌باشد) بدست می‌آوریم:

می‌دانیم واریانس $\overline{X}$ برابر $\frac{\delta^۲}{n}$ می‌باشد بنابراین:

$$\text{var}(\overline{X}) = \frac{\delta^۲}{n} \rightarrow E\left[(\overline{X}-\mu)^۲\right] = \frac{\delta^۲}{n}$$
$$E\left[\overline{X}^۲\right] - E^۲[X] = E\left[\overline{X}^۲\right] - \mu^۲ = \frac{\delta^۲}{n} \quad \Rightarrow \quad E\left[\overline{X}^۲\right] = \frac{\delta^۲}{n} + \mu^۲$$

حال داریم:

$$E\left[S^۲\right] = E\left[\frac{۱}{n-۱} \sum_{i=۱}^{n} (X_i - \overline{X})^۲\right] = \frac{۱}{n-۱} E\left[\sum_{i=۱}^{n} (X_i - \overline{X})^۲\right]$$
$$= \frac{۱}{n-۱} E\left[\sum_{۱}^{n} X_i^۲ - ۲\overline{X} \sum_{۱}^{n} X_i + \sum_{۱}^{n} \overline{X}^۲\right]$$
$$= \frac{n}{n-۱} E\left[\frac{۱}{n} \sum_{۱}^{n} X_i^۲ - ۲\overline{X} \frac{۱}{n} \sum_{۱}^{n} X_i + \frac{۱}{n} \sum_{۱}^{n} \overline{X}^۲\right]$$
$$= \frac{n}{n-۱} E\left[\frac{۱}{n} \sum_{۱}^{n} X_i^۲ - ۲\overline{X}^۲ + \overline{X}^۲\right]$$
$$= \frac{n}{n-۱} E\left[\frac{۱}{n} \sum_{۱}^{n} X_i^۲ - \overline{X}^۲\right] = \frac{n}{n-۱}\left(E\left[\frac{۱}{n} \sum_{۱}^{n} X_i^۲\right] - E\left[\overline{X}^۲\right]\right)$$
$$= \frac{n}{n-۱}\left(E[M_۲] - \left(\frac{\delta^۲}{n} + \mu\right)\right)$$

۹-۵-۲

که در آن M_2

گشتاور دوم نمونه‌ها می‌باشد و داریم: $E[M_2]=m_2$ که m_2 نیز گشتاور دوم متغیر تصادفی X می‌باشد. در نهایت داریم:

$$= \frac{1}{n-1}\left(m_2 - \mu^2 - \frac{\delta^2}{n}\right)$$

اما: $m_2 = \delta^2 + \mu^2$

$$= \frac{n}{n-1}\left(\delta^2 + \mu^2 - \mu^2 - \frac{\delta^2}{n}\right) = \frac{n}{n-1}\left(\delta^2 - \frac{\delta^2}{n}\right) = \delta^2$$

به نتیجه جالبی رسیدیم و آن اینکه میانگین واریانس نمونه‌ها برابر با واریانس جامعه می‌باشد در واقع علت اینکه در محاسبه S^2 از $\frac{1}{n-1}$ استفاده می‌کنیم این است که در محاسبات فوق در نهایت جواب ساده شده و δ^2 به عنوان میانگین S^2 ($E[S^2] = \delta^2$) بدست می‌آید. در حالی که اگر S^2 را بصورت زیر تعریف می‌کردیم $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ جواب نهایی $E[S^2]$ برابر با σ^2 $(\frac{n}{n-1})$ می‌شد که بلیل وجود n در جواب نتیجه مطلوبی نیست. حال به جای بدست آوردن توزیع S^2 توزیع S^2 $(\frac{n-1}{\delta^2})$ را بدست می‌آوریم زیرا در محاسبه مقدار میانگین بصورت زیر ساده‌تر می‌شود.

$$E\left[\frac{n-1}{\delta^2} - S^2\right] = \frac{n-1}{\delta^2} E\left[S^2\right] = \frac{n-1}{\delta^2}\ \delta^2 = n-1$$

برای بدست آوردن توزیع S^2 $(\frac{n-1}{\delta^2})$ ابتدا توزیع خی-دو یا مربع‌کای را معرفی می‌کنیم:

۹-۶

۹.۲.۱ توزیع خی دو

فرض کنید Z یک متغیر تصادفی نرمال استاندارد باشد $Z \sim N(0, 1)$ در این صورت متغیر تصادفی $W = Z^2$ را یک متغیر تصادفی χ^2 (خی دو) با یک درجه آزادی می‌نامیم و به صورت χ_1^2 نمایش می‌دهیم که تابع توزیع آن بصورت زیر بدست می‌آید:

قبلاً نشان دادیم که اگر $Y = X^2$ باشد تابع توزیع $F_Y(t)$ برابر است با:

$$F_Y(t) = F_X(\sqrt{t}) - F_X(-\sqrt{t}) \qquad t \geq 0$$
$$= 0 \qquad t < 0$$

حال داریم: $W = Z^2$

$$F_W(t) = F_Z(\sqrt{t}) - F_Z(-\sqrt{t})$$

اما: $F_Z(\sqrt{t}) = N_Z(t)$ و $N_Z(-a) = 1 - N_Z(a)$ بنابراین:

$$F_W(t) = N_Z(\sqrt{t}) - \left[1 - N_Z(\sqrt{t})\right] = 2N_Z(\sqrt{t}) - 1$$

بنابراین با داشتن جدول مقادیر توزیع نرمال استاندارد به راحتی می‌توان مقادیر توزیع χ_1^2 را به دست آورد. مقدار دقیق تابع چگالی χ_1^2 بصورت زیر می‌باشد:

$$f_X(x)=\frac{1}{\sqrt{2\pi}}x^{-\frac{1}{2}}e^{-\frac{x}{2}} \qquad x>0$$

میانگین، واریانس و تابع مولد گشتاور آن عبارتند از:

$$E[W]=1 \qquad var(W)=2 \qquad m_W(t)=\left(\frac{1}{1-2t}\right)^{\frac{1}{2}}$$

۹-۷

حال فرض کنید $X_1, X_2, \ldots, X_n$ نمونه‌هایی تصادفی از یک متغیر تصادفی نرمال X با میانگین μ و واریانس δ^2 باشند در این صورت:

$$Y=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\delta^2}$$

یک متغیر تصادفی χ^2 با n درجه آزادی نامیده می‌شود و با χ_n^2 نمایش داده می‌شود. به بیانی دیگر اگر $Z_1, Z_2, \ldots, Z_n$ تعداد n متغیر تصادفی نرمال استاندارد $N(0,1)$ باشند در این صورت $Y=Z_1^2, Z_2^2, \ldots, Z_n^2$ یک متغیر تصادفی χ^2 با n درجه آزادی است. تابع چگالی متغیر تصادفی χ_n^2 بصورت زیر می‌باشد:

$$f_Y(y)=\frac{1}{\Gamma(\frac{n}{2})}\left(\frac{1}{2}\right)^{\frac{n}{2}}y^{\frac{n}{2}-1}e^{-\frac{y}{2}} \qquad y>0$$

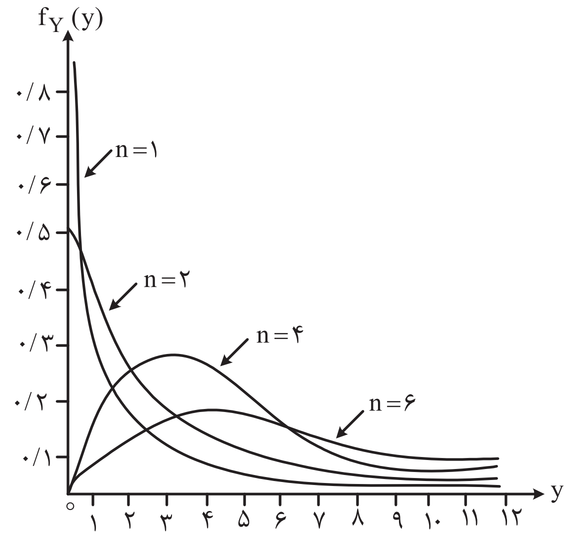
توجه کنید که متغیر تصادفی χ_n^2 حالت خاصی از متغیر تصادفی گاما می‌باشد که در آن $k=\frac{n}{2}$ و $\lambda=\frac{1}{2}$ می‌باشد.

مقادیر میانگین، واریانس و تابع مولد گشتاور متغیر تصادفی χ_n^2 بصورت زیر می‌باشد:

$$E[Y]=n \quad , \quad var(Y)=2n$$

$$m_Y(t)=\left(\frac{1}{1-2t}\right)^{\frac{n}{2}}$$

نمودار منحنی متغیر تصادفی χ_n^2 دارای نقطه ماکزیممی در $y=n-2$ می‌باشد و با افزایش n نمودار منحنی هر چه بیشتر به نمودار نزدیک نزدیکتر می‌شود این مطلب در شکل زیر نشان داده شده است:



نکته: اگر X_n , ... , $X_۲$, $X_۱$ متغیرهای تصادفی خی دو به ترتیب با r_n , ... , $r_۲$, $r_۱$ درجه آزادی باشند و همچنین دو بدو مستقل از یکدیگر باشند در این صورت متغیر تصادفی $Y=\sum_{۱}^{n} X_i$ دارای توزیع خی دو با $R=\sum_{i=۱}^{n} r_i$ درجه آزادی می‌باشد.

۹-۸

مثال ۳: فرض کنید $X_۳$, $X_۲$, $X_۱$ سه متغیر تصادفی مستقل خی دو به ترتیب با ۷ و ۳ و ۱ درجه آزادی باشند، اگر $W = X_۱+X_۲+X_۳$ مطلوبست میانگین و واریانس متغیرهای تصادفی W؟

حل: متغیر تصادفی W یک متغیر تصادفی خی دو با ۱۱ = ۷+۳+۱ درجه آزادی می‌باشد بنابراین میانگین و واریانس آن عبارتست از:

$E[W] = ۱۱$

$var(W) = ۲ \times ۱۱ = ۲۲$

حال به قضیه بسیار مهم توجه کنید:

قضیه: اگر X_n , ... , $X_۲$, $X_۱$ یک نمونه تصادفی از متغیر تصادفی نرمال X با میانگین μ و واریانس $\delta^۲$ باشند آنگاه:

الف) $\overline{X}$ و $S^۲$ از یکدیگر مستقل هستند.

ب) $\frac{(n-۱)}{\delta^۲} S^۲$ یک متغیر تصادفی $\chi^۲$ با $n-۱$ درجه آزادی می‌باشد توجه کنید که $(\frac{n-۱}{\delta^۲}) S^۲$ برابر است با:

$$(\frac{n-۱}{\delta^۲}) S^۲ = \frac{n-۱}{\delta^۲} \frac{۱}{n-۱} \sum_{۱}^{n} (X_i-\overline{X})^۲ = \sum_{۱}^{N} \frac{(X_i-\overline{X})^۲}{\delta^۲}$$

و این با آنچه در تعریف متغیر تصادفی $\chi_n^۲$ متفاوت می‌باشد زیرا در تعریف از میانگین جامعه μ استفاده نمودیم: $\sum_{۱}^{n} \frac{(X_i-\mu)^۲}{\delta^۲}$.

۹-۹

حال نشان می‌دهیم که چرا با جایگزین نمودن $\overline{X}$ به جای μ یک درجه آزادی از n کم می‌شود و درجه آزادی $\frac{(n-۱)}{\delta^۲} S^۲$ برابر با $n-۱$ می‌شود.

$$\sum_{i=۱}^{n} (X_i-\mu)^۲ = \sum_{۱}^{n} \left[(X_i-\overline{X}) + (\overline{X}-\mu)\right]^۲$$

$$=\sum_{۱}^{n} (X_i-\overline{X})^۲ + \sum_{۱}^{n} (\overline{X}-\mu)^۲ + ۲(\overline{X}-\mu) \sum_{۱}^{n} (X_i-\overline{X})$$

$$=\sum_{۱}^{n} (X_i-\overline{X})^۲ + n(\overline{X}-\mu)^۲$$

در نتیجه بدست می‌آید:

$$\frac{\sum (X_i-\mu)^۲}{\delta^۲} = \frac{\sum (X_i-\mu)^۲}{\delta^۲} + \frac{(\overline{X}-\mu)^۲}{\frac{\delta^۲}{n}}$$

حال می‌دانیم که عبارت سمت چپ تساوی فوق یک متغیر تصادفی خی دو با n درجه آزادی است و عبارت $\frac{(\overline{X}-\mu)^{۲}}{\frac{\delta^{۲}}{n}}$ نیز یک متغیر تصادفی $X^{۲}$ با ۱ درجه آزادی است بنابراین جمله $\frac{\sum (X_i-\mu)^{۲}}{\delta^{۲}}$ می‌بایستی یک متغیر تصادفی $\chi^{۲}$ با $n-۱$ درجه آزادی باشد.

مثال ۴: از یک جامعه نرمال با میانگین μ و واریانس $\delta^{۲}$ یک نمونه تصادفی ۱۱تایی انتخاب می‌کنیم. متغیر تصادفی $Y=\frac{۱۰\,S^{۲}}{\delta^{۲}}$ را در نظر بگیرید مطلوبست:

الف) محاسبه $p(۳/۹۴ < Y < ۱۸/۳)$.

ب) با استفاده از احتمال فوق رابطه بین $S^{۲}$ و $\delta^{۲}$ را در جامعه معین کنید.

حل: الف) ابتدا احتمال را بصورت زیر ساده می‌کنیم:

$$p(۳/۹۴ < Y < ۱۸/۳) = p(Y < ۱۸/۳) - p(Y < ۳/۹۴)$$

حال از روی جدول مقادیر احتمال برای متغیر تصادفی خی دو و با توجه به اینکه $Y=\frac{۱۰\,S^{۲}}{\delta^{۲}}$ یک متغیر تصادفی خی دو با ۱۰ درجه آزادی است مقادیر احتمال بصورت زیر بدست می‌آیند:

$$p(Y < ۱۸/۳) - p(Y < ۳/۹۴) = ۰/۹۵ / ۰/۰۵ = ۰/۹$$

ب) می‌دانیم $p(۳/۹۴ < Y < ۱۸/۳) = ۰/۹$ بنابراین:

$$p(۳/۹۴ < Y = \frac{۱۰\,S^{۲}}{\delta^{۲}} < ۱۸/۳) = p(\frac{۳/۹۴\,\delta^{۲}}{۱۰} < S^{۲} < \frac{۱۸/۳\,\delta^{۲}}{۱۰})$$

$$\Rightarrow p(۰/۳۹۴\ \delta^{۲} < S^{۲} < ۱/۸۳\ \delta^{۲}) = ۰/۹$$

حال می‌بینیم که $S^{۲}$ در بازه‌ای نزدیک $\delta^{۲}$ قرار دارد یعنی اگر از یک جامعه نرمال با میانگین μ و واریانس $\delta^{۲}$ تعداد ۱۱ نمونه انتخاب کنیم به احتمال بسیار زیاد $(۰/۹)$ مقدار واریانس نمونه‌ها نزدیک به واریانس جامعه می‌باشد.

۹-۱۱-۱

۹ . ۳ تابع توزیع آمارهٔ مرتب

قبلاً آمارهٔ مرتب را معرفی نمودیم، می‌دانیم آماره‌های مرتب خود یک متغیر تصادفی می‌باشند. بنابراین می‌توانیم تابع توزیع آنها را بدست بیاوریم.
فرض کنید $X_۱, X_۲, \dots, X_n$ تعداد n نمونه تصادفی از متغیر تصادفی X باشند در این صورت تابع توزیع برای آمارهٔ مرتب $X_{(۱)}, X_{(۲)}, \dots, X_{(n)}$ عبارتست از:

$$F_{X(j)}(t) = \sum_{k=j}^{n} \binom{n}{k} (F_X(t))^k (۱-F_X(t))^{n-k} \qquad t \in R$$

$F_X(t)$ تابع توزیع متغیر تصادفی X می‌باشد.

$F_{X(j)}(t)$ به صورت زیر بدست می‌آید:

اگر حداقل j متغیر تصادفی نمونه، کمتر یا مساوی با t باشند در این صورت $X_{(j)}$ یا jامین آمارهٔ مرتب کمتر یا مساوی با t می‌باشد. نمام متغیرهای تصادفی نمونه مستقل می‌باشند و برای تمام متغیرهای تصادفی نمونه $F_X(t)$ برابر است با:

احتمال کمتر یا مساوی با t بودن. بنابراین تعداد متغیرهای تصادفی نمونه که کمتر یا مساوی با t می‌باشند یک متغیر تصادفی دو جمله‌ای با پارامترهای n و $p = F_X(t)$ می‌باشد، در نتیجه:

$$f_{X(j)}(t) = \sum_{k=j}^{n} \binom{n}{k} (F_X(t))^{j-۱}(۱-F_X(t))^{n-j} f_X(t)$$

با قرار دادن $j = ۱$ و $j = n$ تابع توزیع و چگالی برای $\min(X_j)$ و $\operatorname{Max}(X_j)$ بصورت زیر بدست می‌آید:

$$\operatorname{Min}(X_j) \text{ تابع توزیع}: F_{X(۱)}(t) = \sum_{k=۱}^{n} \binom{n}{k} (F_X(t))^k (۱-F_X(t))^{n-k} = ۱-(۱-F_{X(۱)}(t))^n$$

$$\min(X_j) \text{ تابع چگالی}: f_{X(۱)}(t) = \binom{n}{۱} (f_X(t))^{۱-۱} (۱-F_X(t))^{n-۱} f_X(t) = n\,(۱-F_X(t))^{n-۱} f_X(t)$$

همانطور که ملاحظه می‌کنید رابطه $f_X(t) = \dfrac{dF_{X(۱)}(t)}{dt}$ نیز بین تابع توزیع و تابع چگالی برقرار است.

$$\operatorname{Max}(X_j) \text{ تابع توزیع}: F_{X(n)}(t) = \sum_{k=n}^{n} \binom{n}{k} (F_X(t))^k (۱-F_X(t))^{n-k} = (F_X(t))^n$$

$$\operatorname{Max}(X_j) \text{ تابع چگالی}: f_{X(n)}(t) = n \binom{n}{n} (f_X(t))^{n-۱} (۱-F_X(t))^{n-n} f_X(t) = n\,(F_X(t))^{n-۱} f_X(t)$$

مثال ۵: متغیر تصادفی X را مدت زمان بین هر دو تماس تلفنی در یک شرکت در نظر می‌گیریم که بطور میانگین هر ۱۰ دقیقه یکبار رخ می‌دهد اگر بطور تصادفی ۴ نمونه از متغیر تصادفی X اختیار کنیم مطلوبست:

الف) تابع توزیع آمارهٔ مرتب نمونه تصادفی.

ب) تابع توزیع و چگالی ماکزیمم و مینیمم نمونه‌ها.

ج) احتمال اینکه در نمونه‌های مشاهده شده کوچکترین بازه زمانی بین دو تماس بیشتر از ۲ دقیقه باشد چقدر است؟

حل: الف) ابتدا توجه کنید که X یک متغیر تصادفی نمایی با پارامتر $\lambda = \dfrac{۱}{۱۰}$ می‌باشد. بنابراین تابع توزیع X عبارتست از:

$$F_X(t) = ۱ - e^{-\lambda t} = ۱ - e^{-\frac{۱}{۱۰}t} \qquad t > ۰$$

بنابراین تابع توزیع آمارهٔ مرتب برای ۴ نمونه مشاهده شده عبارتست از:

$$F_{X(j)}(t) = \sum_{k=j}^{۴} \binom{n}{k} (۱-e^{\frac{-۱}{۱۰}t})^k (e^{-\frac{۱}{۱۰}t})^{n-k}$$

۴-۱۲-۲

ب) ماکزیمم آماره‌های مرتب در این مثال $X_{(۴)}$ می‌باشد. که تابع توزیع و چگالی آن بدست می‌آید:

$$F_{X(۴)}(t) = (F_X(t))^۴ = (۱-e^{-\frac{۱}{۱۰}t})^۴ \qquad t > ۰$$

$$F_{X(۴)}(t) = ۴\,(۱-e^{-\frac{۱}{۱۰}t})^۳ \frac{۱}{۱۰} e^{-\frac{۱}{۱۰}t}$$

مینیمم آماره‌های مرتب $X_{(۱)}$ می‌باشد که تابع توزیع و چگالی آن نیز بصورت زیر می‌باشد:

$$F_{X(1)}(t)=1-(1-F_X(t))^4=1-(1-1+e^{\frac{1}{10}t})^4=1-e^{-\frac{4}{10}t}$$

$$F_{X(1)}(t)=4\,(1-1+e^{-\frac{1}{10}t})^3\,\frac{1}{10}\,e^{-\frac{1}{10}t}=\frac{4}{10}\,e^{-\frac{3}{10}t}\,e^{-\frac{1}{10}t}=\frac{4}{10}\,e^{-\frac{4}{10}t}$$

توجه کنید که توزیع کوچکترین آماره مرتب در این مثال خود یک متغیر تصادفی نمایی با پارامتر $\lambda=\frac{4}{10}$ می‌باشد.

ج) احتمال اینکه کوچکترین بازه زمانی بین دو تماس در نمونه‌ها بیشتر از ۲ دقیقه باشد برابر است با: $p(X_{(1)}>2)$ و داریم:

$$p(X_{(1)}>2)=1-p(X_{(1)}<2)=1-F_{X(1)}(t=2)$$

$$=1-(1-e^{-\frac{4}{10}\times 2})=e^{-0/8}\approx 0/449$$

۹-۱۳

مثال ۶: برای آماره مرتب $X_{(1)},X_{(2)},\ldots,X_{(n)}$ از یک متغیر تصادفی پیوسته X مقادیر مورد انتظار برای احتمال بزرگترین عضو نمونه بدست بیاورید:

حل: می‌بایستی $E\left[F_X(X_{(n)})\right]$ را بدست بیاوریم داریم:

$$f_{X(n)}(t)=n\,(F_X(t))^{n-1}\,f_X(t)$$

$$E\left[f_x(x_{(n)})\right]=\int_{-\infty}^{+\infty} f_X(t)\;n\;(F_X(t))^{n-1}\,f_X(t)$$

$$=\int_{-\infty}^{+\infty}(f_X(t))^{n+1}\;\int_{-\infty}^{+\infty}=\frac{n}{n+1}-0=\frac{n}{n+1}$$

۹-۱۴

۹ . ۴ توزیع نمونه‌ای $\dfrac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$.

در این فصل نشان دادیم که توزیع $\dfrac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ نرمال استاندارد است، در صورتی که $X_1,X_2,\ \ldots\ ,X_n$ از جامعه‌ای نرمال با میانگین μ و واریانس δ^2 انتخاب شده باشد. همینطور نشان دادیم که در صورتی که تعداد نمونه‌ها زیاد باشد $(n>30)$ توزیع $\dfrac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ مستقل از توزیع جامعه دارای توزیع نرمال استاندارد می‌باشد.

حال اگر واریانس یک جامعه مجهول باشد بناچار می‌بایستی از واریانس نمونه S^2 در محاسبه $\dfrac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ استفاده کنیم. به این ترتیب می‌بایستی ابتدا تابع توزیع $\dfrac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ را بدست بیاوریم.

هر گاه Z یک متغیر تصادفی نرمال استاندارد باشد $(Z \sim N(۰,۱))$ و Y نیز یک متغیر تصادفی خی دو با n درجه آزادی باشد $(Y \sim \chi^{۲}_{n})$ و Z و Y از یکدیگر مستقل باشند در این صورت متغیر تصادفی $T = \dfrac{Z}{\sqrt{\dfrac{Y}{n}}}$ را یک متغیر تصادفی t با n درجه آزادی می‌نامیم و بصورت $T \sim t_{(n)}$ نمایش می‌دهیم.

۹-۱۵

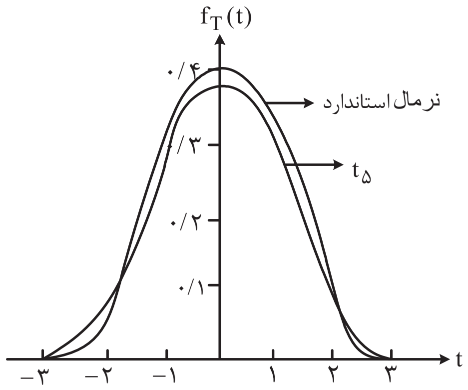
تابع چگالی متغیر تصادفی T بصورت زیر می‌باشد:

$$f_T(t) = \frac{T\left(\frac{n+۱}{۲}\right)}{T\left(\frac{n}{۲}\right)\sqrt{n\pi}} \left[۱ + \frac{t^{۲}}{n}\right]^{-\frac{n+۱}{۲}} \qquad -\infty < t < +\infty$$

میانگین و واریانس متغیر تصادفی t_n عبارتند از:

$$E[t_n] = ۰ \qquad , \qquad var(t_n) = \frac{n}{n-۲} \qquad n > ۲$$

نمودار منحنی متغیر تصادفی t_n بسیار مشابه متغیر تصادفی نرمال استاندارد می‌باشد به شکل زیر که متغیر تصادفی $t_۵$ را با نرمال استاندارد مقایسه می‌کند توجه کنید.



در واقع با افزایش درجه آزادی متغیر تصادفی t_n نمودار منحنی هر چه بیشتر به نمودار منحنی نرمال استاندارد نزدیک می‌شود.

۹-۱۶

توجه: اگر $X \sim t_۱$

توزیع $t_۱$ بصورت زیر می‌باشد:

$$f_X(x) = \frac{۱}{\pi(۱+x^{۲})} \qquad -\infty < x +\infty$$

حال توجه کنید که معکوس توزیع X یعنی $Y = \dfrac{۱}{X}$ برابر است با:

$$f_Y(y) = \frac{۱}{\pi(۱+y^{۲})}$$

به عبارت دقیق‌تر عکس یک متغیر $t_۱$ بازهم $t_۱$ می‌باشد.

حال به محاسبه توزیع $\frac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ می‌پردازیم. طبق تعریف می‌دانیم اگر $Z \sim N(0, 1)$ و $Y \sim \chi_n^2$ متغیر تصادفی $T = \frac{Z}{\sqrt{\frac{Y}{n}}}$ یک متغیر تصادفی t_n می‌باشد حال در نظر بگیرید: $Z = \frac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}$ و $Y = \frac{(n-1)S^2}{\delta^2}$ قبلاً نشان دادیم که Z و Y به ترتیب متغیرهای تصادفی نرمال استاندارد و χ_{n-1}^2 می‌باشند در نتیجه متغیر تصادفی $T = \frac{Z}{\sqrt{\frac{Y}{n-1}}}$ می‌بایستی یک متغیر تصادفی t با $n-1$ درجه آزادی می‌باشد. با ساده نمودن عبارت T داریم:

$$T=\frac{\frac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}}{\sqrt{\frac{(n-1)S^2}{\delta^2(n-1)}}} = \frac{\frac{\overline{X}-\mu}{\frac{\delta}{\sqrt{n}}}}{\sqrt{\frac{S^2}{\delta^2}}} = \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$$

در نتیجه متغیر تصادفی $T=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ یک متغیر تصادفی t با $n-1$ درجه آزادی می‌باشد.

توجه کنید در صورتی که نمونه‌های تصادفی $X_1, X_2, \ldots, X_n$ از جامعه‌ای نرمال باشند رابطه $T=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ یک متغیر تصادفی t_{n-1} می‌باشد.

اما اگر جامعه نرمال نباشد یا توزیع آنرا ندانیم یا n به اندازه کافی بزرگ $(n > 30)$ باشد آنگاه باز هم $T=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ یک متغیر تصادفی t_{n-1} می‌باشد حتی در این حالت می‌توان با توجه به تشابه نمودار منحنی‌های t_{n-1} و نرمال استاندارد می‌توان گفت که توزیع $T=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ برای $n > 30$ همان نرمال استاندارد می‌باشد.

۹-۱۷

محاسبه مقادیر احتمال t_n.

برای بدست آوردن مقادیر مختلف احتمال t_n از جدول احتمالات $F_t(x)$ استفاده می‌کنیم بطوریکه $F_t(x)$ برابر است با:

$$F_t(x) = \int_{-\infty}^{x} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} (1+x^2)^{\frac{-(n+1)}{2}} \, dx = 1-\alpha$$

$$F_t(x) = p(T < x)$$

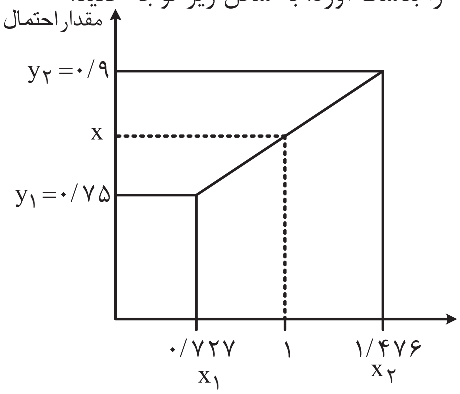
مقادیر هر یک از ردیف یا معادل با درجه آزادی و مقادیر ستون‌ها معادل با مقدار احتمال $F_t(x)$ می‌باشد. جدول زیر بخشی از جدول تابع توزیع t را نشان می‌دهد:

n \ $1-\alpha$	۰/۶	۰/۷۵	۰/۹۰	۰/۹۵	۰/۹۷۵
۱	۰/۳۲۵	۱/۰۰	۳/۰۷۸	۶/۳۱۴	۱۲/۷۰۶
۲	۰/۲۸۹	۰/۸۱۶	۱/۸۸۶	۲/۹۲	۴/۳۰۳
۳	۰/۲۷۷	۰/۷۶۵	۱/۶۳۸	۲/۳۵۳	۳/۱۸۲

حال با توجه به جدول به عنوان مثال $p(t_{۲} < ۲/۹۲)$ برابر با ۰/۹۵ می‌باشد توجه کنید که مقادیر احتمال از عناوین هر یک از ستون‌ها بدست می‌آیند. برای بدست آوردن سایر مقادیر احتمال از تقریب زدن استفاده می‌کنیم.

۹-۱۸

به عنوان مثال برای بدست آوردن مقدار احتمال $p(t_{۵} < ۱)$ از آنجا که در جدول عدد ۱ در ردیف پنجم ما بین دو مقدار ۱/۴۷۶ و ۰/۷۲۷ می‌باشد می‌بایستی با استفاده از روش خطی مقدار این احتمال را تقریب بزنیم در این روش ما بین دو نقطه $(۰/۷۵ , ۰/۷۲۷)$ و $(۰/۹ , ۱/۴۷۶)$ یک خط در نظر می‌گیریم، با نوشتن معادله خط به سادگی می‌توان مقادیر احتمال ما بین این دو نقطه را بدست آورد. به شکل زیر توجه کنید.



معادله خط عبارتست از: $y = \dfrac{y_{۲} - y_{۱}}{x_{۲} - x_{۱}}(x - x_{۱}) + y_{۱}$

بنابراین:

$$y = \frac{۰/۹ - ۰/۷۵}{۱/۴۷۶ - ۰/۷۲۷}\ (x - ۰/۷۲۷) + ۰/۷۵$$

حال برای بدست آوردن احتمال هر یک از مقادیر ما بین $x_{۱} , x_{۲}$ کافیست آن را به جای x در معادله بالا قرار دهیم تا y که مقدار احتمال تقریبی می‌باشد بدست آید.

برای احتمال $p(t_{۵} < ۱)$ داریم:

$$p(t_{۵} < ۱) = y = \frac{۰/۹ - ۰/۷۵}{۱/۴۷۶ - ۰/۷۲۷}\ (۱ - ۰/۷۲۷) + ۰/۷۵ = ۰/۸$$

مثال: تعداد مراجعه کنندگان به یک فروشگاه در طول روز یک متغیر تصادفی نرمال می‌باشد. مدیر فروشگاه می‌داند که بطور متوسط روزانه ۱۲۰ نفر به فروشگاه مراجعه می‌کنند. تعداد مراجعه کنندگان به فروشگاه در طول یک هفته بصورت زیر بدست آمده است:

۹۰ , ۱۱۰ , ۷۵ , ۱۳۰ , ۱۵۰ , ۱۲۰ , ۱۰۰

مطلوبست:

الف) احتمال اینکه میانگین تعداد مراجعه کنندگان حداقل ۱۲۵ نفر در روز باشد؟

ب) احتمال اینکه تفاوت میانگین نمونه‌ها با میانگین جامعه حداکثر ۲ واحد باشد؟

حل: الف) ابتدا از روی نمونه‌های بدست آمده مقدار $\overline{X}$ و $S^۲$ را بدست می‌آوریم. داریم:

$$\overline{X} = \frac{۱}{n}\sum_{i=۱}^{n} x_i = \frac{۱}{۷}(۹۰ + ۱۱۰ + ۷۵ + ۱۳۰ + ۱۵۰ + ۱۲۰ + ۱۰۰) = ۱۱۰/۷$$

$$S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i - \overline{X})^۲ = \frac{۱}{۶}[(۹۰ - ۱۱۰/۷)^۲ + (۷۵ - ۱۱۰/۷)^۲ + (۱۳۰ - ۱۱۰/۷)^۲ + (۱۵۰ - ۱۱۰/۷)^۲$$

$$(۱۲۰ - ۱۱۰/۷)^۲ + (۱۰۰ - ۱۱۰/۷)^۲] = ۶۳۶/۹ \qquad \Rightarrow \qquad S = ۲۵/۲$$

حال می‌بایستی مقدار احتمال $p(\overline{X} > ۱۲۵)$ را بدست بیاوریم: (با توجه به اینکه واریانس جامعه را نمی‌دانیم بنابراین می‌بایستی از توزیع t استفاده کنیم.)

$$p(\overline{X} > ۱۱۰) = p(\overline{X} - \mu > ۱۲۵ - ۱۲۰) = p\left(\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} > \frac{۱۲۵ - ۱۲۰}{\frac{۲۵/۲}{\sqrt{۷}}}\right)$$

$$= p\left(\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} > \frac{۵}{۹/۵۲}\right) = p(t_۶ > ۰/۵۲۵) = ۱ - p(t_۶ < ۰/۵۲۵) = ۱ - ۰/۶۸ = ۰/۳۲$$

ب) در این حالت می‌بایستی احتمال $p(\overline{X} - \mu < ۲)$ را بدست بیاوریم:

$$p(\overline{X} - \mu < ۲) = p\left(\frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} < \frac{۲}{۹/۵۲}\right) = p(t_۶ < ۰/۲۱) \approx ۰/۵۸$$

۹-۲۰

توزیع نمونه‌ای نسبت واریانسهای نمونه

اگر از دو جامعه نرمال دو نمونه مستقل به اندازه‌های $n_۱$ و $n_۲$ بگیریم برای بدست آوردن نسبت واریانسهای دو نمونه ($\frac{S_۱}{S_۲}$ یا $\frac{S_۲}{S_۱}$) از توزیعی به نام F (فیشر) استفاده می‌کنیم.

تعریف: اگر U یک متغیر تصادفی خی دو با m درجه آزادی ($U \sim \chi^۲_m$) و V یک متغیر تصادفی خی دو با n درجه آزادی ($V \sim \chi^۲_m$) باشند. در این صورت توزیع متغیر تصادفی $F = \frac{\frac{U}{m}}{\frac{V}{n}}$ را یک متغیر F با m و n درجه آزادی می‌نامند و با نماد $F \sim F_{m,n}$ نمایش می‌دهند. تابع چگالی احتمال توزیع $F_{m,n}$ بصورت زیر می‌باشد:

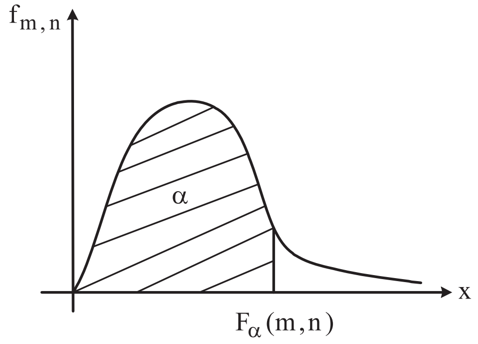
$$f_F(x) = \frac{\Gamma(\frac{m+n}{۲})(\frac{m}{n})^{\frac{m}{۲}}}{\Gamma(\frac{m}{n})\,\Gamma(\frac{n}{۲})}\, x^{\frac{m}{۲}-۱}\,(۱ + \frac{m}{n}x)^{-\frac{m+n}{۲}} \qquad x > ۰$$

۹-۲۱

مقادیر میانگین و واریانس توزیع $F_{m,n}$ عبارتند از:

$$E[X] = \frac{n}{n-۲} \quad n > ۲ \quad , \quad var(X) = ۲\left(\frac{۱}{n} + \frac{۱}{m}\right)$$

شکل نمودار تابع توزیع $F_{m,n}$ بصورت زیر می‌باشد:



حال برای بدست آوردن توزیع نسبت $\frac{S_1^2}{S_2^2}$ بصورت زیر عمل می‌کنیم:

Y و X را دو جمعیت نرمال با واریانسهای δ_1^2 و δ_2^2 در نظر می‌گیریم

فرض: $X \sim N(\mu_1 , \delta_1^2) \qquad ; \qquad Y \sim N(\mu_2 , \delta_2^2)$

می‌دانیم:

$$V = \frac{(n_1-1)\,S_1^2}{\delta_1^2} \sim \chi^2_{(n_1-1)}$$

$$V = \frac{(n_2-1)\,S_2^2}{\delta_2^2} \sim \chi^2_{(n_2-1)}$$

حال با توجه به تعریف تابع توزیع F داریم:

$$F = \frac{\frac{U}{n-1}}{\frac{V}{n_2-1}} = \frac{\frac{(n_1-1)\,S_1^2}{\delta_1^2\,(n_1-1)}}{\frac{(n_2-1)\,S_2^2}{\delta_2^2\,(n_2-1)}} = \frac{\frac{S_1^2}{\delta_1^2}}{\frac{S_2^2}{\delta_2^2}} = \frac{S_1^2\,\delta_2^2}{S_2^2\,\delta_1^2} \sim F_{n_1-1\,,\,n_2-1}$$

۹-۲۲

مثال: اگر X یک متغیر تصادفی t با n درجه آزادی باشد مربع متغیر تصادفی X دارای چه توزیعی می‌باشد؟

حل: فرض می‌کنیم: $Z \sim N(0, 1)$ و $Y = \chi^2_n$ در این صورت داریم:

$$X = t_n \sim \frac{Z}{\sqrt{\frac{Y}{n}}}$$

$$\Rightarrow \quad X^2 = \frac{Z^2}{\frac{Y}{n}}$$

از آنجا که مربع یک متغیر تصادفی نرمال استاندارد یک متغیر تصادفی خی دو با یک درجه آزادی است. بنابراین:

$$X^2 = \frac{\chi_1^2}{\frac{Y}{n}} = \frac{\frac{\chi_1^2}{1}}{\frac{\chi_n^2}{n}} \sim F_{1,n}$$

بنابراین مربع یک متغیر تصادفی t با n درجه آزادی یک متغیر تصادفی F با $n,1$ درجه آزادی می‌باشد.

۹-۲۳

مثال: تعداد تصادفات در دو شهر A و B متغیر تصادفی نرمال به ترتیب با واریانس ۲۰ و ۲۲ تصادف در روز می‌باشد. اگر از دو جامعه فوق دو نمونه به ترتیب با حجم ۲۱ و ۳۱ نمونه انتخاب کرده باشیم و S_B^2 , S_A^2 به ترتیب واریانسهای نمونه‌های گرفته شده از شهر A و B باشد محاسبه احتمال اینکه S_A^2 حداقل ۱/۵ برابر S_B^2 باشد؟

حل: داریم:

$$\sigma_A^2 = 20 \qquad ; \qquad \sigma_B^2 = 22$$

می‌بایستی احتمال $p(S_A^2 \geq 1/5\, S_B^2)$ را محاسبه کنیم داریم:

$$p(S_A^2 \geq 1/5\, S_B^2) = p\left(\frac{S_A^2}{S_B^2} \geq 1/5\right)$$

$$p\left(\frac{\frac{S_A^2}{\sigma_A^2}}{\frac{S_B^2}{\sigma_B^2}} \geq 1/5 \left(\frac{22}{20}\right)\right) = p\left(\frac{S_A^2\,\sigma_B^2}{S_B^2\,\sigma_A^2} \geq 1/65\right)$$

$$= 1 - p(F_{20,30} < 1/65) \approx 1 - 0/9 = 0/1$$

۹-۲۴

تقریب توزیع F به توزیع خی دو.

در صورتی که در یک متغیر تصادفی $F_{m,n}$ شرط $n \to \infty$ برقرار باشد می‌توانیم برای محاسبه مقادیر احتمال mF از متغیر تصادفی χ_m^2 استفاده نمایم. به عبارت دیگر $mF_{m,n}$ با شرط $n \to \infty$ دارای توزیع χ_m^2 می‌باشد.

توزیع نمونه‌ای اختلاف میانگین‌ها

با استفاده از توزیع F احتمالات مربوط به نسبت واریانسهای دو نمونه را بدست آوریم حال می‌خواهیم برای دو نمونه گرفته شده از دو جامعه نرمال توزیع اختلاف میانگین‌ها را بدست آوریم.

از دو جامعه به ترتیب با میانگین‌های μ_2, μ_1 و واریانسهای σ_2^2, σ_1^2 دو نمونه به اندازه n_2, n_1 انتخاب می‌کنیم. در این صورت برای میانگین نمونه‌ها داریم:

$$\overline{X}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$\overline{X}_2 \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

از آنجا که $\overline{X}_2, \overline{X}_1$ از یکدیگر مستقل می‌باشد و $\overline{X}_1 - \overline{X}_2$ نیز یک ترکیب خطی‌از دو متغیر تصادفی نرمال می‌باشد بنابراین $\overline{X}_1 - \overline{X}_2$ نیز نرمال است و توزیع آن عبارتست از:

$$\overline{X}_1 - \overline{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

با تبدیل آن به متغیر تصادفی نرمال استاندارد داریم:

$$Z=\frac{(\overline{X}_۱-\overline{X}_۲)-(\mu_۱-\mu_۲)}{\sqrt{\frac{\sigma_۱^۲}{n_۱}+\frac{\sigma_۲^۲}{n_۲}}}\sim N(۰,۱)$$

حتی اگر دو جمعیت نرمال نباشند با استفاده از قضیه حد مرکزی می‌دانیم که اگر داشته باشیم $n_۱,n_۲\geq ۳۰$ باز هم هر یک از $\overline{X}_۱,\overline{X}_۲$ دارای توزیع تقریبی نرمال می‌باشند. و در نتیجه $\overline{X}_۱-\overline{X}_۲$ نیز دارای توزیع نرمال خواهد بود. و روابط فوق برقرار خواهند بود.

۹-۲۵

مثال: در مثال قبل اگر متوسط تعدد تصادفات در شهر A ۱۰۰ و در شهر B ۱۲۰ تصادف در روز باشند. اگر دو نمونه به حجم ۳۰ انتخاب کنیم در این صورت مطلوبست احتمال اینکه میانگین تعداد تصادفات در شهر B حداقل ۱۸ تصادف از شهر A بیشتر باشد

حل: با توجه به مفروضات مساله داریم:

$$\mu_A=۱۰۰\qquad \sigma_A^۲=۲۰\qquad n_A=۳۰$$
$$\mu_B=۱۲۰\qquad \sigma_B^۲=۲۲\qquad n_B=۳۰$$
$$\overline{X}_A-\overline{X}_B\sim N(۱۰۰-۱۲۰,\frac{۲۰}{۳۰}+\frac{۲۲}{۳۰})\sim N(-۲۰,۱/۴)$$

حال می‌بایستی احتمال $p(\overline{X}_B\geq\overline{X}_A+۱۸)$ را محاسبه کنیم:

$$p(\overline{X}_B\geq\overline{X}_A+۱۸)=p(\overline{X}_A-\overline{X}_B\leq -۱۸)$$
$$=p(\overline{X}_A-\overline{X}_B-(\mu_A-\mu_B))\leq -۱۸-(۱۰۰-۱۲۰))$$
$$=p\left(\frac{\overline{X}_A-\overline{X}_B-(\mu_A-\mu_B)}{\sqrt{\frac{\sigma_A^۲}{n_A}+\frac{\sigma_B^۲}{n_B}}}\leq\frac{۲}{۱/۱۸}\right)$$
$$=p(Z\leq ۱/۶۹)=۰/۹۷۵$$

۹-۲۶

اگر در محاسبه توزیع اختلاف میانگین نمونه‌ها واریانس دو جامعه نامعلوم اما مساوی باشد $(\sigma_۱^۲=\sigma_۲^۲=\sigma^۲)$ می‌توانیم از واریانس نمونه‌ها $(S_۲^۲,S_۱^۲,\ldots)$ به عنوان تخمین $\sigma^۲$ استفاده کنیم. در این حالت از میانگین وزنی این دو واریانس برای برآورد $\sigma^۲$ استفاده می‌کنیم:

$$S_p^۲=\frac{(n_۱-۱)S_۱^۲+(n_۲-۱)S_۲^۲}{(n_۱+n_۲-۲)}$$

در این حالت اگر دو جامعه نرمال باشند داریم:

$$\frac{(n_۱+n_۲-۲)S_p^۲}{\sigma^۲}=\frac{(n_۱-۱)S_۱^۲}{\sigma^۲}+\frac{(n_۲-۱)S_۲^۲}{\sigma^۲}=\chi^۲_{(n_۱+n_۲-۲)}$$

حال با توجه به توزیع t داریم:

$$T=\frac{(\overline{X}_۱-\overline{X}_۲)-(\mu_۱-\mu_۲)}{S_p\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}}\sim t_{(n_۱+n_۲-۲)}$$

بنابراین در محاسبه احتمال مربوط به نمونه‌های بدست آمده از دو جامعه با واریانس نامعلوم اما مساوی می‌بایستی از توزیع t با $n_۱+n_۲-۲$ درجه آزادی استفاده کنیم.

فصل دهم

S-۱

۱- برآورد

در فصل قبل با نمونه‌گیری از یک جامعه با معلوم بودن تابع چگالی و پارامترهای جامعه مثل میانگین و واریانس آشنا شدید. همچنین نحوه بدست آوردن احتمالات مربوط به آماره‌های خاص مثل میانگین، نسبت واریانسها و ..... را نشان دادیم.

در عمل بسیاری ار اوقات می‌دانیم که یک جامعه مثلاً نرمال می‌باشد اما مقدار دقیق پارامترهای جامعه را که μ و $\sigma^{۲}$ می‌باشند نمی‌دانیم، در این حالت با نمونه‌گیری از جامعه و با استنباط آماری روی نمونه‌ها می‌توانیم مقادیر پارامترهای مجهول جامعه را با دقت زیادی بدست بیاوریم. بنابراین در این فصل به بحث پیرامون روشهای برآوردیابی می‌پردازیم.

۲ . ۱ . ۱ برآورد نقطه‌ای

با نمونه‌گیری از جامعه دو پارامتر میانگین نمونه‌ها و واریانس نمونه‌ها را بدست می‌آوریم. این دو پارامتر نقش اصلی را در برآورد میانگین و واریانس جامعه بازی می‌کنند. بطور کلی با استفاده از میانگین و واریانس نمونه‌ها به دو طریق می‌توان پارامترهای مجهول جامعه را برآورد نمود. در روش اول که به برآورد نقطه‌ای معروف است با برابر قرار دادن میانگین و واریانس نمونه‌ها با میانگین جامعه، این دو پارامتر مجهول را بدست می‌آوریم. اما در روش دوم که به برآورد فاصله‌ای معروف است برای پارامتر مجهول جامعه مثل میانگین یک بازه با استفاده از پارامترهای معلوم نمونه بدست می‌آوریم و نشان می‌دهیم که با احتمال زیاد پارامتر مجهول جامعه در این بازه قرار دارد.

در برآورد نقطه‌ای از دو روش زیر استفاده می‌کنیم که در ادامه با آنها آشنا می‌شوید:

۱- برآورد به روش گشتاورها

۲- برآورد به روش حداکثر احتمال (در ...... ماکزیمم).

۳ ۱۰ . ۱ . ۱ برآورد به روش گشتاورها

قبلاً kامین گشتاور متغیر تصادفی X را بصورت زیر معرفی نمودیم:

$$m_k = E\left[X^K\right]$$

به همین ترتیب می‌توانیم kامین گشتاور نمونه‌ها را بصورت زیر معرفی کنیم:

$$m'_k = \frac{۱}{n}\sum_{i=۱}^{n} X_i{}^k$$

n: تعداد نمونه‌ها

با توجه به اینکه kامین گشتاور متغیر تصادفی X به تمام مقادیر احتمال متغیر X وابسته است و همچنین kامین گشتاور نمونه‌ها نیز به هر یک از مقادیر نمونه‌ها وابسته است بنابراین این طور بنظر می‌رسد که با استفاده از گشتاورها بتوان بنحوی پارامترهای مجهول جامعه را تقریب زد.

در این روش به این صورت عمل می‌کنیم که اگر p عدد پارامتر مجهول داشته باشیم به ترتیب گشتاور اول نمونه را با گشتاور اول متغیر تصادفی x، گشتاور دوم نمونه را با گشتاور دوم متغیر تصادفی و گشتاور pام نمونه را با گشتاور pام متغیر تصافی X برابر قرار می‌دهیم به این ترتیب دستگاه معادلات زیر بدست می‌آید:

$$\begin{cases} m_۱ = m'_۱ \\ m_۲ = m'_۲ \\ \quad\vdots \\ m_k = m'_k \end{cases} \qquad k = ۱\,,\,۲\,,\,\dots\,,\,p$$

اگر $\theta_۱, \theta_۲, \dots, \theta_p$ مقادیر مجهول پارامترها باشد با حل دستگاه فوق هر یک از مقادیر θ_i بدست می‌آید.

حال به مثال زیر توجه کنید:

S-۲

۴ مثال ۱: می‌دانیم تعداد مراجعه کنندگان به یک پمپ بنزین در طول روز یک متغیر تصادفی پواسون با پارامتر مجهول λ می‌باشد. اگر در طول ۱۰ روز مشاهدات زیر را برای تعداد مراجعات به پمپ بنزین بدست آورده باشیم مطاوبست تعیین پارامتر مجهول λ؟

۱۲۰ , ۹۰ , ۸۵ , ۱۱۰ , ۱۳۵ , ۹۵ , ۱۰۰ , ۱۱۱ , ۱۱۵ , ۹۸

حل: برای متغیر تصادفی پواسون داریم:

$$m_۱ = E\left[X'\right] = E\left[X\right] = \lambda$$

از آنجا که تنها یک پارامتر مجهول λ داریم استفاده از اولین گشتاور کفایت می‌کند. حال اولین گشتاور نمونه را بدست می‌آوریم:

$$m'_۱ = \frac{۱}{n}\sum_{i=۱}^{n} X_i = \overline{X} = \frac{۱}{۱۰}(۱۲۰+۹۰+۸۵+۱۱۰+۱۳۵+۹۵+۱۰۰+۱۱۱+۱۱۵+۹۸) = ۱۰۵/۹$$

حال با برابر قرار دادن اولین گشتاور نمونه با اولین گشتاور متغیر تصادفی X مقدار λ بدست می‌آید:

$$m_۱ = m'_۱ = ۱۰۵/۹ \quad \Rightarrow \quad \lambda = ۱۰۵/۹$$

۵ مثال ۲: در یک سری مسابقات تیراندازی n مرتبه به هدف شلیک می‌ند که احتمال اصابت گلوله به هدف p می‌باشد. نتایج حاصل از ۱۰ مرتبه شرکت تیرانداز در مسابقات بصورت زیر می‌باشد: (نتایج بر حسب تعداد دفعات اصابت گلوله به هدف می‌باشند)

۷ , ۵ , ۶ , ۸ , ۸ , ۶ , ۷ , ۴ , ۹ , ۸

مطلوبست برآورد پارامترهای مجهول n و p؟

ابتدا توجه می‌نیم که در این مثال تعداد پیروزی‌ها در n مرتبه انجام آزمایشهای مستقل برنولی با احتمال پیروزی p مورد نظر است. بنابراین متغیر تصادفی X دو جمله‌ای با پارامترهای n و p می‌باشد.

در این مثال دو پارامتر مجهول داریم بنابراین از گشتاور مرتبه اول و دوم استفاده می‌کنیم:

$$m_۱ = E[X] = np$$

$$m_۲ = E[X^۲] = \sigma^۲ + \mu^۲ = npq + n^۲p^۲ = n(n-۱)p + np = np\,(np + ۱ - p)$$

توجه کنید که مقادیر اول و دوم گشتاور متغیر تصادفی X همواره با استفاده از میانگین و واریانس متغیر تصادفی X بدست می‌آیند. به عبارتی در برآورد به روش گشتاورها مستقل از اینکه توزیع متغیر تصادفی X چه باشد همواره داریم:

$$\begin{cases} m_۱ = \mu \\ m'_۱ = \overline{X} \end{cases} \quad \Rightarrow \quad \mu = \overline{X}$$

$$\begin{cases} m_۲ = \sigma^۲ + \mu^۲ \\ m'_۲ = \dfrac{۱}{n}\sum\limits_{i=۱}^{n} \overline{X}_i^۲ \end{cases} \quad \Rightarrow \quad \sigma^۲ = \frac{۱}{n}\sum_{i=۱}^{n}(\overline{X}_i - \overline{X})^۲ = \frac{۱}{n}\sum_{i=۱}^{n}\overline{X}_i^{۲} - \overline{X}^۲$$

یعنی برآورد گشتاوری میانگین و واریانس هر توزیعی برابر است با میانگین و واریانس نمونه‌ها. به همین دلیل $\overline{X}$, $\frac{۱}{n}\sum(X_i - \overline{X})^۲$ برآوردهای توزیع – آزاد نامیده می‌شوند، یعنی برآوردهایی که مستقل از توزیع تصادفی X می‌باشند.

۶ حال مقادیر گشتاور اول و دوم نمونه را بدست می‌آوریم:

$$m'_۱ = \frac{۱}{n}\sum_{i=۱}^{n} X_i = \frac{۱}{۱۰}(۷+۵+۶+۸+۸+۶+۷+۴+۹+۸) = ۶/۸$$

$$m'_۲ = \frac{۱}{n}\sum_{i=۱}^{n} X_i^۲ = \frac{۱}{۱۰}(۴۹+۲۵+۳۶+۶۴+۶۴+۳۶+۴۹+۱۶+۸۱+۶۴) = ۴۸/۴$$

به این ترتیب دستگاه معادلات زیر بدست می‌آید:

$$\begin{cases} m_1 = m'_1 & \Rightarrow \quad np = ۶/۸ \\ m_2 = m'_2 & \Rightarrow \quad np = (np + ۱ - p) = ۴۸/۴ \end{cases}$$

با حل دستگاه بدست می‌آوریم:

$$۶/۸\,(۶/۸ + ۱ + p) = ۴۸/۴ \quad \Rightarrow \quad ۵۳/۰۴ - ۶/۸\,p = ۴۸/۴ \quad \Rightarrow \quad p = ۰/۶۸$$

$$np = ۶/۸ \;\rightarrow\; n\,(۰/۶۸) = ۶/۸ \;\Rightarrow\; n = ۱۰$$

بنابراین تیرانداز فوق در هر مسابقه می‌بایستی حدوداً ۱۰ مرتبه به هدف شلیک کرده باشد که احتمال موفقیت وی در هر شلیک ۰/۶۸ می‌باشد.

برآورد به روش گشتاورها همواره نتایج دقیق و رضایت بخشی نمی‌دهد به مثال بعد در این زمینه توجه کنید:

۷ مثال ۳: فرض کنید متغیر تصادفی X در بازه $[\circ, a]$ بصورت یکنواخت توزیع شده باشد در این صورت مقدار a را به ازای نمونه‌های بدست آمده زیر برآورد کنید:

الف) ۵ و ۴ و ۳ و ۱ و ۸ و ۵ و ۴ و ۲ و ۰ و ۱

ب) ۵ و ۱۰ و ۴۰

حل: برای متغیر تصادفی پیوسته یکنواخت داریم:

$$m_1 = E[X] = \int_\circ^a x \frac{۱}{a} dx = \frac{۱}{a} \left(\frac{x^۲}{۲}\right)\Big|_\circ^a = \frac{a}{۲}$$

همینطور برای نمونه‌ها داریم:

$$m'_1 = \frac{۱}{n} \sum_{i=۱}^{n} X_i = \frac{۱}{۱۰} (۱+\circ+۲+۴+۵+۸+۱+۳+۴+۵) = ۳/۳$$

بنابراین در حالت کلی:

$$a = ۲\overline{X} \quad \Leftarrow \quad m'_1 = \overline{X} = \frac{a}{۲}$$

یعنی برآورد a عبارتست از دو برابر میانگین نمونه‌های بدست آمده که در اینجا با توجه به نمونه‌های الف a برابر می‌شود با: $a = ۲\overline{X} = ۲\times۳/۳ = ۶/۶$.

حال توجه کنید که $a = ۶/۶$ یعنی نمونه‌ها در اصل از بازه $[\circ\,,\,۶/۶]$ انتخاب شده‌اند در حالی که در بین نمونه‌ها عدد ۸ موجود می‌باشد که داخل بازه فوق نمی‌باشد و این یعنی برآورد به روش گشتاورها دارای کمبودهایی می‌باشد. همین حالت برای نمونه‌های (ب) نیز صادق است:

ب)

$$\overline{X} = \frac{۱}{۳} (۴۰+۱۰+۵) = ۱۸/۳ \quad \Rightarrow \quad a = ۲\overline{x} = ۳۶/۶$$

بنابراین برآورد به روش گشتاورها در بعضی موارد نتایج دلخواه را بدست نمی‌دهد در نتیجه می‌بایستی از معیاری استفاده کنیم که میزان کارایی یک روش برآورد را نشان دهد تا بتوان میان روشهای مختلف برآورد، بهترین روش را برای مسایل مختلف انتخاب نمود.

S-۳

۲ .۱ .۱ . برآورد به روش حداکثر احتمال

متغیر تصادفی X با تابع چگالی احتمال $f_X(x)$ و پارامترهای مجهول $\theta_1, \theta_2, \ldots, \theta_p$ را در نظر بگیرید. از این متغیر تعداد n نمونه استخراج می‌کنیم که عبارتند از $x_1, x_2, \ldots, x_n$ حال $X_1, X_2, \ldots, X_n$ را متغیرهای تصادفی متناظر با نمونه‌ها در نظر می‌گیریم.

می‌دانیم تابع چگالی احتمال توام متغیرهای تصادفی $X_1, \ldots, X_n$ مقادیر احتمال وقوع نمونه‌ها را بدست می‌دهد. مقادیری از θ_iها که تابع چگالی احتمال توام X_iها را ماکزیمم می‌کند برآورد پارامترهای مجهول θ_i می‌باشد. زیرا به ازای ماکزیمم شدن تابع چگالی احتمال توام X_iها در واقع احتمال وقوع نمونه‌ها به بیشترین مقدار خود می‌رسد. برای بدست آوردن برآورد θ_iها فرض می‌کنیم $L(\theta_1, \theta_2, \ldots, \theta_p)$ تابع چگالی احتمال توام متغیرهای تصادفی X_i باشد در این صورت با توجه به اینکه X_iها از یکدیگر مستقل باشند تابع چگالی احتمال توام آنها عبارتست از:

$$L(\theta_1, \theta_2, \ldots, \theta_p) = f_{X_1}(x_1) \,.\, f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

حال با مشتق‌گیری از $L(\theta_1, \ldots, \theta_p)$ نسبت به θ_iها و برابر صفر قرار دادن این مشتق‌ها دستگاهی از معادلات بدست می‌آید که با حل این دستگاه مقادیر هر یک از θ_iها بدست می‌آید. معمولاً در محاسبات برای سادگی از $Ln(L)$ مشتق می‌گیریم. به مثال زیر توجه کنید:

۹ مثال ۴: نمونه‌های $X_1, X_2, \ldots, X_n$ را از متغیر تصادفی برنولی X داریم مطلوبست: برآورد پارامتر p به روش حداکثر احتمال (MLE)؟

حل: تابع چگالی متغیر تصادفی برنولی X به صورت زیر می‌باشد:

$$f_X(x) = p^x (1-p)^{1-x} \qquad x = 0, 1$$
$$= 0 \qquad \text{سایر مقادیر}$$

به این ترتیب تابع چگالی احتمال توام نمونه‌ها عبارتست از:

$$L(p) = f_{X_1}(x_1) \,.\, f_{X_2}(x_2) \cdots f_{X_n}(x_n) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$

$$p^{\sum_{i=1}^{n} x_i} (1-p)^{\sum_{i=1}^{n}(1-x_i)} \Rightarrow L(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}$$

حال برای اینکه محاسبات ساده‌تر شوند از طرفین رابطه بالا Ln می‌گیریم:

$$Ln(L(p)) = (n(p^{\sum x_i} (1-p)^{n-\sum x_i})$$
$$= \sum x_i \, Ln(p) + (n - \sum x_i)(n(1-p)$$

فرض می‌کنیم $H(p) = Ln(L(p))$ در این صورت می‌بایستی معادله $\frac{dH}{dP} = 0$ را حل کنیم تا مقدار مجهول p بدست آید. بنابراین:

$$\frac{dH}{dP} = 0 \quad \Rightarrow \quad \sum x_i \left(\frac{1}{p}\right) - (n - \sum x_i)\left(\frac{1}{1-p}\right) = 0$$

همانطور که ملاحظه می‌کنید با بکار بردن Ln در محاسبات، مشتق‌گیری بسیار ساده‌تر می‌شود. حال داریم:

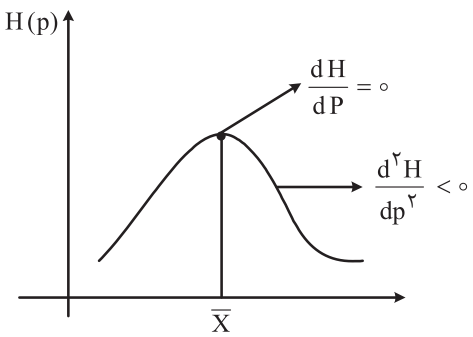
$$\sum x_i \left(\frac{1}{p}\right) = (n - \sum x_i)\left(\frac{1}{1-p}\right)$$
$$\Rightarrow \frac{1-p}{p} = \frac{n - \sum x_i}{\sum x_i} \quad \Rightarrow \quad \frac{1}{p} - 1 = \frac{n}{\sum x_i} - 1$$
$$\Rightarrow \frac{1}{p} = \frac{n}{\sum x_i} \quad \Rightarrow \quad \hat{p} = \frac{1}{n} \sum x_i$$

($\hat{p}$ برآورد پارامتر مجهول p می‌باشد).

بنابراین برآورد پارامتر مجهول p برابر است با $\overline{X} = \frac{1}{n} \sum x_i$ توجه کنید مقدار $\hat{p} = \overline{X}$ بدست آمده تنها در صورتی جواب درستی می‌باشد که نقطه $\overline{X}$ واقعاً مقدار ماکزیمم تابع $H(P)$ باشد. مثلاً ممکن است این نقطه، نقطه مینیمم تابع $H(P)$ باشد. برای اطمینان از اینکه $\hat{p} = \overline{X}$ مقدار ماکزیمم $H(P)$ می‌باشد باید شرط زیر برقرار باشد:

$$\frac{d^2H}{dp^2} < 0$$

یعنی تقعر تابع $H(P)$ حول $\hat{p} = \overline{X}$ روبه پایین باشد به این ترتیب $\hat{p} = \overline{X}$ نقطه ماکزیمم تابع می‌باشد. در این رابطه به شکل زیر توجه کنید:



حال داریم:

$$\frac{d^2H}{dp^2} = \sum x_i \left(\frac{-1}{p^2}\right) - \left(n - \sum x_i\right)\left(\frac{1}{(1-p)^2}\right)$$

که به ازای هر مقدار p منفی است. بنابراین $\overline{X}$ به عنوان برآورد p در متغیر تصادفی برنولی می‌باشد.

۱۱ حال برآورد p را با استفاده از روش گشتاورها بدست می‌آوریم و نتیجه را با روش حداکثر احتمال مقایسه می‌کنیم:

مطابق روش گشتاورها داریم:

$$m_1 = p$$

$$\Rightarrow \quad m_1 = m'_1 \;\Rightarrow\; p = \overline{X}$$

$$m'_1 = \frac{1}{n}\sum x_i$$

همانطور که ملاحظه می‌کنید نتیجه برآورد از هر دو روش یک نتیجه را بدست می‌دهد. بطور کلی از آنجا که روش حداکثر معمولاً برآوردهای بهتری می‌دهد، هرگاه نتیجه دو روش برآورد با یکدیگر برابر نبود، نتیجه روش برآورد حداکثر احتمال را ملاک قرار می‌دهیم.در برآورد به روش گشتاورها نشان دادیم که این روش برای مواردی مثل مثال ۳ برآورد مطلوبی ارایه نمی‌کند بنابراین در مثال بعدی مثال ۳ را با روش برآورد حداکثر احتمال حل می‌کنیم:

S ۴

۱۲ مثال ۵: متغیر X یک متغیر تصادفی یکنواخت در بازه $(0, a)$ می‌باشد. مطلوبست برآورد پارامتر مجهول a به روش حداکثر احتمال؟

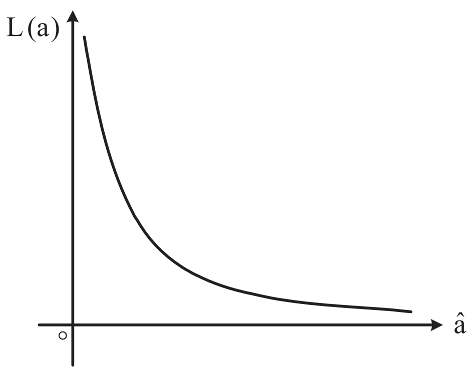
حل: تابع چگالی احتمال X عبارتست از:

$$f_X(x) = \frac{1}{a} \qquad \text{برای} \quad 0 < x < a$$

با در نظر گرفتن یک نمونه تصادفی به اندازه n تابع چگالی احتمال توام بصورت زیر بدست می‌آید:

$$L(a) = \frac{1}{a}\cdot\frac{1}{a}\cdots\frac{1}{a} = \frac{1}{a^n}$$

منحنی نمایش $L(a)$ را در شکل زیر مشاهده می‌کنید.



از روی منحنی پیداست که شیب منحنی در هیچ نقطه‌ای صفر نمی‌شود بنابراین مقدار مشتق $L(a)$ در هیچ نقطه‌ای صفر نمی‌شود. اما با توجه به نمودار می‌بینیم که هر چقدر a به صفر نزدیک می‌شود مقدار $L(a)$ بیشتر می‌شود، و به عبارتی $L(a)$ به ازای کوچکترین مقدار a ماکزیمم می‌شود، از طرفی کوچکترین مقدار a نباید از بزرگترین نمونه بدست آمده کمتر باشد بنابراین برآورد a به روش حداکثر احتمال در این مثال برابر است با بزرگترین نمونه بدست آمده یعنی nامین آماره مرتب $\hat{a} = X_{(n)}$.

بوضوح این برآورد کاراتر از برآورد به روش گشتاورها می‌باشد.

۱۳ مثال ۶: اگر متغیر تصادفی X یک متغیر تصادفی پواسون با پارامتر λ باشد مطلوبست برآورد λ به روش حداکثر احتمال و به روش گشتاورها؟

حل: تابع چگالی متغیر تصادفی پواسون عبارتست از:

$$f_X(x) = \frac{e^{-\lambda}\lambda^{x}}{x_i} \qquad x = 0,1,2,\ldots$$

$$= 0 \qquad \text{سایر مقادیر}$$

$$L(\lambda) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!}.\frac{e^{-\lambda}\lambda^{x_2}}{x_2!}\cdots\frac{e^{-\lambda}\lambda^{x_n}}{x_n!} = \frac{e^{-n\lambda}\ \lambda^{\sum x_i}}{x_1!\,x_2!\cdots x_n!}$$

$$H(\lambda) = \mathrm{Ln}(L(\lambda)) = -n\lambda + \sum x_i\ \mathrm{Ln}(\lambda) - \mathrm{Ln}(x_1!\,x_2!\cdots x_n!)$$

$$\frac{dH}{d\lambda} = -n + \sum x_i\left(\frac{1}{\lambda}\right) = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{n}\sum x_i = \overline{X}$$

حال این مساله را به روش گشتاورها حل می‌کنیم:

$$m_1 = \mu = \lambda$$

$$\Rightarrow \quad m_1 = m_1' \ \Rightarrow \ \hat{\lambda} = \overline{X}$$

$$m_1' = \frac{1}{n}\sum x_i = \overline{X}$$

ملاحظه می‌کنید که نتایج هر دو روش در این حالت نیز یکسان می‌باشند.

۱۴ مثال ۷: اگر متغیر تصادفی X یک متغیر تصادفی نرمال با میانگین μ و واریانس σ^2 باشد مطلوبست برآورد پارامترهای μ و σ^2 به روشهای حداکثر احتمال و گشتاورها؟

حل: تابع چگالی احتمال متغیر تصادفی X عبارتست از:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$L(\mu,\sigma^2) = \prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}\ e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}$$

$$H(\mu,\sigma^2) = \mathrm{Ln}\,(L((\mu,\sigma^2)) = \mathrm{Ln}\left(\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}\right) + \mathrm{Ln}\left(e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}\right)$$

$$= -\frac{n}{2}\,(n(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$$

$$\Rightarrow H(\mu,\sigma^2) = -\frac{n}{2}(n(2\pi) - \frac{n}{2}(n\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$$

$$\frac{\partial H}{\partial \mu} = -\frac{1}{2}(-2) = \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial H}{\partial \sigma^2} = -\frac{n}{2}(\frac{1}{\sigma^2}) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2 (\frac{-1}{(\sigma^2)^2})$$

$$= \frac{-n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

حال می‌بایستی معادلات زیر را حل کنیم:

$$\begin{cases} \frac{1}{\hat{\sigma}^2}\sum(x_i-\hat{\mu}) = 0 \\ -\frac{n}{2\hat{\sigma}^2}+\frac{1}{2(\hat{\sigma}^2)^2}\sum(x_i-\mu)^2 = 0 \end{cases}$$

۱۵ جواب معادله اول عبارتست از:

$$\sum(x_i-\hat{\mu}) = 0 \quad \Rightarrow \quad \sum x_i - \sum\hat{\mu} = 0 \quad \to \quad \sum x_i - n\hat{\mu} = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum x_i = \overline{X}$$

با قرار دادن $\hat{\mu} = \overline{X}$ در معادله دوم داریم:

$$-\frac{n}{2\hat{\sigma}^2} + \frac{1}{2(\hat{\sigma}^2)^2}\sum(x_i-\overline{X})^2 = 0$$

$$n\,2\hat{\sigma}^2 + \sum(x_i-\overline{X})^2 = 0 \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{1}{n}\sum(x_i-\overline{X})^2$$

بنابراین برآورد پارامترهای مجهول μ و σ^2 به روش حداکثر احتمال عبارتند از:

$$\hat{\mu} = \overline{X} \quad , \quad \hat{\sigma}^2 = \frac{n-1}{n}S^2$$

۱۶ حال برآورد را به روش گشتاورها بدست می‌آوریم:

$$m_1 = \mu$$

$$m_2 = \sigma^2 + \mu^2$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n}x_i = \overline{X}$$

$$m'_2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2$$

$$\begin{cases} m_1 = m'_1 \\ m_2 = m'_2 \end{cases} \quad \Rightarrow \quad \begin{cases} \mu = \overline{X} \\ \frac{1}{n}\sum_{i=1}^{n}x_i^2 = \sigma^2 + \mu^2 \end{cases}$$

از معادله اول داریم $\hat{\mu} = \overline{X}$ با جاگذاری در معادله دوم بدست می‌آید:

$$\frac{1}{n}\sum_{i=1}^{n} x_i^{2} = \hat{\sigma}^{2} + \overline{X}^{2} \qquad \Rightarrow \qquad \hat{\sigma}^{2} = \frac{1}{n}\sum_{i=1}^{n} x_i^{2} - \overline{X}^{2}$$

$$\hat{\sigma}^{2} = \frac{1}{n}\,\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^{2} = \frac{n-1}{n} S^{2}$$

در نتیجه برآورد پارامترهای μ و σ^{2} به روش گشتاورها نیز همان نتایج برآورد به روش حداکثر را بدست می‌دهد.

$$S^{2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^{2}$$

S-۵

۱۶ مثال ۸: اگر متغیر تصادفی X بصورت یکنواخت در بازه $[-b\,,b]$ توزیع شده باشد مطلوبست برآورد b به روش گشتاورها و حداکثر احتمال؟

حل: میانگین متغیر تصادفی X عبارتست از:

$$E[X] = \int_{-b}^{b} dx = \frac{1}{2b}\left(\frac{x^{2}}{2}\right)\Big|_{-b}^{b} = 0$$

حال طبق روش گشتاورها می‌بایستی از حل معادله $m_1 = m'_1$ پارامتر b بدست آید اما می‌بینیم که:

$$m_1 = E[X] = 0$$

$$m'_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$$

$$m_1 = m'_1 \;\Rightarrow\; \overline{X} = 0$$

همانطور که ملاحظه می‌کنید $\overline{X} = 0$ کمکی در یافتن پارامتر مجهول به ما نمی‌کند. در واقع این مثال حالت خاصی است که در آن جهت تعین پارامتر مجهول می‌بایستی از گشتاور دوم نمونه و متغیر تصادفی X استفاده کنیم:

$$m_2 = E\left[X^{2}\right] = \int_{-b}^{b} x^{2}\frac{1}{2b}\,dx = \frac{1}{2b}\left(\frac{x^{3}}{3}\right)\Big|_{-b}^{b}$$

$$= \frac{1}{2b}\left(\frac{b^{3}}{3} + \frac{b^{3}}{3}\right) = \frac{b^{2}}{3}$$

به همین ترتیب داریم:

$$m'_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^{2} \quad \Rightarrow m_2 = m'_2 \quad \Rightarrow \quad \frac{b^{2}}{3} = \frac{1}{n}\sum_{i=1}^{n} x_i^{2}$$

$$\Rightarrow b^{2} = 3\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{2}\right) \;\rightarrow\; \hat{b} = \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{2}\right)} = \sqrt{3m'_2}$$

بنابراین برآورد پارامتر مجهول b در این مثال برابر $\sqrt{3m'_2}$ می‌باشد.

۱۷ روش برآورد پارامتر b به روش حداکثر احتمال کاملاً مشابه مثال ۵ می‌باشد به این صورت که:

$$f_X(x) = \frac{1}{2b} \qquad\qquad -b < x < b$$

$$L(b) = \frac{1}{2b}.\frac{1}{2b}\cdots\frac{1}{2b} = \frac{1}{(2b)^{n}}$$

مجدداً از آنجا که $L(b)$ با نزدیک شدن به صفر ماکزیمم می‌شود بنابراین کوچکترین مقداری که b می‌تواند بخود بگیرد برآورد آن می‌باشد و از آنجا که b نمی‌تواند از بزرگترین مقدار نمونه کمتر باشد بنابراین در این مثال نیز برآورد b عبارتست از $\max\left|X_i\right|$ یا همان nامین آماره مرتب $X_{(n)}$.

۲.۱.۱۸ خواص برآورد کننده‌ها

در بخش روشهای گشتاورها و حداکثر احتمال را برای برآورد پارامترهای مجهول جامعه معرفی نمودیم و نشان دادیم که این دو روش همواره برآورد کننده‌های مشابه بدست نمی‌دهند، بنابراین نیازمند یک معیاری برای سنجش میزان کارایی یک برآوردگر نسبت به برآوردگر دیگری می‌باشیم، در این بخش به معرفی خواص برآورد کننده‌ها و روشهایی برای سنجش میزان کارایی آنها می‌پردازیم.

۱.۲.۱.۱۹ متوسط مربع خطا (MSE)

فرض کنید $\hat{\theta}$ برآورد پارامتر θ باشد در این صورت می‌دانیم که $\left|\hat{\theta}-\theta\right|$ مقدار خطای برآورد می‌باشد. معمولاً برای تخمین میزان خطا از مربع $(\hat{\theta}-\theta)$ استفاده می‌شود و از آنجا که در اینجا n نمونه استخراج می‌شود برای هر بار استخراج این n نمونه مقدار $(\hat{\theta}-\theta)^۲$ محاسبه می‌شود و میانگین آن ملاک خواهد بود. به این ترتیب معیار متوسط مربع خطا بصورت زیر بدست می‌آید:

$$MSE\ (\hat{\theta}) = E\left[(\hat{\theta}-\theta)^۲\right]$$

توجه کنید که همواره در عمل با توجه به استخراج n نمونه و برآورد پارامتر θ مقداری خطا در برآورد پارامتر θ وجود دارد، بنابراین برای اینکه بهترین برآورد را داشته باشیم می‌بایستی خطای برآورد را حداقل کنیم. بوضوح اگر $\hat{\theta}=\theta$ باشد خطای برآورد صفر می‌باشد و بهترین حالت بدست آمده است.

برای اینکه بتوان خطای برآورد را حداقل نمود ابتدا مقدار متوسط مربع خطا را ساده می‌کنیم:

$$MSE = E\left[(\hat{\theta}-\theta)^۲\right] = E\left[(\hat{\theta}^۲-۲\theta\hat{\theta}-\theta)^۲\right]$$
$$= E\left[\hat{\theta}^۲\right] - ۲E\left[\theta\hat{\theta}\right] + E\left[\theta^۲\right]$$

حال یک $E^۲\left[\hat{\theta}\right]$ به سمت راست عبارت بالا اضافه و کم می‌کنیم:

$$\Rightarrow MSE = E\left[\hat{\theta}^۲\right] - E^۲\left[\hat{\theta}\right] + E^۲\left[\hat{\theta}\right] - ۲E\left[\theta\hat{\theta}\right] + E\left[\theta^۲\right]$$

توجه کنید که عبارت A همان واریانس $\hat{\theta}$ می‌باشد. می‌دانیم $\hat{\theta}$ خود یک آماره می‌باشد و از آنجا که آماره‌ها خود یک متغیر تصادفی می‌باشند، بنابراین واریانس $\hat{\theta}$ معنی‌دار می‌باشد.

$$\Rightarrow MSE = var(\hat{\theta}) + (E\left[\hat{\theta}\right]-\theta)^۲$$

بنابراین متوسط مربع خطا بصورت فوق خلاصه می‌شود که در آن عبارت $E\left[\hat{\theta}\right]-\theta$ در واقع اندازه اختلاف مرکز توزیع $\hat{\theta}$ از θ را نشان می‌دهد.

۲۰ برای اینکه MSE حداقل شود می‌بایستی مقادیر $(E\left[\hat{\theta}\right]-\theta)^۲,\ var(\hat{\theta})$ حداقل شوند. از طرفی می‌دانیم $(E\left[\hat{\theta}\right]-\theta)^۲,\ var(\hat{\theta})$ هر دو عبارتی همواره مثبت می‌باشند بنابراین MSE در صورتی حداقل می‌شود که این دو عبارت حداقل شوند. حال کمترین مقداری را که این دو عبارت بخود می‌پذیرد را محاسبه می‌کنیم:

$$var(\hat{\theta}) = E\left[\hat{\theta}^۲\right] - E^۲\left[\hat{\theta}\right]$$

اگر $E^۲\left[\hat{\theta}\right]=E\left[\hat{\theta}^۲\right]$ واریانس $\hat{\theta}$ برابر صفر می‌شود که کمترین مقداری است که یک عبارت مثبت بخود می‌پذیرد. بنابراین $\min(var(\hat{\theta})) = ۰$

همینطور حداقل $(E\left[\hat{\theta}\right]-\theta)^۲$ با توجه به مثبت بودن عبارت زمانی رخ می‌دهد که کل عبارت صفر شود یعنی:

$$\min\left[(E\left[\hat{\theta}\right]-\theta)^۲\right] = ۰ \quad \Rightarrow \quad (E\left[\hat{\theta}\right]-\theta)^۲ = ۰ \quad \Rightarrow \quad E\left[\hat{\theta}\right] = \theta$$

می‌دانیم در عمل $var(\hat{\theta})$ هرگز صفر نمی‌شود (مگر اینکه تمام مقادیر نمونه‌ها برابر باشند که آن هم بی‌معنی است). اما رخ دادن تساوی $E\left[\hat{\theta}\right] = \theta$ امکان‌پذیر است بنابراین از میان برآوردگرهای مختلف برای θ برآوردگری مناسب‌تر است که شرایط زیر را داشته باشد.

۱- واریانس برآوردگر $(\hat{\theta})$ کمترین مقدار ممکن را داشته باشد.

۲- $E\left[\hat{\theta}\right] = \theta$ باشد.

شرط دوم به برآورد کننده نااریب معروف است که به صورت زیر تعریف می‌شود:

برآورد کننده نااریب: برآورد کننده $\hat{\theta}$ برای θ نااریب نامیده می‌شود اگر $E\left[\hat{\theta}\right]=\theta$.

برابر تساوی $E\left[\hat{\theta}\right]=\theta$ به این معنی است که اگر مثلاً kبار از جامعه تعداد n نمونه استخراج کنیم و به ازای هر بار استخراج n نمونه مقدار $\hat{\theta}$ را محاسبه کنیم و در نهایت میانگین $\hat{\theta}_i$ها ($E\left[\hat{\theta}\right]=\frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i$) می‌بایستی برابر با پارامتر مجهول θ شود. برای روشن شدن مطلب به مثال زیر توجه کنید.

۲۱ مثال ۹: آیا برآوردگر بدست آمده برای توزیع نرمال، برنولی، پواسون نااریب می‌باشد؟

حل: قبلاً نشان دادیم که برآوردگر μ و σ^2 در توزیع نرمال با استفاده از روشهای گشتاورها و حداکثر احتمال عبارتست از:

$$\hat{\mu}=\overline{X}$$

$$\hat{\sigma}^2=\frac{n-1}{n}S^2$$

برای $\hat{\mu}$ داریم:

$$E\left[\hat{\mu}\right]=E\left[\overline{X}\right]=\mu$$

توجه کنید که قبلاً نشان دادیم برای متغیر تصادفی X با هر توزیعی $E\left[\overline{X}\right]$ برابر μ می‌باشد.

بنابراین $\hat{\mu}=\overline{X}$ در توزیع نرمال یک برآورد نااریب برای μ می‌باشد.

حال برای $\hat{\sigma}^2$ داریم:

$$E\left[\hat{\sigma}^2\right]=E\left[\frac{n-1}{n}S^2\right]=\frac{n-1}{n}\;E\left[S^2\right]$$

۲۲ در فصل قبل نشان دادیم که اگر $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})$ آنگاه داریم:

$$E\left[S^2\right]=\sigma^2$$

بنابراین:

$$E\left[\hat{\sigma}^2\right]=\frac{n-1}{n}\;\sigma^2$$

ملاحظه می‌کنید که برای $\hat{\sigma}^2$ مقدار $E\left[\hat{\sigma}^2\right]$ برابر با σ^2 نمی‌باشد بنابراین برآوردگر $\hat{\sigma}^2-\frac{n-1}{n}\;S^2$ برآوردگر نااریبی برای σ^2 نمی‌باشد. در این حالت می‌گوییم برآوردگر $\hat{\sigma}^2$ یک برآوردگر اریب برای σ^2 می‌باشد.

توجه کنید که $\hat{\sigma}^2$ تنها به دلیل وجود ضریب $\left(\frac{n-1}{n}\right)$ اریب می‌باشد، به عبارتی تعداد نمونه‌ها عامل اصلی اریب بودن این برآوردگر می‌باشد بنابراین اگر تعداد نمونه‌ها زیاد باشد $(n\to\infty)$ داریم:

$$\lim_{n\to\infty}\left(\frac{n-1}{n}\right)\sigma^2=\sigma^2$$

یعنی با شرط $n\to\infty$ برآوردگر $\hat{\sigma}^2$ یک برآوردگر نااریب برای σ^2 می‌باشد. به این نوع برآوردگرها، برآوردگر نااریب مجانبی می‌گویند که بصورت زیر تعریف می‌شود:

برآوردگر نااریب مجانبی: اگر $\hat{\theta}$ یک برآوردگر اریب برای θ بر اساس نمونه تصادفی nتایی باشد، می‌گوییم $\hat{\theta}$ یک برآوردگر نااریب مجانبی برای θ است هرگاه داشته باشیم:

$$\lim_{n\to\infty}E\left[\hat{\theta}\right]=\theta$$

۲۳ برای توزیع برنولی قبلاً بدست آوردیم که:

$$\hat{p}=\overline{X}\quad\Rightarrow\quad E\left[\hat{p}\right]=E\left[\overline{X}\right]=\mu=p$$

همینطور برای توزیع پواسون داریم:

$$\hat{\lambda}=\overline{X}$$

$$E\left[\hat{\lambda}\right]=E\left[\overline{X}\right]=\mu=\lambda$$

یک نتیجه کلی که از برآوردگر نااریب بدست می‌آوریم این است که هرگاه پارامتر مجهول توزیع برابر با میانگین توزیع متغیر تصادفی X باشد

(یعنی $\theta = E[X] = \mu$) همینطور از آنجا که برای هر توزیعی مستقل از نوع توزیع داریم $E[\overline{X}] = \mu$ در این صورت $\overline{X}$ همواره به عنوان یک برآوردگر ناااریب این گونه از توزیعها، منظور می‌شود.

خاصیت ناااریبی اگر چه یک معیار برای محک بهتر بودن یک برآوردگر می‌باشد اما از آنجا که برای یک پارامتر می‌توان چندین برآوردگر ناااریب بدست آورد، بنابراین می‌بایستی معیارهای دیگری وجود داشته باشند تا بتوان با کمک آنها از میان برآوردگرهای ناااریب بهترین را انتخاب نمود. برای این منظور معیار کارایی یا موثرتر بودن را معرفی می‌کنیم:

تعریف: اگر $\hat{\theta}_1, \hat{\theta}_2$ هر دو برآوردگرهایی ناااریب برای پارامتر θ باشند در این صورت می‌گوییم $\hat{\theta}_1$ کاراتر یا موثرتر از $\hat{\theta}_2$ می‌باشد هرگاه:

$$\operatorname{var}(\hat{\theta}_1) < \operatorname{var}(\hat{\theta}_2)$$

علت استفاده از این معیار با توجه به تعریف متوسط مربع خطا (MSE) کاملاً قابل توجیه است.

۲۴ مثال ۱۰: از یک جامعه نرمال با میانگین μ و واریانس σ^2 یک نمونه به حجم n انتخاب شده است. در این صورت کدامیک از برآوردگرهای ناااریب زیر معیار مناسب‌تری برای برآورد پارامتر μ می‌باشد؟

الف) $\hat{\mu}_1 = \overline{X}$

ب) $\hat{\mu}_2 = \frac{1}{3}(X_1 + X_3 + X_5)$

حل: با توجه به اینکه هر دو برآوردگر ناااریب می‌باشند بنابراین واریانس آنها را محاسبه نموده و با یکدیگر مقایسه می‌کنیم:

$$\operatorname{var}(\hat{\mu}_1) = \operatorname{var}(\overline{X}) = \frac{\sigma^2}{n}$$

$$\operatorname{var}(\hat{\mu}_2) = \operatorname{var}\left(\frac{1}{3}(X_1 + X_3 + X_5)\right) = \frac{1}{9}\left(\operatorname{var}(X_1) + \operatorname{var}(X_3) + \operatorname{var}(X_5)\right)$$

$$= \frac{1}{9}(3\sigma^2) = \frac{\sigma^2}{3}$$

با مقایسه دو واریانس بدست آمده داریم:

$$\frac{\sigma^2}{n} < \frac{\sigma^2}{3} \qquad \text{برای } n > 3$$

بنابراین اگر تعداد نمونه‌های استخراج شده از متغیر تصادفی X بیستر از ۳ باشد خواهیم داشت:

$$\operatorname{var}(\hat{\mu}_1) < \operatorname{var}(\hat{\mu}_2)$$

و در نتیجه با توجه به معیار کارایی می‌توان نتیجه گرفت که برآورد کننده $\overline{X}$ بهتر از $\frac{1}{3}(X_1 + X_3 + X_5)$ در این مثال می‌باشد.

تعریف: برای دو برآوردگر $\hat{\theta}_1, \hat{\theta}_2$ نسبت $\frac{\operatorname{var}(\hat{\theta}_1)}{\operatorname{var}(\hat{\theta}_2)}$ را کارایی نسبی برآوردگر $\hat{\theta}_2$ نسبت به برآوردگر $\hat{\theta}_1$ می‌نامند.

بوضوح اگر مقدار کسر از یک بیشتر باشد می‌گوییم برآوردگر $\hat{\theta}_2$ از برآوردگر $\hat{\theta}_1$ کاراتر یا موثرتر است.

S-۷

۲۵ مثال ۱۱: از یک جامعه نمونه‌ای به حجم ۲۰ با میانگین $\overline{X}_1 = 15$ و واریانس $\sigma_1^2 = 4$ استخراج کرده‌ایم. همچنین نمونه دیگری به حجم ۳ با میانگین $\overline{X}_2 = 180$ و واریانس $\sigma_2^2 = 2$ استخراج کرده‌ایم. اگر $\overline{X}$ را به عنوان برآوردگر میانگین جامعه انتخاب کنیم کارایی نسبی $\overline{X}_1$ را محاسبه کنید؟

حل: ابتدا واریانس $\overline{X}_1$ و $\overline{X}_2$ را محاسبه می‌کنیم:

$$n_1 = 20\ ,\quad \overline{X}_1 = 15\quad ,\quad \sigma_1^2 = 4$$
$$n_2 = 30\ ,\quad \overline{X}_2 = 18\quad ,\quad \sigma_1^2 = 2$$

$$\operatorname{var}(\overline{X}_1) = \frac{\sigma_1^2}{n_1} = \frac{4}{20} = \frac{1}{5}$$
$$\operatorname{var}(\overline{X}_2) = \frac{\sigma_2^2}{n_2} = \frac{2}{30} = \frac{1}{15}$$

کارایی نسبی $\overline{X}_1$ نسبت به $\overline{X}_2$ $= \dfrac{\operatorname{var}(\overline{X}_2)}{\operatorname{var}(\overline{X}_1)} = \dfrac{\frac{1}{15}}{\frac{1}{5}} = \dfrac{5}{15} = \dfrac{1}{3}$

بنابراین استفاده از نمونه متوسط استخراج شده دوم برای برآورد میانگین جامعه کارایی بیشتری دارد و موثرتر می‌باشد. همچنین توجه کنید که در محاسبه کارایی از متوسط دو نمونه استخراج شده استفاده‌ای نمی‌شود.

۲۶ تا بحال نشان دادیم که یک برآوردگر به شرطی قابل قبول است که در درجه اول ناریب باشد و همچنین کمترین واریانس را در میان سایر برآوردگرها داشته باشد بنابراین تعریف زیر را داریم:

تعریف: اگر $\hat{\theta}$ یک برآورد کننده ناریب برای θ باشد. در این صورت $\hat{\theta}$ بهترین برآورد کننده ناریب خطی برای θ می‌باشد. هرگاه:

۱- $\hat{\theta}$ تابعی خطی از $X_1, X_2, \ldots, X_n$ باشد.

۲- برای هر θ $\quad E\left[\hat{\theta}\right] = \theta$.

۳- از میان تمام برآورد کننده‌های $\hat{\theta}_i$ که در شرایط ۱ و ۲ صدق می‌کند داشته باشیم:

$$\operatorname{var}(\hat{\theta}) \le \operatorname{var}(\hat{\theta}_i)$$

با توجه به استفاده از روش متوسط مربع خطا (MSE) تعریف بالا و شرایط آن کاملاً موجه می‌باشد. همچنین توجه کنید که اگر $\hat{\theta}$ در شرایط فوق صدق کند در این صورت $\hat{\theta}$ موثرترین برآورد کننده ناریب خطی نیز می‌باشد.

حال به مثال زیر توجه کنید:

۲۷ مثال ۱۲: فرض کنید X یک متغیر تصادفی با میانگین μ و واریانس σ^2 باشد. هرگاه $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از X باشد نشان دهید توزیع متغیر تصادفی X هر چه باشد، برآورد کننده $\overline{X}$ بهترین برآورد کننده ناریب خطی برای μ می‌باشد؟

حل: ابتدا فرض می‌کنیم $\hat{\theta}$ یک برآورد کننده ناریب خطی برای μ باشد در این صورت $\hat{\theta}$ می‌بایستی به فرم زیر باشد:

$$\hat{\theta} = \sum_{i=1}^{n} (a_i\, x_i) + b$$

حال متغیرهای a_i و b را طوری بدست می‌آوریم که $\hat{\theta}$ بهترین برآورد کننده برای μ باشد. برای این منظور مطابق تعریف $\hat{\theta}$ می‌بایستی در شرایط تعریف بهترین برآورد کننده ناریب خطی صدق کند. یعنی: $E\left[\hat{\theta}\right] = \mu$ بنابراین:

$$E\left[\hat{\theta}\right]=E\left[\sum_{i=1}^{n}(a_i\ x_i)+b\right]=E\left[\sum_{i=1}^{n}a_i\ x_i\right]+b$$

$$=E\sum_{i=1}^{n}(a_i\,E\left[X_i\right])+b=\sum_{i=1}^{n}(a_i\ m_1)+b$$

$$=m_1\sum_{i=1}^{n}a_i+b=E\left[X\right]\sum_{i=1}^{n}a_i+b \quad \Rightarrow\ E\left[\hat{\theta}\right]=\mu\sum_{i=1}^{n}a_i+b$$

حال چون بایستی داشته باشیم $E\left[\hat{\theta}\right]=\nu$ بنابراین:

$$\mu\sum_{i=1}^{n}a_i+b=\mu \quad \Rightarrow \quad \begin{cases}\sum\limits_{i=1}^{n}a_i=1\\ b=0\end{cases}$$

۲۸ حال می‌بایستی واریانس برآورد کننده $\hat{\theta}$ را بدست آورده و سپس آنرا حداقل کنیم:

$$\operatorname{var}(\hat{\theta})=\operatorname{var}(\sum_{i=1}^{n}a_i X_i+b)=\operatorname{var}(\sum a_i X_i)$$

$$=\sum\operatorname{var}(a_i X_i)=\sum a_i^2\operatorname{var}(X_i)=\sum a_i^2\,\sigma^2=\sigma^2\sum_{i=1}^{n}a_i^2$$

بنابراین برای حداقل نمودن واریانس $\sum a_i^2$ می‌بایستی حداقل شود به عبارتی باید مقادیر $a_1,a_2,\dots,a_n$ را طوری پیدا نمود که در دو شرط زیر صدق کنند:

۱- $\sum\limits_{i=1}^{n}a_i=1$

۲- $\sum a_i^2$ مینیمم شود.

برای حداقل نمودن $\sum a_i^2$ ابتدا مقدار a_1 را از شرط اول محاسبه نموده و حاصل را در شرط دوم $(\sum a_i^2)$ قرار می‌دهیم. به صورت زیر:

$$\sum_{i=1}^{n}a_i=1 \quad\Rightarrow\quad a_1+\sum_{i=2}^{n}a_i=1 \quad\Rightarrow\quad a_1=1-\sum_{i=2}^{n}a_i$$

فرض می‌کنیم $Q=\sum\limits_{i=1}^{n}a_i^2$ در این صورت باید Q مینیمم شود پس:

$$Q=\sum_{i=1}^{n}a_i^2=a_1^2+\sum_{i=2}^{n}a_i=(1-\sum_{i=2}^{n}a_i)^2+\sum_{i=2}^{n}a_i^2$$

$$\Rightarrow Q=(1-\sum_{i=2}^{n}a_i)^2+\sum_{i=2}^{n}a_i^2$$

حال برای حداقل نمودن Q می‌بایستی مشتقات پاره‌ای Q را نسبت به a_iها محاسبه نموده و برابر صفر قرار دهیم تا مقادیری بشکل $a_1^*,a_2^*,\dots,a_n^*$ که Q را مینیمم می‌کنند بدست آید:

$$\frac{\partial Q}{\partial a_j}=2(-1)(1-\sum_{i=1}^{n}a_i)+2a_j \qquad j=2,3,\cdots,n$$

۲۹ توجه کنید که مقدار a_1^* از روی سایر a_i^*ها $(i=2,3,\cdots,n)$ بشکل زیر بدست می‌آید:

$$a_1^* = 1-\sum_{i=2}^{n} a_i^*$$

رابطه ۱-۱۰

$$\frac{\partial Q}{\partial a_j} = 0 \quad \Rightarrow \quad \begin{cases} a_j^* = (1-\sum\limits_{i=2}^{n} a_i^*) = a_1^* \\ a_1^* = 1-\sum\limits_{i=2}^{n} a_i^* \end{cases}$$

حال با حل دستگاه فوق مقادیر a_j^*ها بدست می‌آید، برای حل ابتدا کل معادلات را جداگانه نوشته و طرفین معادلات را جمع می‌کنیم. به این ترتیب داریم:

$$a_1^* = 1-\sum_{i=2}^{n} a_i^*$$

$$a_2^* = 1-\sum_{i=2}^{n} a_i^*$$

$$a_3^* = 1-\sum_{i=2}^{n} a_i^*$$

$$\vdots$$

$$a_n^* = 1-\sum_{i=2}^{n} a_i^*$$

$$\text{حاصل جمع معادلات} = \sum_{i=1}^{n} a_i^* = n - n\sum_{i=2}^{n} a_i^* = n\,\underbrace{(1-\sum_{i=2}^{n} a_i^*)}_{a_i^*} = n a_1^*$$

بنابراین بدست می‌آید:

$$\begin{cases} \sum\limits_{i=1}^{n} a_i^* = n a_1^* \\ \sum\limits_{i=1}^{n} a_i^* = 1 \end{cases} \quad \Rightarrow \quad n a_1^* = 1 \;\Rightarrow\; a_1^* = \frac{1}{n}$$

و با توجه به رابطه ۱-۱۰ بدست می‌آید: $(j=1,2,\cdots,n) \quad a_j^* = \frac{1}{n}$

بنابراین $\sum\limits_{i=1}^{n} a_i^2$ و با شرط $\sum\limits_{i=1}^{n} a_i = 1$ به ازای $a_j^* = \frac{1}{n}$ حداقل می‌شود. بنابراین برآوردگر $\hat{\theta}$ برابر می‌شود با: $\hat{\theta} = \sum\limits_{i=1}^{n} \frac{1}{n} X_i = \frac{1}{n}\sum\limits_{i=1}^{n} X_i = \overline{X}$

در نتیجه $\overline{X}$ بهترین برآوردگر ناریب خطی برای μ می‌باشد و این مساله مستقل از توزیع متغیر تصادفی X می‌باشد. به همین دلیل این مطلب از اهمیت فراوانی برخوردار است از طرفی می‌توان نشان داد که اگر برآوردگر $\hat{\theta}$ تابعی خطی از X_iها هم نباشد باز هم $\overline{X}$ دارای کوچکترین واریانس می‌باشد. حال می‌توان گفت که $\overline{X}$ برای پارامتر μ در تابع نرمال و برای پارامتر p در توزیع برنولی و برای پارامتر λ در توزیع پواسون بهترین برآوردگر ناریب می‌باشد.

۳۰ . ۱۰ . ۳ برآورد کننده سازگار

یکی دیگر از خواصی که برای برآورد کننده‌ها می‌توان بدست آورد خاصیتی است که به سازگاری معروف است. اگر داشته باشیم:

$$\lim_{n\to\infty} MSE(\hat{\theta}) = \lim E\left[(\hat{\theta}-\theta)^{2}\right] = 0$$

یعنی به ازای تعداد زیادی نمونه از متغیر تصادفی X مقدار متوسط مربع خطا برابر صفر شود، از طرفی طبق نامساوی مارکف (رجوع کنید به ثضیه چبی شف) داریم:

$$p(g(x)\geq k) \leq \frac{E[g(x)]}{k} \Rightarrow 1-p(g(x)\leq k) \leq \frac{E[g(x)]}{k}$$
$$\Rightarrow p(g(x)\geq k) \geq 1-\frac{E[g(x)]}{k}$$

حال اگر $k=\varepsilon^{2}$ $(\varepsilon>0)$, $g(x)=(\hat{\theta}-\theta)^{2}$ باشد در این صورت نامساوی بصورت زیر خواهد بود:

$$p((\hat{\theta}-\theta)^{2}\leq\varepsilon^{2}) \geq 1-\frac{E\left[(\hat{\theta}-\theta)^{2}\right]}{\varepsilon^{2}}$$
$$\Rightarrow p(|\hat{\theta}-\theta|<\varepsilon) \geq 1-\frac{E\left[(\hat{\theta}-\theta)^{2}\right]}{\varepsilon^{2}}$$

حال اگر شرط $\lim_{n\to\infty} MSE = 0$ برقرار باشد می‌بایستی $E\left[(\hat{\theta}-\theta)^{2}\right]$ برابر صفر باشد در این صورت نامساوی بصورت زیر خلاصه می‌شود:

$$p(\left|\hat{\theta}-\theta\right|<\varepsilon) \geq 1-\frac{0}{\varepsilon^{2}} \Rightarrow p(\left|\hat{\theta}-\theta\right|<\varepsilon) \geq 1$$

و از آنجا که احتمال حداکثر برابر ۱ واحد می‌باشد بنابراین:

$$p(\left|\hat{\theta}-\theta\right|<\varepsilon)=1 \qquad \forall\varepsilon>0$$

یا به بیان دیگر:

$$p(\left|\hat{\theta}-\theta\right|\geq\varepsilon)=0 \qquad \forall\varepsilon>0$$

در صورتی که احتمالات فوق برای یک برآورد کننده $\hat{\theta}$ برقرار باشند می‌گوییم برآورد کننده دارای خاصیت سازگاری می‌باشد.
بوضوح خاصیت سازگاری زمانی که تعداد نمونه‌ها به اندازه کافی بزرگ باشد نشان دهنده برتری یک برآورد کننده نسبت به دیگری می‌باشد. اما در حالتی که تعداد نمونه‌ها کم باشد، سازگار بودن یک برآورد کننده معیاری برای بهتر بودن آن نمی‌باشد.
قبل از ارایه یک مثال به قضیه زیر نیز توجه کنید:

۳۱ قضیه: اگر داشته باشیم: $\lim_{n\to\infty} E\left[\hat{\theta}\right]=\theta$ و $\lim_{n\to\infty} var\left[\hat{\theta}\right]=0$ آنگاه $\hat{\theta}$ یک برآورد کننده سازگار برای θ می‌باشد. این قضیه از آنجا نتیجه می‌شود که در تعریف سازگاری داشتیم:

$$\lim_{n\to\infty} MSE=0 \Rightarrow \lim_{n\to\infty} E[(\hat{\theta}-\theta)^{2}]=0$$

$$\Rightarrow \lim_{n\to\infty}(var(\hat{\theta})+(E[\hat{\theta}]-\theta)^{2})=0 \Rightarrow \begin{cases}\lim_{n\to\infty} var(\hat{\theta})=0\\ \lim_{n\to\infty} E[\hat{\theta}]=\theta\end{cases}$$

۳۲ مثال ۱۳: اگر X یک متغیر تصادفی با میانگین μ و واریانس σ^{2} باشد نشان دهید توزیع X هر چه باشد $\overline{X}$ یک برآورد کننده سازگار از روی نمونه‌های $X_1,X_2,\ldots,X_n$ برای μ می‌باشد؟

حل: می‌دانیم:

$$E[\overline{X}] = \mu$$

$$var(\overline{X}) = \frac{\sigma^2}{n}$$

به این ترتیب:

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} \mu = \mu$$

$$\lim_{n\to\infty} var(\overline{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

در نتیجه $\overline{X}$ یک برآورد کننده سازگار برای μ می‌باشد.

۳۳ مثال: اگر نمونه‌های $X_1, X_2, \ldots, X_n$ را از یک جامعه نرمال با میانگین μ و واریانس σ^2 استخراج کنیم در این صورت نشان دهید که S^2 یک برآورد کننده سازگار برای σ^2 می‌باشد؟

حل: قبلاً نشان دادیم که $\frac{(n-1)S^2}{\sigma^2}$ یک متغیر تصادفی خی دو با $n-1$ درجه آزادی است (χ^2_{n-1}) و بنابراین میانگین آن $n-1$ و واریانس آن $2(n-1)$ می‌باشد:

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1$$

$$var(\frac{(n-1)S^2}{\sigma^2}) = 2(n-1)$$

حال با ساده نمودن روابط فوق داریم:

$$\frac{n-1}{\sigma^2} E[S^2] = n-1 \quad \Rightarrow \quad E[S^2] = \frac{n-1}{n-1}\sigma^2 \quad \Rightarrow \quad E[S^2] = \sigma^2$$

$$var(\frac{(n-1)S^2}{\sigma^2}) = 2(n-1) \quad \Rightarrow \quad \frac{(n-1)^2}{\sigma^4} var(S^2) = 2(n-1)$$

$$var(S^2) = \frac{2\sigma^4}{n-1} \quad \Rightarrow \quad \lim_{n\to\infty} var(S^2) = \lim_{n\to\infty} \frac{2\sigma^4}{n-1} = 0$$

بنابراین S^2 یک برآورد کننده سازگار برای σ^2 می‌باشد.

S-۹

۳۴ ۱۰ . ۴ آماره کافی

اگر بتوان یک آماره یافت که تمام اطلاعات را در بازه یک پارامتر مجهول خلاصه کند آنگاه آنرا یک آماره کافی برای پارامتر مجهول می‌نامیم. آماره‌های کافی به صورت زیر تعریف می‌شود:

تعریف: اگر نمونه تصادفی $X_1, X_2, \ldots, X_n$ را از یک متغیر تصادفی با پارامتر مجهول θ داشته باشیم در این صورت آمارهٔ $y=g(X_1, X_2, \ldots, X_n)$ را یک آماره کافی برای θ می‌نامیم اگر و فقط اگر برای هر مقدار $Y=y$ توزیع مشروط هر آماره بفرم $h(X_1, X_2, \ldots, X_n)$ و با شرط $Y=y$ به θ وابسته نباشد.

به این معنی که اگر مقدار آماره کافی معلوم باشد $(Y=y)$ آنگاه به خود مقادیر نمونه نیازی نیست و مقادیر نمونه اطلاعات بیشتری در بازه پارامتر مجهول θ نسبت به آماره کافی بدست آمده، نمی‌دهند.

۳۵ مثال ۱۴: فرض می‌کنیم $X_۲,X_۱$ یک نمونه از متغیر تصادفی برنولی با پارامتر مجهول p باشند در این صورت نشان دهید که $Y=X_۱+X_۲$ یک آماره کافی برای p می‌باشد؟

حل: تابع احتمال متغیر تصادفی برنولی بفرم زیر است:

$$p_X(x) = p^x (۱-p)^{۱-x} \qquad x=۰,۱$$
$$= ۰ \qquad \text{سایر مقادیر}$$

حال تعریف می‌کنیم:

$Y=X_۱+X_۲=g(X_۱,X_۲)$ باید نشان دهیم اگر $V=h(X_۱,X_۲)$ یک آماری دیگر باشد در این صورت $P_{V,Y}(V|Y=y)$ به θ وابسته نمی‌باشد. برای این منظور ابتدا لازم است توزیع توام V و Y را محاسبه کنیم که عبارتست از:

$$P_{V,Y}(V=h(۰,۰)\,,\,Y=۰) = (۱-p)(۱-p) = (۱-p)^۲$$
$$P_{V,Y}(V=h(۰,۱)\,,\,Y=۱) = p(۱-p)$$
$$P_{V,Y}(V=h(۱,۰)\,,\,Y=۰) = (۱-p)p$$
$$P_{V,Y}(V=h(۱,۱)\,,\,Y=۲) = p.p = p^۲$$

حال توجه کنید که توزیع $Y=X_۱+X_۲$ یک توزیع دو جمله‌ای با پارامتر $n=۲$ و p می‌باشد بنابراین: $.P_Y(y) = \binom{۲}{y} p^y (۱-p)^{۲-y}$

مقدار توزیع مشروط برای هر یک از مقادیر $Y=y$ عبارتست از:

$$P_{V,Y}(V=h(۰,۰)\,|\,Y=۰) = \frac{(۱-p)^۲}{\binom{۲}{۰} p^۰ (۱-p)^{۲-۰}} = ۱$$

$$P_{V,Y}(V=h(۰,۱) \text{ یا } h(۱,۰)\,|\,Y=۱) = \frac{p(۱-p)}{\binom{۲}{۱} p(۱-p)} = \frac{۱}{۲}$$

$$P_{V,Y}(V=h(۱,۱)\,|\,Y=۲) = \frac{p^۲}{\binom{۲}{۲} p^۲ (۱-p)^۰} = ۱$$

بنابراین اگر V هر آماره دلخواهی باشد، تابع احتمال مشروط آن به p وابسته نمی‌باشد در نتیجه $Y=X_۱+X_۲$ یک آماره کافی برای پارامتر مجهول p می‌باشد.

۳۶ محاسبه کافی بودن یک آماره با توجه به تعریف کار بسیار مشکلی است، اما با استفاده از روش تجزیه فیشر- نیمن یک راه نسبتاً ساده برای اثبات کافی بودن یک آماره وجود دارد:

قضیه: $\hat{\theta}$را یک برآورگر کافی θ می‌نامیم اگر و فقط اگر تابع چگالی احتمال توام نونه‌ها را بتوان بفرم زیر تجزیه کرد:

$$f(x_۱,x_۲,\ldots,x_n.\theta) = \prod_{i=۱}^{n} f_X(x_i) = g(\hat{\theta},\theta)\,.\,h(x_۱,x_۲,\ldots,x_n)$$

مثال ۱۵: اگر یک نمونه nتایی از یک متغیر تصادفی برنولی داشته باشیم نشان دهید $Y=\sum_{i=۱}^{n} X_i$ یک آماره کافی برای p می‌باشد؟

حل:

$$\prod_{i=1}^{n} P_X(x_i) = \prod_{i=1}^{n} \left[p^{x_i} (1-p)^{1-x_i}\right] = p^{\sum x_i} (1-p)^{n-\sum x_i}$$

حال داریم:

$$\begin{cases} g(\hat{\theta},\theta) = p^{\sum x_i} (1-p)^{1-x_i} \\ h(x_1, x_2, \ldots, x_n) = 1 \end{cases}$$

بنابراین $Y = \sum_{i=1}^{n} X_i$ یک آماره کافی برای p می‌باشد.

۳۷ توجه کنید هرگاه $\hat{\theta}$ یک برآورد کننده نااریب برای θ بر حسب یک آماره کافی باشد آنگاه می‌توان نشان داد که $\hat{\theta}$ دارای کمترین واریانس در میان سایر برآورد کننده‌های نااریب برای θ می‌باشد. به همین دلیل است که برآورد کننده‌ها را بر پایه آماره‌های کافی بدست می‌آورند.

قبلاً نشان دادیم که $(\frac{n-1}{n})\, S^2$ یک برآورد کننده نااریب برای σ^2 متغیر تصادفی نرمال می‌باشد، حال با استفاده از برآورد کننده کافی یک برآورد کننده نااریب و کافی برای متغیر تصادفی نرمال بدست می‌آوریم.

برای متغیر تصادفی نرمال با میانگین μ و واریانس σ^2 داریم:

$$\prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{\sigma}\right)^n \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} e^{\frac{-\sum(x_i-\mu)^2}{2\sigma^2}}$$

$$= g(\hat{\theta},\theta) \,.\, h(x_1, x_2, \cdots, x_n)$$

که در آن:

$$h(x_1, x_2, \cdots, x_n) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}}$$

از طرفی به جای μ نیز می‌توان از برآورد کافی آن $\overline{X}$ استفاده نمود. در نتیجه $\sum (x_i - \overline{X})^2$ یک آماره کافی برای σ^2 می‌باشد. از طرفی قبلاً نشان دادیم که:

$$E\left[\frac{1}{n} \sum (x_i - \overline{X})^2\right] = \frac{n-1}{n}\sigma^2$$

$$\Rightarrow E\left[\frac{n-1}{n} \underbrace{\left(\frac{1}{n-1}\sum (x_i-\overline{X})^2\right)}_{S^2}\right] = E\left[\frac{n-1}{n} S^2\right] = \frac{n-1}{n}\sigma^2$$

$$\Rightarrow \frac{n-1}{n} E[S^2] = \frac{n-1}{n}\sigma^2 \quad \Rightarrow \quad E[S^2] = \sigma^2$$

بنابراین خود S^2 یک برآورد کننده نااریب و کافی برای σ^2 می‌باشد. و دارای کوچکترین واریانس در میان چنین برآورد کننده‌هایی می‌باشد.

۳۸ ۱۰ . ۵ خاصیت پایایی برآورد کننده‌ها

یکی از نتایج کلی که از برآورد کننده‌های حداکثر احتمال بدست می‌آید خاصیت پایایی برآورد کننده حداکثر احتمال است. اگر $\hat{\theta}$ برآورد کننده حداکثر احتمال θ باشد و بخواهیم تابعی از θ مثل $\gamma = h(\theta)$ را برآورد کنیم کافیست به جای θ مقدار برآورد شده آنرا که $\hat{\theta}$ می‌باشد جایگزین کنیم در این صورت برآورد γ عبارتست از:

$$\hat{\gamma} = h(\hat{\theta})$$

این مطلب نشان می‌دهد که برای محاسبه برآورد تابعی از پارامتر مجهول نیازی به محاسبه مجدد برآورد کننده نمی‌باشد بلکه کافیست یکبار برآورد برای پارامتر θ بدست آید. توجه کنید که این مطلب فقط برای برآورد کننده‌هایی که به روش حداکثر احتمال بدست می‌آیند صادق می‌باشد.

مثال ۱۶: اگر X یک متغیر تصادفی هندسی نوع اول باشد مطلوبست:

الف) برآورد پارامتر p به روش حداکثر احتمال.

ب) برآورد تابع $h(p) = p(X \le ۲)$ با استفاده از خاصیت پایایی.

حل: تابع احتمال توزیع هندسی نوع اول عبارتست از:

$$P_X(x) = p(۱-p)^x \qquad x = ۰,۱,۲,\cdots$$
$$= ۰ \qquad \text{سایر مقادیر}$$

$$L(p) = \prod_{i=۱}^{n} p(۱-p)^{x_i} = p^n (۱-p)^{\sum x_i}$$

$$H(p) = Ln(L(p)) = n\, Ln(p) + \sum x_i\, Ln(۱-p)$$

$$\frac{dH}{dP} = \frac{n}{p} - \frac{\sum x_i}{۱-\hat{p}} \quad \Rightarrow \quad \frac{۱-\hat{p}}{\hat{p}} = \frac{۱}{n}\sum x_i = \overline{X} \quad \Rightarrow \quad \frac{۱}{\hat{p}} - ۱ = \overline{X} \quad \Rightarrow \quad \hat{p} = \frac{۱}{۱+\overline{X}}$$

بنابراین برآورد p به روش حداکثر احتمال برابر $\frac{۱}{۱+\overline{X}}$ می‌باشد.

ب)

$$p(X \le ۲) = p(X=۰) + p(X=۱) + p(X=۲)$$
$$= p + p(۱-p) + p(۱-p)^۲$$

حال داریم:

$$\gamma = h(p) = p(۱+(۱-p)+(۱-p)^۲)$$

بنابراین خاصیت پایایی $\hat{\gamma} = h(\hat{p})$ بنابراین:

$$\hat{\gamma} = h(\hat{p}) = \frac{۱}{۱+\overline{X}}\left(۱+\left(۱-\frac{۱}{۱+\overline{X}}\right) + \left(۱-\frac{۱}{۱+\overline{X}}\right)^۲\right)$$

S-۱۰

۴۰ ۱۰ . ۶ برآورد فاصله‌ای

در برآورد نقطه‌ای با توجه به نمونه‌های بدست آمده نشان دادیم برای پارامترهای مجهول جامعه می‌توان عددی بدست آورد که با خطای کمی بیانگر مقدار واقعی پارامتر مجهول باشد. اما واقعاً نمی‌دانیم با چه مقدار خطایی پارامتر مجهول را برآورد کرده‌ایم برآورد فاصله‌ای روشی است که در آن برای پارامتر مجهول یک بازه در نظر می‌گیریم و نشان می‌دهیم که با احتمالی معین پارامتر مجهول در بازه مورد نظر قرار دارد.

فرض کنید نمونه‌های $X_۱, X_۲, \ldots, X_n$ را از یک متغیر تصادفی X بدست آورده باشیم در این صورت فاصله $(L_۱, L_۲)$ را طوری می‌سازیم که با احتمال بسیار زیاد مثلاً ۹۵٪ شامل پارامتر مجهول جامعه (مثلاً μ) باشد. به فاصله $(L_۱, L_۲)$ یک فاصله اطمینان می‌گویند زیرا مطمئن هستیم که به احتمال زیاد (مثلاً ۹۵٪) این فاصله شامل پارامتر مجهول می‌باشد. مقادیر $L_۱, L_۲$ هر کدام آماره‌ای هستند که بر اساس مقادیر معلوم نمونه‌ها بدست می‌آیند. حال به تعریف دقیق فاصله اطمینان توجه کنید:

هرگاه مقادیر $X_۱, X_۲, \ldots, X_n$ تعداد n نمونه تصادفی بدست آمده از یک متغیر تصادفی X با پارامتر مجهول θ باشند. در این صورت می‌گوییم دو آماره $L_۱, L_۲$ یک $(۱-\alpha)$ ۱۰۰ درصد فاصله اطمینان برای θ تشکیل می‌دهند. هرگاه:

$$p(L_۱ \le \theta \le L_۲) \ge ۱-\alpha$$

توجه کنید که $۱-\alpha$ مقدار احتمال می‌باشد و بنابراین در بازه ۰ تا ۱ قرار دارد، در نتیجه برای بیان آن بصورت درصد آنرا درصد ضرب کرده‌ایم و به صورت $(۱-\alpha)$ ۱۰۰٪ می‌نویسیم.

در برآورد فاصله‌ای از آماره‌های بدست آمده از فصل قبل استفاده می‌کنیم. اینکه کدام پارامتر جامعه مجهول باشد، تعین می‌کند که از کدام آماری در ساختن فاصله اطمینان استفاده کنیم. مثال بعد روش ساختن فاصله اطمینان برای جامعه نرمال را نشان می‌دهد.

۴۱ مثال ۱۷: در شهر A بارش باران یک متغیر تصادفی نرمال با میانگین μ سانتی‌متر و واریانس ۴ می‌باشد اگر میزان بارش را در طول ۱۶ روز اندازه بگیریم و میانگین ۱۰ سانتی‌متر را بدست بیاوریم، در این صورت با احتمال ۹۵٪ میانگین بارش باران در شهر A در چه فاصله اطمینانی قرار می‌گیرد؟

حل: معلومات مساله عبارتند از:

$n = ۱۶$

$\overline{X} = ۱۰$

$\mu =$ مجهول

$\sigma^۲ = ۴ \rightarrow \sigma = ۲$

ابتدا توجه کنید که با توجه به جدول متغیر تصادفی نرمال استاندارد رابطه زیر برقرار می‌باشد:

$$p(-۱/۹۶ \le Z \le ۱/۹۶) = ۰/۹۵ \quad (\text{رابطه } ۱۰-۲)$$

زیرا داریم:

$$N_Z(۱/۹۶) = p(Z \le ۱/۹۶) = ۰/۹۷۵$$

$$N_Z(-۱/۹۶) = ۱ - N_Z(۱/۹۶) = ۰/۰۲۵$$

$$\Rightarrow p(-۱/۹۶ \le Z \le ۱/۹۶) = p(Z \le ۱/۹۶) - p(Z \le ۱/۹۶) = ۰/۹۷۵ - ۰/۰۲۵ = ۰/۹۵$$

۴۲ در فصل نمونه‌گیری نشان دادیم که اگر واریانس جامعه معلوم باشد آمارهٔ $Z = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد و از طرفی با توجه به رابطه ۱۰ − ۲ چون Z نیز یک متغیر تصادفی نرمال استاندارد می‌باشد می‌توان نوشت:

$$P(-۱/۹۶ \le Z \le ۱/۹۶) = P\left(-۱/۹۶ \le \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le ۱/۹۶\right)$$

با ساده نمودن نامساوی $-۱/۹۶ \le \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le ۱/۹۶$ بدست می‌آوریم:

$$\overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

بنابراین رابطه ۱۰ − ۲ بصورت زیر بدست می‌آید:

$$p\left(\overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)\right) = ۰/۹۵$$

همانطور که ملاحظه می‌کنید برای μ یک فاصله اطمینان ۹۵ درصدی به شکل $(L_۱, L_۲)$ بدست آوردیم که در آن $L_۱, L_۲$ عبارتند از:

$$L_۱ = \overline{X} - ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$L_۲ = \overline{X} + ۱/۹۶\left(\frac{\sigma}{\sqrt{n}}\right)$$

حال با جایگذاری $\overline{X}$ و σ و n در رابطه بدست می‌آوریم:

$$L_۱ = ۱۰ - ۱/۹۶\left(\frac{۲}{\sqrt{۱۶}}\right) = ۹/۰۲$$

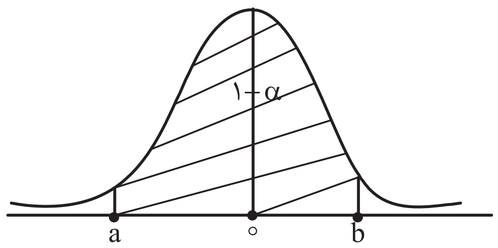
$$L_۲ = ۱۰ + ۱/۹۶\left(\frac{۲}{\sqrt{۱۶}}\right) = ۱۰/۹۸$$

در نتیجه:

$$p(۹/۰۲ \le \mu\ ۱۰۹۸) = ۰/۹۵$$

یعنی ۹۵ درصد اطمینان داریم که میانگین جامعه در بازه $(۹/۰۲\ ,\ ۱۰/۹۸)$ قرار دارد حال برای اینکه بتوانیم یک فاطله اطمینان $۱۰۰(۱-\alpha)$ درصدی در حالت کلی برای جامعه نرمال با میانگین مجهول و واریانس معلوم بسازیم ابتدا مفهوم $Z_{۱-\frac{\alpha}{۲}}$ را معرفی می‌کنیم.

۴۳ نمودار زیر منحنی نمایش توزیع نرمال استاندارد می‌باشد ملاحظه می‌کنید که مساحت مشخص شده در شکل برابر $۱-\alpha$ می‌باشد:



برای پیدا کردن فاصله اطمینان می‌بایستی اعداد a و b را طوری پیدا کرد که:

$$p(a \le Z \le b) = p(a \le \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le b) = ۱-\alpha$$

با فرض اینکه فاصله متقارن باشد مساحت زیر منحنی بعد از نقطه b و قبل از نقطه a برابر $\frac{\alpha}{۲}$ می‌باشد. در نتیجه مساحت کل زیر منحنی قبل از نقطه b برابر است با $۱-\frac{\alpha}{۲} = ۱-\alpha + \frac{\alpha}{۲}$ که آنرا با $Z_{۱-\frac{\alpha}{۲}}$ نمایش می‌دهیم. به این ترتیب $Z_{۱-\frac{\alpha}{۲}}$ نقطه‌ای روی محور افقی است که مساحت زیر منحنی قبل از آن برابر $۱-\frac{\alpha}{۲}$ می‌باشد. این مطلب با توجه به تعریف تابع توزیع نرمال استاندارد بصورت زیر نیز قابل بیان است:

$$p(Z \le Z_{۱-\frac{\alpha}{۲}}) = ۱-\frac{\alpha}{۲}$$

حال می‌توان به جای b و a در $p(a \le Z \le b)$ مقادیر $Z_{۱-\frac{\alpha}{۲}}$ و $-Z_{۱-\frac{\alpha}{۲}}$ را جایگزین نمود و با ساده نمودن نامساوی خواهیم داشت:

$$p(-Z_{۱-\frac{\alpha}{۲}} \le Z \le Z_{۱-\frac{\alpha}{۲}}) = p(-Z_{۱-\frac{\alpha}{۲}} \le \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le Z_{۱-\frac{\alpha}{۲}})$$

$$p=(-Z_{۱-\frac{\alpha}{۲}}\ (\frac{\sigma}{\sqrt{n}}) \le \overline{X}-\mu \le Z_{۱-\frac{\alpha}{۲}}\ (\frac{\sigma}{\sqrt{n}}))$$

$$p=(\overline{X}-Z_{۱-\frac{\alpha}{۲}}\ (\frac{\sigma}{\sqrt{n}}) \le \mu \le \overline{X}+Z_{۱-\frac{\alpha}{۲}}\ (\frac{\sigma}{\sqrt{n}}))=۱-\alpha$$

بنابراین با توجه به آخرین رابطه یک فاصله $۱۰۰(۱-\alpha)$ درصد برای میانگین (μ) جامعه نرمال با معلوم بودن $\sigma^۲$ بدست آوردیم.

۴۴ مثال ۱۸: میزان فروش ماهانه یک فروشگاه یک متغیر تصادفی نرمال با میانگین مجهول μ و انحراف معیار ۵۰ هزار تومان در ماه می‌باشد. نتایج حاصل از ۲۵ ماه فروش نشان می‌دهند که میانگین فروش در طول این ۲۵ ماه عبارتست از ۵۵۰ هزار تومان. اگر ۹۰ درصد اطمینان داشته باشیم میانگین (μ) در فاصله $(a\ ,b)$ می‌باشد مقادیر a و b را مشخص کنید؟

حل: معلومات مثال عبارتند از:

$n = ۲۵$

$\overline{X} = ۵۵۰ \qquad \mu = \text{مجهول} \qquad \sigma = ۵۰$

می‌دانیم فاصله اطمینان از رابطه زیر بدست می‌آید:

$$p\left(\overline{X} - Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right) \leq \mu \leq \overline{X} + Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right)\right) = ۱-\alpha$$

ابتدا می‌بایستی مقدار α را بدست بیاوریم و سپس از روی آن $Z_{۱-\frac{\alpha}{۲}}$ بدست می‌آید.

$$\%۱۰۰\,(۱-\alpha)\ \%۹۰ \quad \Rightarrow \quad ۱-\alpha = ۰/۹ \quad \Rightarrow \quad \alpha = ۰/۱$$

$$Z_{۱-\frac{\alpha}{۲}} \overset{\alpha=۰/۱}{=} Z_{۱-\frac{۰/۱}{۲}} = Z_{۰/۹۵}$$

۴۵ با توجه به جدول توزیع نرمال استاندارد $Z_{۰/۹۵}$ عبارتست از:

$$N_Z(Z_{۰/۹۵}) = ۰/۹۵ \quad \Rightarrow \quad Z_{۰/۹۵} = ۱/۶۴۵$$

بنابراین بدست می‌آوریم:

$$p\left(۵۵۰ - ۱/۶۴۵\left(\frac{۵۰}{\sqrt{۲۵}}\right) \leq \mu \leq ۵۵۰ + ۱/۶۴۵\left(\frac{۵۰}{\sqrt{۲۵}}\right)\right) = ۰/۹$$

$$\Rightarrow p\,(۵۳۳/۵۵ \leq \mu \leq ۵۶۶/۴۵) = ۰/۹$$

بنابراین با احتمال ۰/۹ متوسط درآمد فروشگاه در بازه (۵۳۳/۵۵ , ۵۶۶/۴۵) قرار دارد. توجه کنید که طول بازه فاصله اطمینان در حالت کلی برابر است با:

$$L_۲ - L_۱ = \left(\overline{X} + Z_{۱-\frac{\alpha}{۲}}\frac{\sigma}{\sqrt{n}}\right) - \left(\overline{X} - Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right)\right) \quad \Rightarrow L_۲ - L_۱ = ۲Z_{۱-\frac{\alpha}{۲}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

ملاحظه می‌کنید که طول بازه مستقل از $\overline{X}$ می‌باشد و تنها به تعداد نمونه‌ها وابسته می‌باشد. یعنی می‌توان تعداد نمونه‌ها را طوری انتخاب نمود که هر فاصله اطمینان دلخواهی برای μ بدست آید.

به عنوان مثال در مثال قبل با ۲۵ نمونه طول بازه عبارتست از:

$$L_۲ - L_۱ = ۲(۱/۶۴۵)\left(\frac{۵۰}{\sqrt{۱۰۰}}\right) = ۱۶/۴۵$$

پیداست که با ۱۰۰ نمونه طول بازه نصف شده است.

۱۰ . ۶ . ۱ فاصله اطمینان برای جامعه نرمال با میانگین مجهول μ و واریانس مجهول $\sigma^۲$. حال فاصله اطمینان را برای حالتی بدست می‌آوریم که علاوه بر میانگین جامعه، واریانس نیز مجهول باشد. از فصل نمونه‌گیری می‌دانیم متغیر تصادفی $T = \dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ یک متغیر تصادفی t با $n-۱$ درجه آزادی می‌باشد بنابراین مجدداً مشابه روشی که قبلاً برای بدست آوردن فاصله اطمینان ارایه کردیم داریم:

$$p\left(-t_{۱-\frac{\alpha}{۲}} \leq t_{(n-۱)} \leq t_{۱-\frac{\alpha}{۲}}\right) = p\left(-t_{۱-\frac{\alpha}{۲}} \leq \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \leq t_{۱-\frac{\alpha}{۲}}\right)$$

$$= p\left(\overline{X} - t_{۱-\frac{\alpha}{۲}}\left(\frac{S}{\sqrt{n}}\right) \leq \mu \leq \overline{X} + t_{۱-\frac{\alpha}{۲}}\left(\frac{S}{\sqrt{n}}\right)\right) = ۱-\alpha$$

که در آن $t_{1-\frac{\alpha}{2}}$ عبارتست از: $t(n-1)\ (T \le t_{1-\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$.

بنابراین فاصله اطمینان در این حالت عبارتست از:

$$L_1 = \overline{X} - t_{1-\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right)$$

$$L_2 = \overline{X} + t_{1-\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right)$$

۴۷ **مثال ۱۹:** میزان مصرف بنزین در هر ماه یک متغیر تصادفی نرمال با کیانگین مجهول μ و واریانس σ^2 می‌باشد اگر اطلاعات زیر از یک نمونه در طول ۶ ماه بدست آمده باشد مطلوبست یک فاصله اطمینان ۹۰ درصدی برای متوسط میزان مصرف بنزین؟

$$\sum_{i=1}^{6} x_i = 246 \qquad \sum_{i=1}^{6} x_i^2 = 11200$$

حل:

$$\overline{X} = \frac{1}{n}\sum_{1}^{n} x_i = \frac{1}{6}(146) = 41$$

$$S^2 = \frac{1}{n-1}\left(\sum_{1}^{n}(x_i-\overline{X})^2\right) = \frac{n}{n-1}\left(\frac{1}{n}\sum_{1}^{n}(x_i-\overline{X})^2\right)$$

$$= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{X}^2\right)$$

$$\Rightarrow S^2 = \frac{1}{5}(11200 - 6(41)^2) = \frac{1}{5}(1114) = 222/8$$

$$\Rightarrow S = 14/92$$

حال برای یک فاصله اطمینان ۹۰ درصدی داریم:

$$1-\alpha = 0/9 \Rightarrow \alpha = 0/1$$

$$t_{1-\frac{\alpha}{2}} = t_{0/95}$$

۴۸ با توجه به جدول t با ۵ درجه آزادی مقدار $t_{0/95}$ برابر ۲/۰۱۵ بدست می‌آید.

در نتیجه:

$$L_1 = \overline{X} - t_{0/95}\left(\frac{S}{\sqrt{n}}\right) = 41 - 2/015\left(\frac{14/92}{\sqrt{6}}\right) = 28/72$$

$$L_2 = \overline{X} + t_{0/95}\left(\frac{S}{\sqrt{n}}\right) = 41 + 2/015\left(\frac{14/92}{\sqrt{6}}\right) = 53/27$$

بنابراین با ۹۰ درصد اطمینان می‌توانیم بگوییم که متوسط مصرف بنزین ماهانه مابین ۲۸/۷۲ و ۵۳/۲۷ میلیون لیتر خواهد بود.

برای حالتی که σ^2 مجهول باشد طول بازهٔ فاصله اطمینان عبارتست از:

$$L_2 - L_1 = 2t_{1-\frac{\alpha}{2}}\left(\frac{S}{\sqrt{n}}\right)$$

اما توجه می‌کنید که در این حالت طول بازه به متغیر S نیز وابسته می‌باشد و در نتیجه به هیچ عنوان نمی‌توان اندازه نمونه را طوری انتخاب نمود که دقیقاً یک بازه با طول معین بدست آید.

نکته: قبلاً نشان دادیم که اگر تعداد نمونه‌ها زیاد باشد $(n > ۳۰)$ $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ و $\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ هر دو به توزیع نرمال استاندارد میل می‌کنند، در نتیجه تمام مطالب گفته شده در حالتی که $n > ۳۰$ باشد برای متغیر تصادفی X با هر توزیعی برقرار می‌باشد.

۴۹ ۲ . ۶ . ۱۰ فاصله اطمینان برای واریانس جامعه نرمال با میانگین مجهول μ و واریانس $\sigma^۲$.

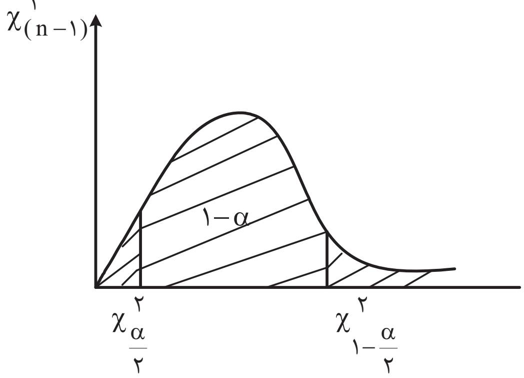
روش بدست آوردن فاصله اطمینان برای واریانس یک جامعه نرمال با میانگین مجهول μ و واریانس $\sigma^۲$ کاملاً مشابه روشی است که برای میانگین جامعه استفاده شد. در این حالت از آماره $\frac{(n-۱)\,S^۲}{\sigma^۲}$ استفاده می‌کنیم. می‌دانیم که $\frac{(n-۱)\,S^۲}{\sigma^۲}$ یک متغیر تصادفی خی دو با $n-۱$ درجه آزادی است.

می‌دانیم:

$$p\,(\chi^۲_{\frac{\alpha}{۲}} \le \chi^۲_{(n-۱)} \le \chi^۲_{۱-\frac{\alpha}{۲}}) = ۱-\alpha$$

$$\Rightarrow p\,(\chi^۲_{\frac{\alpha}{۲}} \le \frac{(n-۱)S^۲}{\sigma^۲} \le \chi^۲_{۱-\frac{\alpha}{۲}}) = ۱-\alpha \qquad (۳-۱۰)$$

توجه کنید که مقدار ابتدایی بازه $(L_۱)$ برابر $\chi^۲_{\frac{\alpha}{۲}}$ می‌باشد زیرا مطابق شکل منحنی توزیع $\chi^۲_{(n-۱)}$ مساحت زیر منحنی قبل از $\chi^۲_{\frac{\alpha}{۲}}$ برابر $\frac{\alpha}{۲}$ می‌باشد.



به عبارتی $p\,(\chi^۲_{(n-۱)} \le \chi^۲_{\frac{\alpha}{۲}}) = \frac{\alpha}{۲}$.

حال با ساده نمودن رابطه (۱۰-۳) داریم:

$$p(\frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} \le \sigma^۲ \le \frac{(n-۱)\,S^۲}{\chi^۲_{\frac{\alpha}{۲}}}) = ۱-\alpha$$

بنابراین بازه زیر یک فاصله اطمینان $(۱-\alpha)۱۰۰$ درصدی برای واریانس بدست می‌دهد:

$$L_۱ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} \qquad , \qquad L_۲ = \frac{(n-۱)\,S^۲}{\chi^۲_{\frac{\alpha}{۲}}}$$

۵۰ مثال ۲۰: کارخانه‌ای ظروف پلاستیکی تولید می‌کند، ضایعات تولید شده توسط ماشین‌های تولیدی یک متغیر تصادفی نرمال با میانگین μ و واریانس $\sigma^۲$ می‌باشد، اگر پراکندگی ضایعات تولید شده در روز زیاد شود (واریانس تعداد ضایعات زیاد شود). می‌بایستی ماشین‌های تولیدی را تعمیر

نمود. برای تخمین واریانس یک نمونه ۲۵ تایی از تعداد ضایعات تولید شده در هر روز می‌گیریم و نتایج $\sum x_i = ۱۷۰$, $\sum x_i^۲ = ۲۷۵۰$ را بدست می‌آوریم.
در صورتی که یک فاصله اطمینان ۹۰ درصدی برای واریانس بدست بیاوریم و طول فاصله اطمینان بیشتر از ۵۰ شود می‌بایستی ماشینهای تولیدی را تعمیر کنیم، آیا نیازی به تعمیر می‌باشد؟
حل: معلومات مثال عبارتند از:

$$n = ۲۵$$

$$\overline{X} = \frac{۱}{۲۵}(۱۷۰) = ۶/۸$$

$$S^۲ = \frac{۱}{n-۱}\left(\sum x_i^{\,۲} - n\overline{X}^۲\right) = \frac{۱}{۲۴}\left(۲۷۵۰ - ۲۵(۶/۸)^۲\right) = ۶۶/۴۱$$

$$\Rightarrow S = ۸/۱۴$$

$$۱-\alpha = ۰/۹ \quad \rightarrow \quad \alpha = ۰/۱ \quad \rightarrow \quad \chi^۲_{\frac{\alpha}{۲}}(n-۱) = \chi^۲_{۰/۰۵}(۲۴) = ۱۳/۸$$

$$\chi^۲_{۱-\frac{\alpha}{۲}}(n-۱) = \chi^۲_{۰/۹۵}(۲۴) = ۳۶/۴$$

بنابراین:

$$L_۱ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} = \frac{(۲۴)\,(۶۶/۴۱)}{۳۶/۴} = ۴۳/۷۸$$

$$L_۲ = \frac{(n-۱)\,S^۲}{\chi^۲_{۱-\frac{\alpha}{۲}}} = \frac{۲۴\,(۶۶/۴۱)}{۱۳/۸} = ۱۱۵/۴۹$$

$$L_۲ - L_۱ = ۱۱۵/۴۹ - ۴۳/۷۸ = ۷۱/۷$$

با توجه به اینکه طول بازه فاصله اطمینان بیشتر از ۵۰ واحد شده است بنابراین می‌بایستی ماشینهای تولیدی تعمیر شوند.

۵۱ . ۱۰ . ۶ . ۳ فاصله اطمینان برای متغیر تصادفی نمایی با پارامتر مجهول λ

می‌دانیم متغیر تصادفی نمایی عموماً مدلی برای طول عمر وسایل الکتریکی، تجهیزات و ... می‌باشد. بنابراین اگر بخواهیم اطمینان داشته باشیم که متوسط طول عمر یک متغیر تصادفی نمایی حداقل برابر تعداد ساعات معین است در این صورت می‌بایستی $\frac{۱}{\lambda}$ از عدد معین بیشتر باشد بنابراین برای متغیر تصادفی نمایی تنها یک فاصله اطمینان یک طرفه می‌سازیم.
هرگاه $X_۱, X_۲, \ldots, X_n$ یک نمونه تصادفی از یک متغیر تصادفی نمایی با پارامتر λ باشد آنگاه:

$$p\left(\lambda \le \frac{\chi^۲_{۱-\alpha}}{۲n\overline{X}}\right) = ۱-\alpha$$

یک فاصله اطمینان $(۱-\alpha)$ ۱۰۰ درصدی برای λ تشکیل می‌دهد که در آن $\chi^۲_{۱-\alpha}$ مقداری از توزیع $\chi^۲$ با $۲n$ درجه آزادی است که مساحت زیر منحنی نمایش آن برابر $۱-\alpha$ باشد.

۵۲ مثال ۲۱: طول عمر یک مدار الکتریکی یک متغیر تصادفی نمایی با پارامتر λ می‌باشد تعداد ۲۰ عدد از این مدارها را آزمایش می‌کنیم تا زمانی که از کار بیفتند. و ملاحظه می‌کنیم که مجموع طول آنها برابر ۲۳۶۰ ساعت می‌شود مطلوبست: تخمین متوسط طول عمر مدار مربوط با فاصله اطمینان ۹۰ درصدی:

حل: معلومات مساله عبارتند از:

$$n = ۲۰$$

$$\sum_{i=۱}^{۲۰} x_i = ۲۳۶۰ \quad \rightarrow \quad \overline{X} = \frac{۱}{n} \sum_{۱}^{n} x_i = \frac{۱}{۲۰}(۲۳۶۰) = ۱۱۸$$

$$۱-\alpha = ۰/۹ \quad \rightarrow \quad \chi^۲_{۱-\alpha}(۲n) = \chi^۲_{۰/۹}(۴۰) = ۵۱/۸$$

بنابراین یک فاصله اطمینان یک طرفه ۹۰ درصدی برای λ عبارتست از:

$$p\,(\lambda \le \frac{\chi^۲_{۱-\alpha}}{۲n\overline{X}}) = ۱-\alpha \;\Rightarrow\; L_۲ = \frac{۵۱/۸}{۲(۲۰)\,(۱۱۸)} = ۰/۰۱۰۹$$

توجه کنید که مقدار متوسط طول عمر برای توزیع نمایی $E[X] = \frac{۱}{\lambda}$ می‌باشد بنابراین حداقل متوسط طول عمر مدار عبارتست از:

$$\frac{۱}{۰/۰۱۰۹} = ۹۱/۱۱$$

به عبارتی ۹۰ درصد اطمینان داریم که هر مدار حداقل ۹۱/۱۱ ساعت کار می‌کند.

S-۱۳

۵۳ ۱۰ . ۶ . ۴ فاصله اطمینان برای نسبت p با تعداد نمونه‌های زیاد در توزیع برنولی

نمونه‌های $X_۱, X_۲, \ldots, X_n$ را از یک متغیر تصادفی برنولی با پارامتر مجهول p استخراج می‌کنیم اگر تعداد نمونه‌ها زیاد باشد در این صورت می‌دانیم متغیر تصادفی $\overline{X}$ تقریباً بصورت نرمال با میانگین p و واریانس $\frac{p(۱-p)}{n}$ توزیع می‌شود. به این ترتیب می‌توان نشان داد که یک فاصله اطمینان $(۱-\alpha)$ ۱۰۰ درصدی برای p عبارتست از:

$$p\,(\,\overline{X} - Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \le p \le \overline{X} + Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}(۱-\overline{X})}{n}}\;) = ۱-\alpha$$

توجه کنید که در محاسبه مقدار $\overline{X}$ برابر است با نسبت تعداد موفقیتها به کل تعداد نمونه‌ها به عنوان مثال اگر از ۲۰ مرتبه تیراندازی به هدف ۱۵ بار هدف مورد اصابت قرار گیرد در این صورت $\overline{X} = \hat{p} = \frac{۱۵}{۲۰}$ خواهد بود.

۵۴ مثال ۲۲: احتمال اینکه یک بازیکن فوتبال یک ضربه پنالتی را گل کند یک متغیر تصادفی برنولی با پارامتر p می‌باشد. بازیکن مربوطه در تمرینات از تعداد ۲۰۰ ضربه پنالتی ۱۶۰ مرتبه را موفق به زدن گل شده است. اگر بخواهیم در مسابقات بازیکن مربوطه را مامور زدن ضربه پنالتی کنیم با چه احتمالی می‌توانیم ۹۵ درصد اطمینان داشته باشیم که وی گل می‌زند؟

حل: معلومات مساله عبارتند از:

$$n = ۲۰۰$$

$$\overline{X} = \hat{p} = \frac{۱۶۰}{۲۰۰} = \frac{۱۶}{۲۰} = ۰/۸$$

$$۱ - \alpha = ۰/۹۵ \rightarrow \alpha = ۰/۰۵ \rightarrow Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۷۵} = ۱/۹۶$$

بنابراین یک فاصله ۹۵ درصدی برای p عبارتست از:

$$L_۱ = \overline{X} - Z_{۱-\frac{\alpha}{۲}} \left(\sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \right) = ۰/۸ - ۱/۹۶ \sqrt{\frac{۰/۸\,(۰/۲)}{۲۰۰}} = ۰/۷۴$$

$$L_۲ = \overline{X} + Z_{۱-\frac{\alpha}{۲}} \left(\sqrt{\frac{\overline{X}(۱-\overline{X})}{n}} \right) = ۰/۸ + ۱/۹۶ \sqrt{\frac{۰/۸\,(۰/۲)}{۲۰۰}} = ۰/۸۵$$

بنابراین با ۹۵ درصد اطمینان می‌توانیم بگوییم که احتمال گل شدن ضربه پنالتی توسط بازیکن مربوطه در بازه $(۰/۸۵\,,\,۰/۷۴)$ قرار دارد.

۵۵ ۱۰ . ۶ . ۵ فاصله اطمینان برای تفاضل نسبت دو جامعه با تعداد نمونه‌های زیاد

اگر دو نمونه با حجم‌های $n_۲\,,n_۱$ از دو متغیر تصادفی برنولی در اختیار داشته باشیم، در این صورت برای مقایسه پارامترهای $p_۲\,,p_۱$ متغیرهای تصادفی از تفاضل این دو نسبت استفاده می‌کنیم به این ترتیب می‌توان نشان داد که یک فاصله اطمینان $(۱-\alpha)۱۰۰$ درصدی برای تفاضل دو نسبت از دو متغیر تصادفی برنولی عبارتست از:

$$p\left(\overline{X}_۱ - \overline{X}_۲ - Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}_۱(۱-\overline{X}_۱)}{n_۱} + \frac{\overline{X}_۲(۱-\overline{X}_۲)}{n_۲}} \le p_۱ - p_۲ \le \overline{X}_۱ - \overline{X}_۲ + Z_{۱-\frac{\alpha}{۲}} \sqrt{\frac{\overline{X}_۱(۱-\overline{X}_۱)}{n_۱} + \frac{\overline{X}_۲(۱-\overline{X}_۲)}{n_۲}}\right) = ۱-\alpha$$

که در آن $\overline{X}_۱ = \hat{P}_۱\,,\,\overline{X}_۲ = \hat{P}_۲$

۵۶ مثال ۲۳: در دو شهر A و B احتمال متولد شدن نوزادان پسر را با متغیرهای تصادفی برنولی $X_۲\,,X_۱$ نشان می‌دهیم. در شهر A از ۳۰۰ نوزاد متولد شده ۱۸۰ نوزاد پسر می‌باشند و در شهر B از ۴۰۰ نوزاد متولد شده ۱۹۰ نوزاد پسر متولد شده‌اند. می‌خواهیم احتمال متولد شدن نزاد پسر را در دو شهر A و B مقایسه کنیم. برای این منظور از یک فاصله اطمینان ۹۵ درصدی استفاده می‌کنیم مطلوبست حدود فاصله اطمینان؟

حل: معلومات مساله عبارتند از:

$n_1 = ۳۰۰$

$\overline{X}_1 = \hat{p}_1 = \frac{۱۸۰}{۳۰۰} = \frac{۱۸}{۳۰} = ۰/۶$

$n_2 = ۴۰۰$

$\overline{X}_2 = \hat{p}_2 = \frac{۱۹۰}{۴۰۰} = ۰/۴۷۵$

$1-\alpha = ۰/۹۵ \rightarrow \alpha = ۰/۰۵ \rightarrow Z_{1-\frac{\alpha}{2}} = Z_{۰/۹۷۵} = ۱/۹۶$

$$L_1 = \overline{X}_1 - \overline{X}_2 - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{(1-\overline{X}_1)\,\overline{X}_1}{n_1} + \frac{\overline{X}_2(1-\overline{X}_2)}{n_2}}$$

$$= ۰/۶ - ۰/۴۷۵ - ۱/۹۶ \sqrt{\frac{(۰/۶)۰/۴}{۳۰۰} + \frac{(۰/۴۷۵)۰/۵۲۵}{۴۰۰}} = ۰/۰۵$$

$$L_2 = \overline{X}_1 - \overline{X}_2 + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\overline{X}_1(1-\overline{X}_1)}{n_1} + \frac{\overline{X}_2(1-\overline{X}_2)}{n_2}}$$

$$= ۰/۶ - ۰/۴۷۵ + ۱/۹۶ \sqrt{\frac{(۰/۶)۰/۴}{۳۰۰} + \frac{(۰/۴۷۵)۰/۵۲۵}{۴۰۰}} = ۰/۱۹۸$$

فاصله اطمینان ۹۵٪ $\quad p(۰/۰۵ \le p_1 - p_2 \le ۰/۱۹۸) = ۰/۹۵$

همچنین می‌توان نوشت: $\quad \Rightarrow p(p_2 + ۰/۰۵ \le p_1 - p_2 + ۰/۱۹۸) = ۰/۹۵$

بر اساس مساوی فوق احتمال متولد شدن نوزاد پسر در شهر A با اطمینان ۹۵ درصد در بازه $(p_2 + ۰/۰۵ \;,\; p_2 + ۰/۱۹۸)$ قرار دارد.

S-۱۴

۵۷ . ۱۰ . ۶ . ۶ فاصله اطمینان برای نسبت واریانسهای دو متغیر تصادفی نرمال

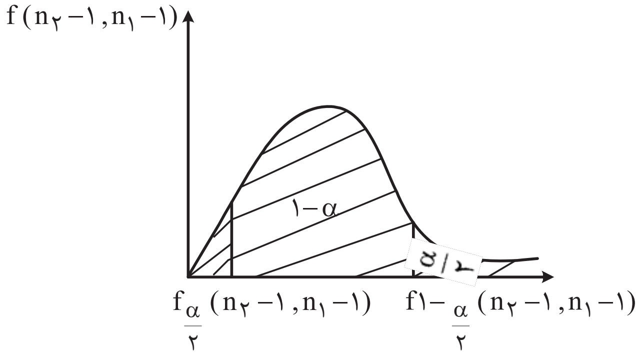
دو نمونه به حجم n_1 و n_2 را با واریانس نمونه‌ای S_1^2 و S_2^2 از دو جامعه نرمال با واریانس σ_1^2 و σ_2^2 استخراج می‌کنیم. برای مقایسه واریانسهای دو جامعه از نسبت آنها $\left(\frac{S_1^2}{S_2^2}\right)$ استفاده می‌کنیم. قبلاً نشان دادیم که متغیر تصادفی $\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2}$ دارای توزیع F با $n_1 - 1 \,,\, n_2 - 1$ درجه آزادی است به این ترتیب می‌توان فاصله اطمینان $۱۰۰(1-\alpha)$ درصدی را برای نسبت $\frac{\sigma_1^2}{\sigma_2^2}$ بصورت زیر بدست آورد:

$$p\left(\frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}\right) = 1-\alpha$$

با توجه به اینکه $f_{\frac{\alpha}{2}}(n_1, n_2) = \frac{1}{f_{1-\frac{\alpha}{2}}(n_2, n_1)}$ بنابراین رابطه فوق به صورت زیر ساده می‌شود:

$$p\left(\frac{S_1^2}{S_2^2}\,\frac{1}{f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{S_1^2}{S_2^2}\, f_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)\right) = 1-\alpha$$

که در آن $f_{\frac{\alpha}{۲}}\ (n_۱-۱, n_۲-۱)$ مقداری از متغیر تصادفی F با $n_۱-۱, n_۲-۱$ درجه آزادی است بطوریکه سطح زیر منحنی به ازای مقادیر کوچکتر از آن برابر $\frac{\alpha}{۲}$ باشد. به شکل زیر توجه کنید:



توجه کنید در تمام فواصل اطمینان که در روابط آنها از $S^۲$ استفاده شده است اگر میانگین جامعه (μ) معلوم باشد به جای استفاده از $S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\overline{X})^۲$ از $S^۲$ استفاده می‌شود که به جای $\overline{X}$ مقدار μ جایگزین شده است یعنی : $S^۲ = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\mu)^۲$.

۵۸ مثال ۲۴: در یک کارخانه از دو ماشین A و B برای برش میله‌های فولادی استفاده می‌شود در صورتی که واریانس طول میله‌های بریده شده توسط هر یک از ماشینها سه برابر دیگری شود تصمیم به تعمیر ماشین مربوطه می‌گیریم. در یک نمونه ۲۵ تایی از میله‌های بریده شده توسط ماشین A مقدار $S_۱{}^۲$ برابر ۱۷۰ می‌باشد، همچنین در یک نمونه ۳۱ تایی از میله‌های بریده شده توسط B مقدار $S_۲{}^۲$ برابر ۲۵ می‌باشد. با اطمینان ٪۹۵ برای نسبت $\frac{\sigma_۱{}^۲}{\sigma_۲{}^۲}$ چه تصمیمی می‌توان در مورد ماشین A و B گرفت؟

حل: داریم:

$$n_۱ = ۲۵ \qquad n_۲ = ۳۱$$

$$S_۱{}^۲ = ۱۷۰ \qquad S_۲{}^۲ = ۲۵$$

$$۱-\alpha = ۰/۹ \rightarrow \alpha = ۰/۱ \rightarrow f_{۱-\frac{\alpha}{۲}}(n_۱-۱, n_۲-۲) = f_{۰/۹۵}(۲۴, ۳۰) = ۱/۹۴$$

$$f_{۱-\frac{\alpha}{۲}}(n_۱-۱, n_۲-۲) = f_{۰/۹۵}(۳۰, ۲۴) = ۱/۸۹$$

به این ترتیب یک فاصله اطمینان ۹۵ درصدی عبارتست از:

$$p\left(\frac{۱۷۰}{۲۵}\frac{۱}{۱/۹۴} \le \frac{\sigma_۱{}^۲}{\sigma_۲{}^۲} \le \frac{۱۷۰}{۲۵}۱/۸۹\right) = ۰/۹۵$$

$$\Rightarrow p\left(۳/۵ \le \frac{\sigma_۱{}^۲}{\sigma_۲{}^۲} \le ۱۲/۸۵\right) = ۰/۹۵$$

بنابراین نسبت واریانسها با ۹۵ درصد اطمینان در بازه $(۱۲/۸۵\ ,\ ۳/۵)$ قرار دارد.

و در نتیجه:

$$p(۳/۵\ ,۱۲/۸۵\ \sigma_۲{}^۲) = ۰/۹۵$$

یعنی با ۹۵ درصد اطمینان می‌توان گفت که واریانس ماشین A سه برابر واریانس ماشین B می‌باشد بنابراین ماشین A بایستی تعمیر شود.

۵۹ ۱۰ . ۶ . ۷ فاصله اطمینان برای تفاضل میانگین‌های دو جامعه نرمال

از دو جامعه نرمال با واریانسهای ${\sigma_1}^2$ و ${\sigma_2}^2$ دو نمونه به حجم n_1 و n_2 می‌گیریم. قبلاً نشان دادیم که $Z=\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{{\sigma_1}^2}{n_1}+\dfrac{{\sigma_2}^2}{n_2}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد. بنابراین می‌توانیم یک فاصله اطمینان $(1-\alpha)\,100$ درصد برای تفاضل میانگین‌های دو جامعه نرمال با واریانس معلوم بصورت زیر بدست آوریم:

$$p\left(\overline{X}_1-\overline{X}_2-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}\ \le\ \mu_1-\mu_2\le\overline{X}_1-\overline{X}_2+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}\right)=1-\alpha$$

و در صورتی که واریانسهای دو جامعه نامعلوم اما مساوی باشند ابتدا واریانس مشترک دو نمونه را بصورت زیر برآورد می‌کنیم:

$${\sigma_1}^2={\sigma_1}^2=\sigma^2$$

$$\hat{\sigma}^2=S_p^2=\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2}$$

فاصله اطمینان در این حالت از رابطه زیر بدست می‌آید:

$$p\left(\overline{X}_1-\overline{X}_2-t_{1-\frac{\alpha}{2}}\,(n_1+n_2-2)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\ \le\mu_1-\mu_2\le\overline{X}_1-\overline{X}_2+t_{1-\frac{\alpha}{2}}\,(n_1+n_2-2)\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

۶۰ مثال ۲۵: میزان برق مصرفی دو شهر A و B یک متغیر تصادفی نرمال می‌باشد برای مقایسه متوسط مصرف یک نمونه ۱۵ تایی از دو شهر انتخاب می‌کنیم و نتایج زیر را بدست می‌آوریم:

$$\overline{X}_A=2590\qquad \sigma_A^2=150$$
$$\overline{X}_B=2560\qquad \sigma_B^2=100$$

مطلوبست:

الف) با فاصله اطمینان ۹۹ درصدی چه نظری می‌توان درباره تفاضل متوسط مصرف در دو شهر داشت؟

ب) اگر برای دو نمونه داشته باشیم $S_A^2=165\ ,\ S_B^2=120$ یک فاصله اطمینان ۹۹ درصدی برای حالتی که واریانسها نامعلوم باشند و مساوی بسازید؟

حل:الف)

$$1-\alpha=0/99\quad\rightarrow\quad\alpha=0/01\quad\rightarrow\quad Z_{1-\frac{\alpha}{2}}=Z_{0/99}=2/575$$

$$\rightarrow\ L_1=2590-2560-2/575\sqrt{\frac{150}{15}+\frac{100}{15}}=19/48$$

$$L_2=2590-2560-2/575\sqrt{\frac{150}{15}+\frac{100}{15}}=40/51$$

بنابراین با ۹۹ درصد اطمینان می‌توانیم بگوییم که $19/48\le\mu_1-\mu_2\le 40/51$ و چون $\mu_1-\mu_2$ در یک بازه مثبت قرار گرفته‌اند می‌توانیم با ۹۹ درصد اطمینان بگوییم که متوسط مصرف شهر A از شهر B بیشتر است.

۶۱ ب)

$$S_p^2=\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2}=\frac{14\,(120+165)}{15+15-2}=142/5\quad\Rightarrow\quad S_p=11/93$$

$$1-\alpha = ۰/۹۹ \quad \rightarrow \quad \alpha = ۰/۰۱ \quad \rightarrow \quad t_{۱-\frac{\alpha}{۲}}(n_۱+n_۲-۲) = t_{۰/۹۹۵}(۲۸) = ۲/۷۶$$

$$L_۱ = \overline{X}_۱ - \overline{X}_۲ - t_{۱-\frac{\alpha}{۲}}(n_۱+n_۲-۲)\, S_p \sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}} = ۲۵۹۰ - ۲۵۶۰ - ۲/۷۶\,(۱۱/۹۳)\sqrt{\frac{۱}{۱۵}+\frac{۱}{۱۵}} = ۱۷/۹۷$$

$$L_۲ = \overline{X}_۱ - \overline{X}_۲ - t_{۱-\frac{\alpha}{۲}}(n_۱+n_۲-۲)\, S_p \sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}} = ۲۵۹۰ - ۲۵۶۰ + ۲/۷۶\,(۱۱/۹۳)\sqrt{\frac{۱}{۱۵}+\frac{۱}{۱۵}} = ۴۲$$

ملاحظه می‌کنید که در حالت دوم که واریانس مشترک را بصورت تقریبی محاسبه نمودیم باز هم فاصله اطمینان مقداری نزدیک به حالت الف شده است.

فصل یازدهم

۱-۱۱ آزمون فرض

در فصل قبل با روشهای برآورد پارامترهای مجهول جامعه آشنا شدید، به عبارتی نشان دادیم که چگونه می‌توان با استفاده از نتایج حاصل از نمونه‌گیری پارامترهای مجهول یک جامعه آماری با توزیع معلوم را بدست آورد. حال در این فصل نشان می‌دهیم که چگونه می‌توان در مورد صحت یک فرضیه بر اساس یک جامعه آماری قضاوت نمود. به عنوان مثال فرض می‌کنیم که یک داروی بخصوص بر روی یک بیماری اثر مثبتی دارد به این ترتیب می‌بایستی برای اثبات این فرضیه، عده‌ای از بیماران را انتخاب کرده و دارو را روی آنها آزمایش می‌کنیم و با توجه به نتایج بدست آمده در مورد صحت یا عدم صحت فرضیه مورد نظر تصمیم می‌گیریم. روشی که در آن فرض مربوطه را پذیرفته یا رد می‌کنیم به آزمون فرض معروف می‌باشد، در این فصل نحوه آزمون فرضهای مختلف را معرفی می‌کنیم.

۲-۱۱

۱۱ . ۱ فرضها و آزمونها

تعریف دقیق فرض و آزمون فرض عبارتند از:

فرض: یک فرض عبارتست از گزاره‌ای درباره قانون احتمال یک متغیر تصادفی (که قابل نمونه‌گیری باشد)

آزمون فرض: نمونه‌گیری از متغیر تصادفی مربوطه و تصمیم‌گیری در مورد پذیرفتن یا رد فرض مورد نظر بر اساس نمونه بدست آمده.

بطور کلی نتیجه یک آزمون فرض یکی از چهار مورد زیر می‌باشد:

۱- فرض را رد می‌کنیم هنگامی که واقعاً صحیح باشد.

۲- فرض را رد می‌کنیم هنگامی که واقعاً صحیح نباشد.

۳- فرض را رد می‌پذیریم هنگامی که واقعاً صحیح باشد.

۴- فرض را رد می‌پذیریم هنگامی که واقعاً صحیح نباشد.

برای وشن شدن مطلب به مثال زیر توجه کنید:

۳-۱۱ مثال ۱: گروهی از پزشکان معتقدند یک دارو نیروزا می‌باشد. برای تصمیم‌گیری در این مورد آنرا روی ۲۰ نفر از ورزشکاران آزمایش می‌کنند. در صورتی که دارو در بیشتر از ۷۰ درصد مواقع اثر مثبتی بر عملکرد عادی ورزشکاران داشته باشد، پزشکان آنرا جزء داروهای نیروزا طبقه بندی می‌کنند. متغیر تصادفی X را برابر تعداد ورزشکارانی در بین ۲۰ نفر در نظر می‌گیریم که دارو تاثیر مثبتی بر عملکرد آنها داشته است. در این صورت X یک متغیر تصادفی دو جمله‌ای با پارامترهای $X \sim B(p\ ,\ ۲۰)$ می‌باشد. که در آن p درصد موثر بودن دارو می‌باشد و مقداری نامعلوم می‌باشد. نیروزا بودن دارو به این معنی است که $p > ۰/۷۰$ به عبارتی اگر از روی نمونه‌ها به این نتیجه برسیم که $p > ۰/۷۰$ می‌باشد در این صورت می‌توان گفت که دارو نیروزا بوده است. نتایج حاصل از نمونهٔا یکی از دو حالت (فرض) زیر را نتیجه می‌دهد:

۱- داروی مورد نظر نیروزا می‌باشد : $p > ۰/۷۰$

۲- داروی مورد نظر نیروزا نمی‌باشد : $p > ۰/۷۰$

۴-۱۱ در روش آزمون یکی از فرضها را فرض خنثی یا فرض صفر در نظر می‌گیریم و آنرا با $H_۰$ نمایش می‌دهیم به عبارتی $H_۰$ در این مثال همان فرض است که مورد ادعای پزشکان می‌باشد یعنی فرض داروی مورد نظر نیروزا می‌باشد $(p > ۰/۷۰)$ فرض $H_۰$ خواهد بود.

در مقابل هر فرضی که فرض $H_۰$ را رد کند فرض مخالف یا مکمل نامیده می‌شود و با $H_۱$ نمایش داده می‌شود. که در این مثال فرض داروی مورد نظر نیروزا نمی‌باشد $(p \leq ۰/۷۰)$ فرض $H_۱$ می‌باشد. بنابراین:

دارو نیروزاست $H_۰ = p > ۰/۷۰$

دارو نیروزا نیست $H_1 = p \le ۰/۷۰$

توجه کنید که فرض مخالف صورتهای متفاوتی می‌تواند داشته باشد مثلاً اگر فرض H_0 برابر $p = ۰/۷$ می‌بود در این صورت فرض H_1 هر یک از حالات زیر می‌توانست باشد:

$H_1 : p > ۰/۷$

$H_1 : p \ne ۰/۷$

$H_1 : p = ۰/۵$

تصمیم‌گیری در مورد اینکه کدام فرض را فرض مخالف در نظر بگیریم اهمیت زیادی دارد و معمولاً به صورت مساله بستگی دارد.

در ایم مثال پزشکان برای اثبات ادعای خود ۲۰ ورزشکار را انتخاب نموده و با توجه به متغیر تصادفی X این احتمال وجود دارد که نتایج طوری بدست آیند که پرشکان دارو را نیروزا معرفی کنند در حالی که واقعاً دارو نیروزا بوده است.

۵-۱۱ به عنوان مثالی دیگر فرض کنید یک سکه را ۲۰ مرتبه پرتاب کنیم و ۱۶ شیر مشاهده کنیم بوضوح مشاهدهٔ ۱۶ شیر تنها با ۲۰ مرتبه پرتاب سکه دلیلی بر ناسالم بودن سکه نمی‌باشد در مثال داروی نیروزا هم این مطلب صدق می‌کند. بنابراین در حالت کلی چهار حالت زیر بدست می‌آید:

۱- H_0 را می‌پذیریم هنگامی که واقعاً صحیح باشد.

۲- H_0 را رد می‌کنیم هنگامی که واقعاً صحیح باشد.

۳- H_0 را می‌پذیریم هنگامی که واقعاً صحیح نباشد.

۴- H_0 را رد می‌کنیم هنگامی که واقعاً صحیح نباشد.

مشاهده می‌کنید که حالت دوم و سوم نشان دهنده خطا در تصمیم‌گیری می‌باشند حالت دوم را خطای نوع اول (ریسک فروشنده می‌نامیم و حالت سوم را خطای نوع دوم (ریسک مشتری) می‌نامیم. خطای نوع اول را با α نمایش می‌دهند و آنرا سطح معنی‌دار یا سطح تشخیص آزمون می‌نامند و خطای نوع دوم را با B نمایش می‌دهند.

۶-۱۱ ۱۱ . ۲ ناحیه بحرانی و آماره آزمون

برای بیان ناحیه بحرانی و آماره آزمون مجدداً به مثال ۱ توجه کنید. در مثال ۱ نشان دادیم که:

$H_0 : p > ۰/۷۰$

$H_1 : p \le ۰/۷۰$

چگونگی انجام آزمون به این ترتیب است که متغیر تصادفی X را برابر تعداد ورزشکاران که دارو بروی عملکرد آنها تاثیر مثبتی داشته در نظر می‌گیریم در این صورت اگر مقدار مشاهده شده X مثلاً برابر ۱۵ باشد در نتیجه $p = \frac{۱۵}{۲۰} = \frac{۳}{۴} = ۰/۷۵$ و فرض H_0 را قبول می‌کنیم زیرا $p > ۰/۷۰$ اما

اگر مقدار مشاهده شده X مثلاً برابر ۵ باشد بوضوح فرض H_0 را رد می‌کنیم و فرض H_1 را می‌پذیریم زیرا $p = \frac{۵}{۲۰} = ۰/۲۵$ و $p \le ۰/۷۰$.

در حالت کلی زمانی فرض H_0 را می‌پذیریم که مقدار مشاهده شدهٔ X از عدد ۱۴ بیشتر باشد زیرا: $p = \frac{۱۴}{۲۰} = ۰/۷۰$

و اگر مقدار مشاهده شده X کوچکتر یا مساوی ۱۴ باشد فرض H_0 را رد می‌کنیم یعنی اگر مقادیر مشاهده شده X متعلق به مجموعه $C = \{x \mid x \le ۱۴\}$ باشد آنگاه H_0 را رد می‌کنیم و در غیر این صورت H_0 را می‌پذیریم. به آماره X که بر اساس مقادیر مشاهده شده آن H_0 را رد می‌کنیم یا می‌پذیریم آماره آزمون گویند و به ناحیه C که به ازای مقادیر آن H_0 رد می‌شود ناحیه بحرانی می‌گوییم. به تعریف زیر توجه کنید:

۱۱-۷ تعریف: به آماره $T = T(X_1, X_2, \cdots, X_n)$ که بر اساس مقادیر مشاهده شده آن یک فرض را می‌پذیریم یا رد می‌کنیم، آماره آزمون گویند. و به ازای مجموعه مقادیری از این آماره که بر اساس آنها فرض H_0 رد می‌شود ناحیه بحرانی آزمون می‌گوییم و با نماد C نمایش می‌دهیم و به متمم ناحیه بحرانی که با C' نمایش می‌دهیم ناحیه پذیرش آزمون می‌گوییم.

با توجه به تعریف در آزمون یک فرض به این صورت عمل می‌کنیم که ابتدامقادیر نمونه‌های $X_1, X_2, \cdots, X_n$ را جمع آوری می‌کنیم و بر اساس آنها آماره آزمون را که بشکل $t = T(x_1, x_2, \cdots, x_n)$ می‌باشد تشکیل می‌دهیم حال اگر $t \in C$ باشد فرض H_0 را رد می‌کنیم و در غیر این صورت آن را می‌پذیریم.

۸ . ۱۱ . ۳ خطای آزمون

اشاره کردیم که رد تصمیم‌گیری ممکن است دو نوع خطای نوع اول (α) و نوع دوم (β) رخ دهد، که در هر دو حالت نتیجه مطلوب نمی‌باشد و در عمل می‌خواهیم تا جای ممکن مقدار خطای α و β را کاهش دهیم. با استفاده از احتمالات شرطی خطای نوع اول و دوم بصورت زیر بدست می‌آیند:

$$\alpha = p(\text{رد شود } H_0 \text{ فرض} \mid \text{درست باشد } H_0) = p(T(X_1, \cdots X_n) \in C \mid \text{درست باشد } H_0)$$

$$\beta = p(\text{پذیرفته شود } H_0 \text{ فرض} \mid \text{درست باشد } H_1) = p(T(X_1, \cdots X_n) \notin C \mid \text{درست باشد } H_1)$$

اگر در آزمون فرض به این نتیجه برسیم که H_0 باید رد شود و در حالی که H_1 واقعاً درست باشد یعنی مقدار احتمال زیر را داشته باشیم:

$$p(\text{رد شود } H_0 \mid \text{درست باشد } H_1)$$
$$= p(\text{رد شود } H_0 \mid \text{واقعاً نادرست باشد } H_0)$$

آشکار است که هر چه مقدار این احتمال بیشتر باشد آزمون نتیجه دقیق‌تری بدست داده است به همین دلیل به این احتمال توان آزمون گفته می‌شود. و با علامت β^* نمایش داده می‌شود.

توجه کنید که مقدار β^* برابر با $1-\beta$ می‌باشد:

$$\beta^* = p(\text{رد شود } H_0 \mid \text{درست باشد } H_1) = 1 - p(\text{پذیرفته شود } H_0 \mid \text{درست باشد } H_1)$$
$$= 1 - \beta$$

۹ مثال ۲: شخصی ادعا می‌کند که سکه استفاده شده برای تعین زمین بازی در یک مسابقه فوتبال یکنواخت نبوده و با احتمال ۰/۶ شیر می‌آمده است. برای تحقیق صحت ادعای وی سکه را ۲۰ مرتبه پرتاب می‌کنیم مطلوبست:

الف) تعین ناحیه بحرانی آزمون و فرضهای آزمون.

ب) محاسبه مقدار احتمال خطای نوع اول و خطای نوع دوم.

ج) محاسبه توان آزمون.

حل: الف) X را تعداد شیرهای بدست آمده در ۲۰ مرتبه پرتاب سکه در نظر می‌گیریم در این صورت X یک متغیر تصادفی دو جمله‌ای با پارامتر p و ۲۰ می‌باشد. $X \sim \beta(20, p)$.

طبق تعریف فرض صفر خلاف ادعای مطرح شده می‌باشد بنابراین:

$$\begin{cases} H_0 : \ p = \frac{1}{2} \\ H_1 : \ p = \frac{6}{10} \end{cases}$$

با توجه به اینکه در متغیر تصادفی برنولی $\mu = np$ می‌باشد بنابراین اگر در ۲۰ مرتبه پرتاب سکه حداقل تعداد $12 = 20 \times \frac{6}{10}$ شیر مشاهده کنیم می‌بایستی فرض H_0 را رد کنیم به این ترتیب ناحیه بحرانی عبارتست از:

$$C = \{x \mid x \geq 12\}$$

ب) محاسبه خطای نوع اول:

$$\alpha = p\,(X \in C \mid \text{درست باشد}) = p(X \geq ۱۲ \mid p = \frac{۱}{۲})$$

$$۱-p\,(X \leq ۱۱ \mid p = \frac{۱}{۲}) = ۱-۰/۷۴۸۳ = ۰/۲۵۱۷$$

محاسبه خطای نوع دوم:

$$\beta = p\,(X \notin C \mid \text{درست } H_۱) = p(X < ۱۲ \mid p=۰/۶)$$
$$= p(X \leq ۱۱ \mid p=۰/۶) = ۰/۴۰۴۴$$

ج) محاسبه توان آزمون:

$$\beta^* = ۱-\beta = ۰/۵۹۵۶$$

ملاحظه می‌کنید که احتمال خطای نوع اول کم و احتمال خطای نوع دوم زیاد می‌باشد.

۱۰-۱۱ مثال ۳: در مثال قبل اگر ناحیه بحرانی بصورت $C = \{x \mid x \geq ۱۳\}$ باشد احتمال خطای نوع اول و دوم و توان آزمون را محاسبه کنید؟

حل:

$$\alpha = p(X \geq ۱۳ \mid p = \frac{۱}{۲}) = ۱-p\,(X \leq ۱۲ \mid p = \frac{۱}{۲}) = ۱-۰/۸۶۸۴ = ۰/۱۳۱۶$$
$$\beta = p(X < ۱۳ \mid p = ۰/۶) = p\,(X \leq ۱۲ \mid p = ۰/۶) = ۰/۵۸۴۱$$
$$\beta^* = ۱-\beta = ۰/۴۱۵۹$$

ملاحظه می‌کنید که در حالت دوم با متغیر ناحیه بحرانی مقدار احتمال نوع اول کاهش یافت اما این مساله موجب افزایش احتمال خطای نوع دوم شده است.

بنابراین می‌بایستی ناحیه بحرانی طوری انتخاب شود که همزمان با در نظر گرفتن یک مقدار حداکثر برای خطای نوع اول باعث حداقل نمودن نوع دوم شود. به این ترتیب توان آزمون نیز حداکثر می‌شود.

در مثال بعد با قرار دادن یک مقدار معین برای احتمال خطای نوع اول (α) تلاش می‌نیم احتمال خطای نوع دوم را تا جای ممکن کاهش دهیم و در نتیجه توان آزمون را افزایش دهیم.

۱-۱۱-۱۱ مثال ۴: X یک متغیر تصادفی نرمال با میانگین مجهول μ و واریانس ۹ می‌باشد. آزمون فرض زیر را در نظر بگیرید:

$$\begin{cases} H_۰ : \ \mu = ۰ \\ H_۱ : \ \mu = ۱ \end{cases}$$

اگر از X یک نمونه ۲۵ تایی بگیریم و ناحیه بحرانی بصورت $C = \{(X_۱,\cdots,X_{۲۰}) \mid \overline{X} \geq C\}$ باشد مقدار C را طوری بدست بیاورید که $\alpha = ۰/۱$ باشد و سپس احتمال خطای نوع دوم و توان آزمون را محاسبه کنید؟

حل: ابتدا توجه کنید که $\overline{X}$ دارای میانگین μ و واریانس $\frac{\sigma^۲}{n}$ می‌باشد. یعنی $\overline{X} \sim (\mu, \frac{۹}{۲۵})$ به این ترتیب:

$$\alpha = ۰/۱ = p(\overline{X} > C \mid \mu = ۰) = p\,(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{C-\mu}{\frac{\sigma}{\sqrt{n}}} \mid \mu = ۰)$$

$$= p(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{C-۰}{\frac{۳}{۵}}) = p\,(Z > ۱/۶۶\,C) = ۰/۱$$

$$\Rightarrow ۱-p\,(Z \leq ۱/۶۶\,C) = ۰/۱ \quad \Rightarrow p(Z \leq ۱/۶۶\,C) = ۰/۹$$

رابطه اخیر معادل است با اینکه $Z_{۰/۹} = ۱/۶۶\,C$ از جدول مقدار $Z_{۰/۹}$ برابر است با ۱/۲۸ بنابراین:

$$Z_{0/9} = 1/28 = 1/66\,C \quad \Rightarrow \quad C = 0/77$$

به این ترتیب ناحیه بحرانی بصورت زیر بدست می‌آید:

$$C = \left\{(X_1,\cdots,X_{25}) \mid \overline{X} > 0/77\right\}$$

ملاحظه می‌کنید مه در این مثال برای اینکه بتوانیم مقدار خطای نوع اول را به دلخواه کاهش دهیم به ناچار می‌بایستی بازه ناحیه بحرانی را متغیر فرض می‌کردیم، حال مقدار خطای نوع دوم را بدست می‌آوریم:

$$\beta = p\,(\overline{X} \le C \mid \mu = 1) = p\,\left(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{0/77-\mu}{\frac{\sigma}{\sqrt{n}}} \mid \mu = 1\right)$$

$$= p\left(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{0/77-1}{\frac{3}{5}}\right) = p\,(Z \le -0/38) = N_Z(-0/38) = 0/3520$$

توان آزمون:

$$\beta^* = 1 - \beta = 1 - 0/3520 = 0/648$$

۱۲-۱۱ ۱۱ . ۴ انواع فرضها

فرض $H_0 : \mu = 2$ را در نظر بگیرید، این فرض با بیان اینکه میانگین جامعه برابر ۲ می‌باشد توزیع جامعه را معلوم می‌کند و نشان می‌دهد که پارامتر مجهول جامعه میانگین برابر عدد ۲ می‌باشد، به این نوع فرضها که توزیع جامعه را کاملاً مشخص می‌سازند فرضهای ساده می‌کوییم.
حال فرض $H_1 : \mu \ge 2$ را در نظر بگیرید، واضح است که اگر این فرض درست باشد با بیان یک بازه برای پارامتر مجهول جانعه، توزیع را بصورت دقیق مشخص نمی‌کند به این فرضها که با درست بودنشان پارامتر مجهول جامعه و در نتیجه توزیع جامعه بصورت دقیق مشخص نمی‌شود فرضهای مرکب می‌گوییم.
به عنوان مثال در مثال ۴ فرضهای $H_0 : \mu = 0$ و $H_1 : \mu = 1$ فرضهای ساده می‌باشند اما در مثال ۱ فرضهای $H_0 : p \ge 0/6$ و $H_1 : p < 0/6$ فرضهای مرکب می‌باشند.

۱۳-۱۱ ۱۱ . ۵ انواع آزمونها

اگر θ یک پارامتر نامعلوم جامعه باشد و بخواهیم آزمونهایی در مورد این پارامتر انجام دهیم، آزمون هر فرض آماری که در آن فرضیه مقابل (H_1) یک طرفه باشد آزمون یک طرفه نامیده می‌شود. به عنوان مثال برای مقدار ثابت θ_0 از θ آزمونهای زیر همگی یک طرفه می‌باشند:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \ge \theta_0 \end{cases} \qquad \begin{cases} H_0 : \theta \ge \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$$

اگر در یک آزمون فرضیه مقابل (H_1) دو طرفه باشد آن آزمون را آزمون دو طرفه می‌نامیم. مانند آزمون زیر:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \ne \theta_0 \end{cases}$$

۱۴-۱۱ ۱۱ . ۶ مراحل انجام یک آزمون فرض

برای انجام یک آزمون فرض می‌بایستی مراحل زیر را بصورت گام به گام طی نمود:

۱- تعین فرضهای صفر H_0 و فرض مقابل H_1.

۲- تعین یک طسح معنی‌دار α که معمولاً یکی از اعداد ۰/۰۱ ، ۰/۰۵ یا ۰/۱ در نظر گرفته می‌شود.

۳- تعین آماره آزمون $T = T(X_1,\cdots,X_n)$ که عموماً بر اساس برآوردگر نقطه‌ای پارامتر مجهول θ بدست می‌آید.

۴- تعین ناحیه بحرانی آزمون که از روی آماره آزمون، فرض مقابل آزمون و سطح معنی‌دار α بدست می‌آید.

۵- محاسبه مقدار آماره آزمون که از روی نمونه‌های تصادفی $X_1, X_2, \cdots, X_n$ بدست می‌آید.

۶- نتیجه‌گیری – اگر مقدار محاسبه شده آماره آزمون درون ناحیه بحران باشد فرض H_0 را رد و در غیر این صورت فرض H_0 را می‌پذیریم.
توجه کنید که عموماً در حل مسایل آزمون فرض، با در نظر گرفتن انواع آزمون فرضها و تعین آماره آزمون و ناحیه بحرانی برای هر یک از آزمون فرض‌ها محاسبه مراحل ۳ و ۴ بسیار ساده‌تر می‌شود.

۱۵-۱۱ ۱۱ . ۷ آزمون فرضهای آماری برای پارامترهای جامعه

برای سادگی انجام آزمونهای فرض بهترین راه محاسبه آماره آزمون و ناحیه بحرانی برای فرضهای متفاوت می‌باشد. بنابراین در این بخش ابتدا آزمون فرض را روی میانگین یک جمعیت را زمانی که واریانس معلوم می‌باشد انجام می‌دهیم:

فرض می‌کنیم X یک متغیر تصادفی با میانگین مجهول μ و واریانس معلوم σ^2 باشد. نمونه‌های تصادفی $X_1, X_2, \cdots, X_n$ را انتخاب می‌کنیم، می‌خواهیم آزمونهایی را روی میانگین μ انجام دهیم. در این صورت سه حالت کلی زیر را در نظر می‌گیریم:

۱- آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ که در آن H_0 رد می‌شود اگر $\overline{X} > C$

۲- آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ که در آن H_0 رد می‌شود اگر $\overline{X} > C$

۳- آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$ که در آن H_0 رد می‌شود اگر $\overline{X} < C_1$ یا $\overline{X} > C_2$

توجه کنید که در هر حالت مقدار C که بیانگر ناحیه بحرانی می‌باشد با توجه به سطح معنی‌دار α بدست می‌آید. حال برای هر حالت آماره آزمون و مقدار C را محاسبه می‌کنیم:

۱۶-۱۱ حالت اول: H_0 رد می‌شود اگر $\overline{X} > C$ بنابراین خطای نوع اول عبارتست از: $\alpha = p(\overline{X} > C \mid \mu = \mu_0)$

اگر X یک متغظر تصادفی نرمال باشد یا $n \geq 30$ باشد قبلاً نشان دادیم که $Z = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ یک متغیر تصادفی نرمال استاندارد می‌باشد بنابراین:

$$Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$\alpha = p(\overline{X} > C \mid \mu = \mu_0) = p\left(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = p\left(Z > \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$\Rightarrow 1 - p\left(Z > \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = \alpha \quad \Rightarrow \quad p\left(Z > \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = 1 - \alpha$$

$$\Rightarrow \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}} = Z_{1-\alpha} \quad \Rightarrow \quad C = \frac{\sigma}{\sqrt{n}}(Z_{1-\alpha}) + \mu_0$$

حال با توجه به مقدار C ناحیه بحرانی ازمون بصورت زیر بدست می‌آید:

$$\overline{X} > C \quad \rightarrow \quad \overline{X} > \mu_0 + \frac{\sigma}{\sqrt{n}} Z_{1-\alpha} \quad \text{یا} \quad \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > Z_{1-\alpha}$$

نتیجه می‌گیریم که:

در آزمون $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$Z=\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} > Z_{1-\alpha}$$

توجه کنید که اگر فرض H_0 بصورت $H_0 : \mu \leq \mu_0$ باشد می‌توان نشان داد که ناحیه بحرانی باز هم به صورت رابطه بالا می‌باشد.

۱۱-۱۷ حالت دوم: آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ H_0 را رد می‌کنیم اگر $\overline{X} < C$ مشابه حالت قبل اگر C را محاسبه کنیم مقدار ناحیه بحرانی بصورت زیر بدست می‌آید:

$$C = \mu_0 - Z_{1-\alpha}\frac{\sigma}{\sqrt{n}} \quad \Rightarrow \quad \overline{X} < C \;\rightarrow\; \overline{X} < \mu_0 - \frac{\sigma}{\sqrt{n}} Z_{1-\alpha}$$

$$Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < -Z_{1-\alpha}$$

بنابراین در حالت کلی در آزمون $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < -Z_{1-\alpha}$ در این حالت هم اگر فرض H_0 بصورت $H_0 : \mu \geq \mu_0$ می‌بود باز هم ناحیه بحرانی بصورت رابطه فوق بدست می‌آمد.

۱۱-۱۸ حالت سوم: آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$ که در آن H_0 رد می‌شود اگر $\overline{X} > C_2$ یا $\overline{X} < C_1$ باشد بنابراین خطای نوع اول عبارتست از:

$$\alpha = p(\overline{X} < C_1 \text{ یا } \overline{X} > C_2 \mid \mu = \mu_0) \;\rightarrow\; 1-\alpha = p(C_1 < \overline{x} < C_2 \mid \mu = \mu_0)$$

$$p = \left(\frac{C_1-\mu_0}{\frac{\sigma}{\sqrt{n}}} < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{C_2-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= p\left(\frac{C_1-\mu_0}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{C_2-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = 1-\alpha$$

که در آن:

$$\frac{C_2-\mu_0}{\frac{\sigma}{\sqrt{n}}} = -\frac{C_1-\mu_0}{\frac{\sigma}{\sqrt{n}}} < Z_{1-\alpha}$$

و در نتیجه بدست می‌آوریم:

$$C_1 = \mu_0 - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$C_2 = \mu_0 + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

و ناحیه بحرانی به صورت زیر بدست می‌آید:

$$\overline{X} > \mu_0 + Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \qquad \text{یا} \qquad \overline{X} < \mu_0 - Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

که معادل است با:

$$\left|\frac{X'-\mu_{\circ}}{\frac{\sigma}{\sqrt{n}}}\right| > Z_{1-\frac{\alpha}{2}}$$

بنابراین در حالت کلی در آزمون فرض $\begin{cases} H_{\circ} : \mu = \mu_{\circ} \\ H_{1} : \mu \neq \mu_{\circ} \end{cases}$ فرض $H_{\circ}$ رد می‌شود اگر وفقط اگر:

$$|Z| = \left|\frac{X'-\mu_{\circ}}{\frac{\sigma}{\sqrt{n}}}\right| > Z_{1-\frac{\alpha}{2}}$$

در مثالهای بعدی هر حالت را مورد بررسی قرار می‌دهیم.

۱۱-۱۹ مثال ۵: توان شکنندگی (مقدار نیروی لازم برای شکستن) کابلهایی که توسط یک شرکت تولید می‌شوند دارای میانگین ۱۸۰۰ پوند و انحراف معیار ۱۰۰ می‌باشند. محققان شرکت با اعمال تکنیک جدیدی که در مراحل ساخت اعمال کرده‌اند ادعا کرده‌اند که توان شکنندگی افزایش یافته است. برای آزمودن این ادعا یک نمونه ۵۰ تایی از کابلها تحت آزمون قرار می‌گیرند و میانگین توان شکنندگی ۱۸۵۰ پوند بدست می‌آید. آیا در سطح معنی‌دار ۰/۰۱ این ادعا پذیرفته است؟

حل: ابتدا هر یک ار فرضهای صفر و مقابل را تعریف می‌کنیم:

$H_{\circ}:$ $\mu = ۱۸۰۰$ هیچ تغییری در توان شکنندگی رخ نداده است.

$H_{1}:$ $\mu > ۱۸۰۰$ توان شکنندگی افزایش یافته است.

ملاحظه می‌کنید که حالت اول رخ داده است و با توجه به مطالب ارایه شده آماره آزمون و ناحیه بحرانی بصورت زیر خواهند بود:

$$Z = \frac{X'-\mu_{\circ}}{\frac{\sigma}{\sqrt{n}}} > Z_{1-\frac{\alpha}{2}}$$

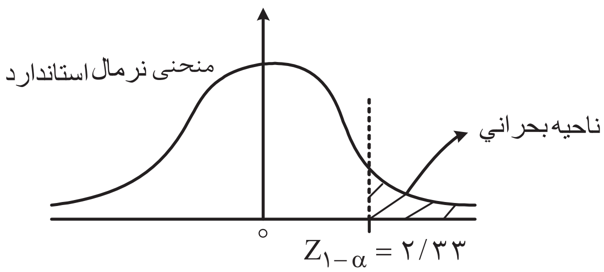
که در آن:

$$\alpha = ۰/۰۱$$
$$\Rightarrow 1-\alpha = 1-۰/۰۱ = ۰/۹۹ \quad \Rightarrow \quad Z_{1-\alpha} = Z_{۰/۹۹} = ۲/۳۳$$

یعنی ناحیه بحرانی بصورت $Z > ۲/۳۳$ می‌باشد و اگر مقدار آماره Z بزرگتر از ۲/۳۳ باشد فرض $H_{\circ}$ را رد می‌کنیم و به عبارتی ادعای محققان را می‌پذیریم. مقدار آماره Z برابر است با:

$$Z = \frac{X'-\mu_{\circ}}{\frac{\sigma}{\sqrt{n}}} >= \frac{۱۸۵۰ - ۱۸۰۰}{\frac{۱۰۰}{\sqrt{۵۰}}} = ۳/۵۵$$

بنابراین در سطح معنی‌دار $\alpha = ۰/۰۱$ نتایج نشان می‌دهند که $Z > ۲/۳۳$ می‌باشد و در نتیجه ادعای محققان را در مورد افزایش توان شکنندگی کابلها می‌پذیریم. در نمودار زیر ناحیه رد فرض $H_{\circ}$ را ملاحظه می‌کنید:



۲۰–۱۱ مثال ۶: عمر متوسط ۱۰۰ عدد از لامپهای مهتابی تولید شده توسط یک کارخانه برابر ۱۵۷۰ ساعت با انحراف معیار ۱۲۰ ساعت بدست آمده است. کارخانه تولید کننده ادعا می‌کند که عمر متوسط لامپها برابر ۱۶۰۰ ساعت می‌باشد در حالی که مصرف کنندگان این ادعا را قبول ندارند. صحت ادعای کارخانه سازنده را در سطح $\alpha = ۰/۰۵$ و $\alpha = ۰/۰۱$ بررسی کنید؟

حل: در این حالت دو فرض زیر را پیش رد داریم:

$$\begin{cases} H_{\circ} : \mu = ۱۶۰۰ \\ H_{۱} : \mu \neq ۱۶۰۰ \end{cases}$$

ملاحظه می‌کنید که در این حالت یک آزمون دو طرفه انجام می‌گیرد که در آن آماره آزمون و ناحیه بحرانی بصورت زیر می‌باشد:

$$|Z| = \left| \frac{X' - \mu_{\circ}}{\frac{\sigma}{\sqrt{n}}} \right| > Z_{۱-\frac{\alpha}{۲}}$$

الف) ابتدا آزمون فرض را در سطح معنی‌دار $\alpha = ۰/۰۵$ انجام می‌دهیم:

$$\alpha = ۰/۰۵ \quad \rightarrow \quad ۱-\frac{\alpha}{۲} = ۰/۹۷۵$$
$$\rightarrow \quad Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۷۵} = ۱/۹۶$$

بنابراین فرض $H_{\circ}$ را رد می‌کنیم اگر $|Z| > ۱/۹۶$ باشد یا به عبارتی:

$$|Z| > ۱/۹۶ \quad \text{یا} \quad Z < -۱/۹۶$$

۲۱–۱۱ حال مقدار آماره آزمون را بدست می‌آوریم:

ابتدا توجه کنید که مقدار واقعی انحراف معیار طول عمر لامپها را نداریم بنابراین از واریانس یا انحراف معیار $\overline{X}$ برای تخمین واریانس واقعی طول عمر لامپها استفاده می‌کنیم:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{۱۲۰}{\sqrt{۱۰۰}} = ۱۲$$

به این ترتیب آماره آزمون بصورت زیر بدست می‌آید:

$$Z = \frac{X' - \mu_{\circ}}{\frac{\sigma}{\sqrt{n}}} > \frac{۱۵۷۰ - ۱۶۰۰}{۱۲} = -۲/۵$$

ملاحظه می‌کنید که ۲/۵- خارج از بازه $(-۱/۹۶ \,,\, ۱/۹۶)$ قرار دارد بنابراین $H_{\circ}$ را در سطح معنی‌دار ۰/۰۵ رد می‌کنیم یعنی ادعای مصرف کنندگان در مورد عدم صحت میانگین طول عمر مطرح شده توسط کارخانه سازنده، صحیح می‌باشد.

ب) حال آزمون فرض را در سطح معنی‌دار ۰/۰۱ انجام می‌دهیم:

$$\alpha = ۰/۰۱ \quad \rightarrow \quad ۱-\frac{\alpha}{۲} = ۰/۹۹۵$$
$$\rightarrow \quad Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۹۵} = ۲/۵۸$$

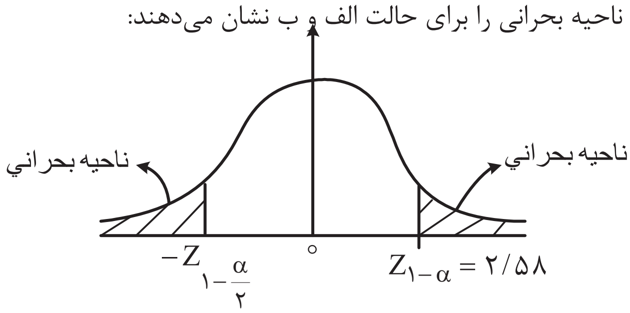
در این حالت اگر آماره آزمون در خارج از بازه $(-۲/۵۸ \,,\, ۲/۵۸)$ قرار داشته باشد $H_{\circ}$ را رد می‌کنیم. مقدار آماره آزمون در این حالت برابر همان مقدار بدست آمده در بند الف مثال می‌باشد:

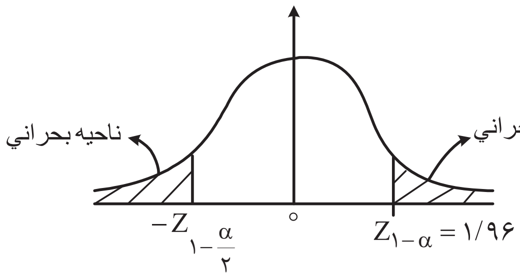
$$Z = -۲/۵ \quad \Rightarrow \quad |Z| > ۲/۵۸$$

۲۲–۱۱ از آنجا که مقدار Z برابر ۲/۵ می‌باشد و در بازه $(-۲/۵۸ \,,\, ۲/۵۸)$ قرار دارد بنابراین در سطح معنی‌دار $\alpha = ۰/۰۱$ و فرض $H_{\circ}$ را می‌پذیریم.

ملاحظه می‌کنید که با تغییر سطح معنی‌داری و به دنبال آن تغییر ناحیه بحرانی، این احتمال وجود دارد که تصمیم‌گیری در مورد رد یا پذیرش فرض H_0 کاملاً عوض شود. توجه کنید که در حالت دوم $(\alpha = 0/01)$ فرض H_0 را می‌پذیریم اما این به معنی رد فرض H_1 نمی‌باشد بلکه در این حالت می‌گوییم نمی‌توان در مورد رد فرض H_1 نظری داد یا به بعارت دیگر هیچ تصمیمی نمی‌گیریم.

دو نمودار زیر ناحیه بحرانی را برای حالت الف و ب نشان می‌دهند:





۱۱-۲۳ آزمون فرض برای میانگین نرمال با واریانس نامعلوم

در صورتی که واریانس جامعه نامعلوم باشد نشان دادیم که آماره $T = \dfrac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}$ دارای توزیع t با $n-1$ درجه آزادی است، بنابراین در این حالت و با توجه به روشهای ارایه شده برای بدست آوردن ناحیه بحرانی در بخش قبل، می‌توان نشان داد که آماره آزمون و ناحیه بحرانی برای هرحالت بصورت زیر بدست می‌آیند:

۱- در آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \text{ یا } \mu \le \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$ H_0 رد می‌شود اگر و فقط اگر:

$$T = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} > t_{1-\alpha}(n-1)$$

۲- در آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \text{ یا } \mu \ge \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$ H_0 رد می‌وشد اگر و فقط اگر:

$$T = \frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}} < -t_{1-\alpha}(n-1)$$

۳- در آزمون فرض $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \ne \mu_0 \end{cases}$ H_0 رد می‌شود اگر و فقط اگر:

$$|T| = \left|\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}\right| > t_{1-\frac{\alpha}{2}}(n-1)$$

۱۱-۲۴ مثال ۷: تعداد نامه‌های رسیده به یک شرکت در طول ۱۰ روز عبارتست از:

۱ , ۳ , ۲ , ۵ , ۶ , ۴ , ۵ , ۸ , ۹ , ۱

اگر توزیع تعداد نامه‌های رسیده از متغیر تصادفی نرمال پیروی کند، آیا در سطح معنی‌دار $\alpha = 0/01$ می‌توان ادعا نمود که بطور متوسط روزانه حداقل ۴ نامه به شرکت پست ارسال می‌شود؟

حل: فرضهای زیر را در نظر می‌گیریم:

$$\begin{cases} H_0 : \mu = 4 \\ H_1 : \mu > 4 \end{cases}$$

واریانس جامعه نامعلوم است بنابراین برای محاسبه از آماره $T = \dfrac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}$ استفاده می‌کنیم که در آن $\overline{X}$ و S برابر هستند با:

$$\overline{X} = \frac{۱}{۱۰}(۱+۳+۲+۵+۶+۴+۵+۸+۹+۱) = ۴/۴$$

$$S^۲ = \frac{۱}{n-۱}\left(\sum_{i=۱}^{n} X_i{}^۲ - \frac{\overline{X}^۲}{n}\right) = \frac{۱}{۹}\left\{(۱+۹+۴+۲۵+۳۶+۱۶+۲۵+۶۴+۸۱+۱) - \frac{(۴/۴)^۲}{۱۰}\right\}$$

$$= \frac{۱}{۹}\left(۲۶۲ - \frac{(۴/۴)^۲}{۱۰}\right) = ۲۸/۸ \quad \Rightarrow \quad S = \sqrt{۲۹} = ۵/۳۶$$

فرض $H_\circ$ در صورتی رد می‌شود که:

$$T > t_{۱-\alpha}(n-۱)$$

$$\alpha = ۰/۰۱ \quad \rightarrow \quad ۱-\alpha = ۰/۹۹$$

$$t_{۱-\alpha} = t_{۰/۹۹}(n-۱) = t_{۰/۹۹}(۹) = ۲/۸۲$$

مقدار آماره آزمون عبارتست از:

$$T = \frac{\overline{X} - \mu_\circ}{\frac{S}{\sqrt{n}}} = \frac{۴/۴ - ۴}{\frac{۵/۳۶}{\sqrt{۱۰}}} = ۰/۲۳$$

از آنجا که مقدار آماره آزمون کمتر از ۲/۸۲ می‌باشد بنابراین فرض $H_\circ$ را می‌پذیریم.

۱۱-۲۵ ۱۱ . ۹ آزمون فرض برای واریانس یک جامعه نرمال

در صورتی که بخواهیم ادعاهایی را در مورد واریانس یک جامعه نرمال بررسی کنیم می‌دانیم که آماره $\chi^۲ = \frac{(n-۱)S^۲}{\sigma^۲}$ دارای توزیع خی دو با $n-۱$ درجه آزادی است بنابراین می‌توان حالتهای زیر را برای هر حالت بدست آورد:

۱- در آزمون $\begin{cases} H_\circ : \sigma^۲ = \sigma_\circ^۲ \text{ یا } \sigma^۲ \le \sigma_\circ^۲ \\ H_۱ : \sigma^۲ > \sigma_\circ^۲ \end{cases}$ $H_\circ$ رد می‌شود اگر و فقط اگر: $\chi^۲ = \frac{(n-۱)S^۲}{\sigma_\circ^۲} > \chi^۲_{۱-\alpha}(n-۱)$

۲- در آزمون فرض $\begin{cases} H_\circ : \sigma^۲ = \sigma_\circ^۲ \text{ یا } \sigma^۲ \ge \sigma_\circ^۲ \\ H_۱ : \sigma^۲ < \sigma_\circ^۲ \end{cases}$ $H_\circ$ رد می‌شود اگر و فقط اگر: $\chi^۲ = \frac{(n-۱)S^۲}{\sigma_\circ^۲} < \chi^۲_{\alpha}(n-۱)$

۳- در آزمون فرض $\begin{cases} H_\circ : \sigma^۲ = \sigma_\circ^۲ \\ H_۱ : \sigma^۲ \ne \sigma_\circ^۲ \end{cases}$ $H_\circ$ رد می‌شود اگر و فقط اگر:

$$\chi^۲ = \frac{(n-۱)S^۲}{\sigma_\circ^۲} > \chi^۲_{۱-\frac{\alpha}{۲}}(n-۱) \quad \text{یا} \quad \chi^۲ = \frac{(n-۱)S^۲}{\sigma_\circ^۲} < \chi^۲_{\frac{\alpha}{۲}}(n-۱)$$

۱۱-۲۶ مثال ۹: یک کارخانه تولید نوشابه ماشین آلات جدیدی برای پرنمودن شیشه‌ها خریداری نموده است که ادعا شده است این ماشین‌ها با انحراف معیار ۱۰ میلی لیتر شیشه‌ها را پر می‌کنند. برای بررسی صحت ادعا یک نمونه تصادفی ۱۰ تایی از شیشه‌های پر شده را انتخاب می‌کنیم و نتایج $\overline{X} = ۲۵۵$ و $S^۲ = ۱۸۰$ بدست امده است. در سطح معنی‌دار $\alpha = ۰/۰۵$ آیا ادعای مطرح شده صحیح است؟

حل: فرضها عبارتند از:

$$\begin{cases} H_0 : \sigma^2 = ۱۰۰ \\ H_1 : \sigma^2 \neq ۱۰۰ \end{cases}$$

با توجه به آزمونها آماره آزمون و ناحیه بحرانی بصورت زیر خواهند بود:

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) < \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{یا} \quad \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{\frac{\alpha}{2}}(n-1)$$

در سطح معنی‌دار ۰/۰۵ داریم:

$$\alpha = ۰/۰۵ \quad \rightarrow \quad 1-\frac{\alpha}{2} = ۰/۹۷۵ \quad \frac{\alpha}{2} = ۰/۰۲۵$$

پس:

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{۰/۹۷۵}(۹) = ۱۹$$

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{۰/۰۲۵}(۹) = ۲/۷$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{۹(۱۸۰)}{۱۰۰} = ۱۶/۲$$

از آنجا که آماره $\chi^2 = ۱۶/۲$ درون بازه (۱۹ , ۲/۷) قرار دارد بنابراین ادعای H_0 رد نمی‌شود و در نتیجه ادعا مطرح شده در سطح معنی‌دار ۰/۰۵ قابل قبول است.

۲۷-۱۱ ۱۱ . ۹ آزمون فرض برای تفاضل میانگین‌ها

یک نمونه تصادفی n_1 تایی از جامعه نرمال با میانگین μ_1 و واریانس σ_1^2 انتخاب می‌کنیم، همچنین یک نمونه n_2 تایی از جامعه نرمال دیگری با میانگین μ_2 و واریانس σ^2 انتخاب می‌کنیم بطوریکه دو نمونه از یکدیگر مستقل باشند در این حالت برای آزمون تفاضل میانگین‌های دو جامعه از آزمون‌های زیر می‌توانیم استفاده کنیم که در سه حالت اول واریانس دو جامعه معلوم و در سه حالت بعدی واریانس مجهول فرض شده است:

۱- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \text{ یا } \mu_1 - \mu_2 \le d_0 \\ H_1 : \mu_1 - \mu_2 > d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > Z_{1-\alpha}$$

۲- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \text{ یا } \mu_1 - \mu_2 \ge d_0 \\ H_1 : \mu_1 - \mu_2 < d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < Z_{1-\alpha}$$

۳- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 \neq d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$|Z| = \left| \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right| > Z_{1-\alpha}$$

۲۸-۱۱ در حالتی که واریانسهای دو جامعه نامعلوم اما برابر باشند $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$ روابط آزمونهای فرض بصورت زیر خواهند بود:

۱- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \text{ یا } \mu_1 - \mu_2 \le d_0 \\ H_1 : \mu_1 - \mu_2 > d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{1-\alpha}(n_1 + n_2 - 2)$$

که در آن $S_P^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$ می‌باشد.

۲- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \text{ یا } \mu_1 - \mu_2 \ge d_0 \\ H_1 : \mu_1 - \mu_2 < d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{1-\alpha}(n_1 + n_2 - 2)$$

۳- در آزمون فرض $\begin{cases} H_0 : \mu_1 - \mu_2 = d_0 \\ H_1 : \mu_1 - \mu_2 \ne d_0 \end{cases}$ فرض H_0 رد می‌شود اگر و فقط اگر:

$$|T| = \left| \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| > t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$$

۲۹–۱۱ مثال ۱۰: یک امتحان از دو کلاس A و B گرفته شده است. کلاس A شامل ۴۰ دانشجو و کلاس B شامل ۵۰ دانشجو می‌باشد. میانگین نمرات در کلاس A ۷۴ با انحراف معیار ۸ و در کلاس B ۷۸ با انحراف معیار ۷ می‌باشد آیا نمرات دانشجویان این دو کلاس متفاوت معنی‌داری در سطوح ۰/۰۵ و ۰/۰۱ با هم دارند؟

حل: هر یک از دو کلاس را دو جامعه با میانگین‌های μ_1 و μ_2 در نظر می‌گیریم داریم:

$\overline{X}_1 = 74 \qquad \sigma_1 = 8$

$\overline{X}_2 = 78 \qquad \sigma_2 = 7$

$n_1 = 40$

$n_2 = 50$

فرض‌های آزمون عبارتند از:

$H_0 : \mu_1 = \mu_2$ اختلاف تنها ناشی از شانس است.

$H_1 : \mu_1 \ne \mu_2$ بین دو کلاس اختلاف معنی‌داری وجود دارد.

در این حالت آماره آزمون و ناحیه بحرانی عبارتند از:

$$|Z| = \left| \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right| > Z_{1-\alpha}$$

که در این مثال مقدار $d_0 = 0$ می‌باشد.

۱۱-۳۰ الف) در سطح معنی‌دار ۰/۰۵ داریم:

$$\alpha = ۰/۰۵ \quad \rightarrow \quad ۱-\frac{\alpha}{۲} = ۰/۹۷۵$$

$$Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۷۵} = ۱/۹۶$$

در این حالت فرض $H_۰$ رد می‌شود اگر $|Z| > ۱/۹۶$ حال داریم:

$$Z = \frac{(\overline{X}_۱ - \overline{X}_۲) - d_۰}{\sqrt{\frac{\sigma_۱^۲}{n_۱} + \frac{\sigma_۲^۲}{n_۲}}} = \frac{۷۴ - ۷۸}{\sqrt{\frac{۸۲}{۴۰} + \frac{۷۲}{۵۰}}} = \frac{-۴}{۱/۶۰۶} = -۲/۴۹$$

از آنجا که $Z = -۲/۴۹$ در بازه $(-۱/۹۶ \,,\, ۱/۹۶)$ قرار ندارد بنابراین در سطح معنی‌دار ۰/۰۵ اختلاف معنی‌داری بین دو کلاس A و B وجود دارد و فرض $H_۰$ رد می‌شود. به این ترتیب به احتمال بیشتر کلاس دوم دارای نتایج بهتری می‌باشد.

ب) برای سطح معنی‌دار ۰/۰۱ داریم:

$$\alpha = ۰/۰۱ \quad \rightarrow \quad ۱-\frac{\alpha}{۲} = ۰/۹۹۵$$

$$Z_{۱-\frac{\alpha}{۲}} = Z_{۰/۹۹۵} = ۲/۵۸$$

در این حالت مقدار آماره آزمون برابر با مقدار بدست آمده از بند الف می‌باشد که برابر $-۲/۴۹$ می‌باشد. و از آنجا که $-۲/۴۹$ در بازه $(-۲/۵۸ \,,\, ۲/۵۸)$ قرار دارد بنابراین در سطح معنی‌دار ۰/۰۱ اختلاف معنی‌داری بین دو کلاس وجود ندارد. توجه کنید که در حالت کلی نتایج حاصل از سطح معنی‌دار ۰/۰۵ را ملاک تصمیم‌گیری قرار می‌دهند. زیرا آماردانان نشان داده‌اند که تقریباً بهترین سطح برای ملاک بودن در تصمیم‌گیری‌ها، سطح معنی‌دار ۰/۰۵ می‌باشد.

۱۱-۳۱ مثال ۱۱: در کشاورز A و B در مزرعه‌های خود گندم می‌کارند، کشاورز A از نوعی ضد آفت جدید استفاده می‌کند. میانگین برداشت محصول در هر کیلومتر مربع از ۱۲ کیلومتر مربع از زمین کشاورز A برابر ۱۳۹ کیلو با انحراف معیار ۱۰ کیلو می‌باشد و در زمین کشاورز B برابر ۱۳۱ کیلو با انحراف معیار ۱۱ کیلو می‌باشد. آیا می‌توان ادعا کرد که در سطح معنی‌دار ۰/۰۵ و ۰/۰۱.

حل: فرض آزمون عبارتند از:

$H_۰ : \quad \mu_۱ = \mu_۲$ اختلاف در تولید تنها ناشی از شانس است.

$H_۱ : \quad \mu_۱ > \mu_۲$ ضد آفت در افزایش تولید موثر بوده است.

$\mu_۱$ میانگین تولید محصول کشاورز A و $\mu_۲$ میانگین تولید محصول کشاورز B می‌باشد.

آماره آزمون و ناحیه بحرانی در این حالت عبارتند از:

$$T = \frac{(\overline{X}_۱ - \overline{X}_۲) - d_۰}{S_P \sqrt{\frac{۱}{n_۱} + \frac{۱}{n_۲}}} > t_{۱-\alpha}(n_۱ + n_۲ - ۲)$$

الف) در سطح‌معنی‌دار ۰/۰۵ داریم:

$$\alpha = ۰/۰۵$$

$$۱-\alpha = ۰/۹۵ \qquad t_{۱-\alpha}(n_۱+n_۲-۲) = t_{۰/۹۵}(۱۲+۱۲-۲)$$

$$= t_{۰/۹۵}(۲۲) = ۱/۷۲$$

$$\overline{X}_۱ = ۱۳۹ \qquad S_۱ = ۱۰ \qquad n_۱ = ۱۲$$

$$\overline{X}_۲ = ۱۳۱ \qquad S_۲ = ۱۱ \qquad n_۲ = ۱۲$$

مقدار S_p برابر است با:

$$S_p = \sqrt{\frac{(n_۱-۱)\,S_۱^۲ + (n_۲-۱)\,S_۲^۲}{n_۱+n_۲-۲}} = \sqrt{\frac{۱۱(۱۰)^۲+۱۱(۱۱)^۲}{۱۲+۱۲-۲}} = ۱۰/۵۱$$

۱۱-۳۲ حال مقدار آماره آزمون برابر می‌شود با:

$$T = \frac{\overline{X}_۱-\overline{X}_۲}{S_p\sqrt{\frac{۱}{n_۱}+\frac{۱}{n_۲}}} = \frac{۱۳۹-۱۳۱}{۱۰/۵۱\sqrt{\frac{۱}{۱۲}+\frac{۱}{۱۲}}} = ۱/۸۵$$

با توجه به مقدار آماره آزمون $T = ۱/۸۵$ که از $t_{./۹۵}\,(۲۲) = ۱/۷۲$ بیشتر است می‌توان نتیجه گرفت که فرض $H_۰$ در سطح معنی‌دار ۰/۰۵ رد می‌شود و به این ترتیب ضدآفت در تولید محصول بیشتر موثر بوده است.

ب) در سطح معنی‌دار ۰/۰۱ داریم:

$$\alpha = ۰/۰۱ \quad \rightarrow \quad ۱-\alpha = ۰/۹۹$$

$$t_{۱-\alpha}\,(n_۱+n_۲-۲) = t_{./۹۹}\,(۲۲) = ۲/۵۱$$

از آنجا که $T = ۱/۸۵ \not> ۲/۵۱$ بنابراین فرض $H_۰$ رد نمی‌شود و تفاوت معنی‌داری در سطح ۰/۰۱ در تولید محصول وجود ندارد. اما از آنجا که ۰/۰۵ را ملاک تصمیم‌گیری می‌گیریم بنابراین در حالت کلی ضد آفت در تولید محصول بیشتر موثر بوده است.

۱۱-۳۳ ۱۱ . ۱۰ آزمون فرض برای واریانسهای دو جامعه

اگر دو جامعه نرمال داشته باشیم و $n_۱$ نمونه از جامعه اول با انحراف معیار $S_۱$ و $n_۲$ نمونه با انحراف معیار $S_۲$ از جامعه دوم انتخاب کنیم در این صورت برای انجام آزمونهایی در مورد واریانسهای دو جامعه می‌توانیم از آماره و ناحیه بحرانی زیر استفاده کنیم:

۱- در آزمون فرض $\begin{cases} H_۰ : \sigma_۱^۲ = \sigma_۲^۲ \text{ یا } \sigma_۱^۲ \le \sigma_۲^۲ \\ H_۱ : \sigma_۱^۲ > \sigma_۲^۲ \end{cases}$ فرض $H_۰$ رد می‌شود اگر و فقط اگر: $F = \frac{S_۱{}^۲}{S_۲{}^۲} > F_{۱-\alpha}\,(n_۱-۱\,,\,n_۲-۱)$

۲- در آزمون فرض $\begin{cases} H_۰ : \sigma_۱^۲ = \sigma_۲^۲ \text{ یا } \sigma_۱^۲ \ge \sigma_۲^۲ \\ H_۱ : \sigma_۱^۲ < \sigma_۲^۲ \end{cases}$ فرض $H_۰$ رد می‌شود اگر و فقط اگر: $F = \frac{S_۱{}^۲}{S_۲{}^۲} < F_{\alpha}\,(n_۱-۱\,,\,n_۲-۱)$

۳- در آزمون فرض $\begin{cases} H_۰ : \sigma_۱^۲ = \sigma_۲^۲ \\ H_۱ : \sigma_۱^۲ \ne \sigma_۲^۲ \end{cases}$ فرض $H_۰$ رد می‌شود اگر و فقط اگر:

$$F = \frac{S_۱{}^۲}{S_۲{}^۲} < F_{\frac{\alpha}{۲}}\,(n_۱-۱\,,\,n_۲-۱) \text{ یا } F = \frac{S_۱{}^۲}{S_۲{}^۲} > F_{۱-\frac{\alpha}{۲}}\,(n_۱-۱\,,\,n_۲-۱)$$

۱۱-۳۴ مثال ۱۲: استادی یک درس را در دو کلاس A و B تدریس می‌کند. کلاس A تعداد ۱۶ دانشجو و کلاس B تعداد ۲۵ دانشجو دارد. استاد از هر دو کلاس یک امتحان را می‌گیرد. میانگین نمرات دانشجویان در هر کلاس تقریباً برابر است اما واریانس کلاس A برابر ۸۶/۴ و واریانس کلاس B برابر ۱۵۰ می‌باشد. (نمرات از ۱۰۰ واحد می‌باشند) آیا در دو سطح معنی‌دار ۰/۰۵ و ۰/۰۱ می‌توان نتیجه گرفت که واریانس کلاس B از واریانس کلاس A بیشتر می‌باشد؟

حل: فرضهای آزمون عبارتند از:

$H_۰ :\quad \sigma_۱ = \sigma_۲$ اختلاف واریانس‌ها ناشی از شانس می‌باشد.

H_1 : $\sigma_1 > \sigma_2$ واریانس کلاس B از واریانس کلاس A بیشتر است.

داریم:

$$n_1 = ۱۶ \qquad S_1^2 = ۸۶/۴$$
$$n_2 = ۲۵ \qquad S_2^2 = ۱۵۰$$

آماره آزمون و ناحیه بحرانی عبارتند از:

$$F = \frac{S_1^2}{S_2^2} < F_\alpha(n_1 - ۱, n_2 - ۱)$$

الف) در سطح معنی‌دار ۰/۰۵ داریم:

$$\alpha = ۰/۰۵$$

$$F_\alpha(n_1 - ۱, n_2 - ۱) = \frac{۱}{F_{۱-\alpha}(n_2 - ۱, n_1 - ۱)}$$

$$\Rightarrow F_{/۰۵}(۱۵, ۲۴) = \frac{۱}{F_{/۹۵}(۲۴, ۱۵)} = \frac{۱}{۲/۱۱} = ۰/۴۷۳$$

۱۱-۳۵ حال آماره آزمون را بدست می‌آوریم:

$$F = \frac{S_1^2}{S_2^2} = \frac{۸۶/۴}{۱۵۰} = ۰/۵۷۶$$

از آنجا که $f = ۰/۵۷۶ \not< ۰/۴۷۳$ بنابراین نمی‌توانیم فرض H_0 را رد کنیم و در نتیجه می‌توان گفت که واریانس دو کلاس در سطح معنی‌دار ۰/۰۵ تفاوت معنی‌داری با یکدیگر ندارند.

ب) در سطح معنی‌دار ۰/۰۱ داریم:

$$\alpha = ۰/۰۱$$

$$F_\alpha(n_1 - ۱, n_2 - ۱) = \frac{۱}{F_{۱-\alpha}(n_2 - ۱, n_1 - ۱)}$$

$$\Rightarrow F_{/۰۱}(۱۵, ۲۴) = \frac{۱}{F_{/۹۹}(۲۴, ۱۵)} = \frac{۱}{۲/۸۹} = ۰/۳۴۶$$

باز هم $f = ۰/۵۷۶ \not< ۰/۳۴۶$ بنابراین در سطح معنی‌دار ۰/۰۱ هم می‌توانیم H_0 را رد کنیم یعنی در این حالت هم تفاوت معنی‌داری میان واریانسهای دو کلاس وجود ندارد.

فصل دوازدهم

S-۱

۱- دگرسیون-۱

دگرسیون

در فصلهای قبلی بیشتر تجزیه و تحلیل‌های آماری روی یک صفت از جامعه یا متغیر تصادفی متمرکز بود در این فصل قصد داریم بصورت همزمان دو صفت از جامعه یا دو متغیر تصادفی را مورد بررسی قرار دهیم. این بررسی شامل پیش بینی مقادیر یکی از متغیرۀا از روی مقادیر متغیر دیگراست که به مساله برگشت یا دگرسیون معروف می‌باشد.

به عنوان مثال فرض کنید بخواهیم تاثیر میزان مصرف شیر را در افزایش قد بدست بیاوریم و یا بخواهیم میزان وزن فرزند را از روی وزن پدرش پیش بینی کنیم. ملاحظه می‌کنید که در این گونه مسایل دو متغیر تصادفی مورد مطالعه به نوعی به یکدیگر وابسته می‌باشند. به عبارت دقیقتر یک متغیر تصادفی مثل X را مستقل و متغیر تصادفی Y را وابسته به آن در نظر می‌گیریم و یا برعکس Y را مستقل و X را وابسته به آن در نظر می‌گیریم.

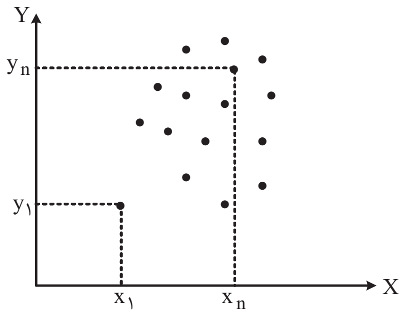
آشکار است که انتخاب هر یک از دو حالت به نوع مساله بستگی دارد.

در مسایل دگرسیون برای یافتن رابطه بین متغیر تصادفی مستقل X و متغیر وابسته Y ابتدا یک نمونه nتایی از متغیر تصادفی x جمع آوری می‌کنیم که نتایج آن بصورت $x_1, x_2, \cdots, x_n$ می‌باشند. سپس مقادیر متناظر با هر یک از نمونه‌های بدست آمده (x_i ها) را که همان مقادیر معادل متغیر تصادفی وابسته Y می‌باشند بدست می‌آوریم. به این ترتیب برای x_i ها مقادیر متناظر $y_1, y_2, \cdots, y_n$ بدست می‌آیند. که می‌توانیم نتیجه را بصورت زوج‌های مرتب زیر نشان دهیم.

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \cdots, (x_n, y_n)$$

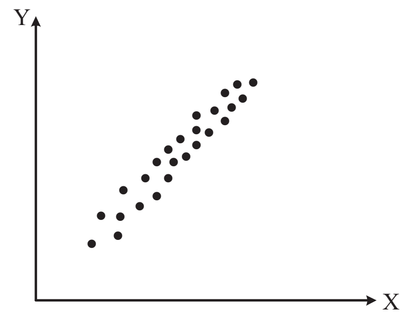
۲- دگرسیون-۲ به عنوان مثال مقادیر x_i ها می‌توانند میزان طول قد افراد یک جامعه و مقادیر y_i ها میزان مصرف شیر هر یک از نمونه‌ها باشند. به این ترتیب زوج مرتب $(۱۸۰,۲)$ بیانگر این است که در نمونه‌گیری یکی از افراد جامعه دارای صول قد $cm۱۸۰$ بوده است و و وی روزانه دو لیوان شیر مصرف کرده است.

پس از بدست آمدن زوج‌های مرتب (x_i, y_i) ملاحظه می‌نید که هر زوج مرتب می‌تواند معادل یک نقطه در صفحه باشد. با رسم نقاط مورد نظر در صفحه یک تصویر کلی از رابطه X و Y بدست می‌آوریم. به شکلهای زیر که برای سه نمونه جداگانه می‌باشند توجه کنید:



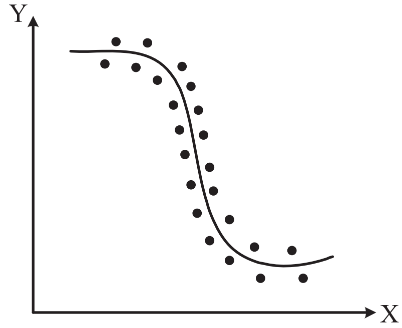
نمودار ۱

در این نمودار ملاحظه می‌کنید که زوج‌های مرتب (x_i, y_i) بصورت کاملاً پراکنده توزیع شده‌اند و به این ترتیب نتیجه می‌گیریم که رابطه‌ای بین متغیرهای تصادفی X و Y وجود ندارد.



نمودار ۲

ملاحظه می‌کنید که در این نمونه یک رابطه خطی بین مقادیر x_i و y_i وجود دارد و می‌توان نوشت $y_i = a x_i + b$.



نمودار ۳

در این نمودار X و Y به یکدیگر وابسته می‌باشند اما این وابستگی از نوع غیر خطی می‌باشد.

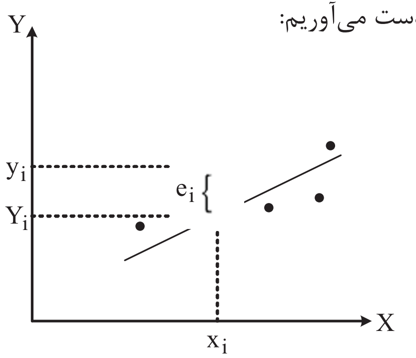
با توجه به دو نمودار ۲ و ۳ این طور به نظر می‌رسد که در حالت کلی دو متغیر تصادفی X و Y اگر از یکدیگر مستقل نباشند یا بصورت خطی و یا بصورت غیر خطی به یکدیگر وابسته می‌باشند.

قبلاً نشان دادیم که برای اندازه‌گیری میزان وابستگی دو متغیر X و Y می‌توان از ضریب همبستگی استفاده نمود. اما در مسایل دگرسیون پیش بینی متغیر Y از روی X و یا بالعکس از اهمیت ویژه‌ای برخوردار است. بنابراین نیازمند روشی هستیم که بتوان در صورت نیاز با ثابت در نظر گرفتن یکی از مقادیر X یا Y مقدار دیگری را بدست بیاوریم. برای این منظور مفهوم بردازش منحنی را مطرح می‌کنیم.

۱- دگرسیون خطی – ۳ ۱۲ . ۱ دگرسیون خطی

اگر بین دو متغیر X و Y یک رابطه خطی وجود داشته باشد می‌توانیم یک خط را طوری رسم کنیم که نقاط (x_i , y_i) کمترین فاصله را با خط مورد نظر داشته باشند. به این عمل بردازش منحنی می‌گویند. معادله خط را بصورت $Y = aX + b$ در نظر می‌گیریم که در آن a و b مقادیر مجهول می‌باشند و در این حالت X متغیر مستقل و Y متغیر وابسته به آن در نظر گرفته می‌شود. مقادیر aو b می‌بایستی طوری محاسبه شوند که مجموع فاصله نقاط (x_i , y_i) از خط $Y = aX + b$ حداقل شود. در این حالت به $Y = aX + b$ معادله دگرسیون Y می‌گویند.

برای حداقل نمودن فاصله نقاط (x_i , y_i) از خط دگرسیون مقدار خطای e_i را مطابق نمودار زیر بدست می‌آوریم:



با توجه به نمودار Y_i مقدار پیش بینی شده توسط خط دگرسیون می‌باشد که با مقدار واقعی y_i به اندازه $e_i = |y_i - Y_i|$ فاصله دارد که این فاصله همان خطای پیش بینی می‌باشد.

برای بدست آوردن بهترین نتیجه، مجموع مربعات خطا را حداقل می‌کنیم که عبارتست از:

$$SSE = E = \sum_{i=۱}^{n} (y_i - Y_i)^۲$$

که در آن:

$$Y = aX + b \Rightarrow Y_i = aX_i + b \Rightarrow E = \sum_{i=۱}^{n} (aX_i + b - Y_i)^۲$$

۲- دگرسیون خطی-۴ برای مینیمم نمودن E مشتقات پاره‌ای آنرا نسبت به a و b بدست آورده و برابر صفر قرار می‌دهیم:

$$\frac{\partial E}{\partial a}=\sum_{i=۱}^{n} ۲x_i(ax_i+b-y_i)=۰$$

$$\frac{\partial E}{\partial b}=\sum_{i=۱}^{n} ۲(ax_i+b-y_i)=۰$$

به این ترتیب دستگاه معادلات زیر بدست می‌آید:

$$\begin{cases}\sum_{i=۱}^{n} ۲x_i(ax_i+b-y_i)=۰\\ \sum_{i=۱}^{n} ۲(ax_i+b-y_i)=۰\end{cases}$$

با حل این دستگاه مقادیر مجهول a و b بدست می‌آیند. که عبارتند از:

$$a=\frac{S_{xy}}{S_{xx}} \quad , \quad b=\overline{y}-a\overline{X}$$

که در آن:

$$S_{xy}=\sum_{i=۱}^{n}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=۱}^{n}x_i y_i-n\overline{x}\,\overline{y}$$

$$S_{xx}=\sum_{i=۱}^{n}(x_i-\overline{x})^۲=\sum_{i=۱}^{n}x_i^۲-n\overline{x}^۲$$

به معادله $y=ax+b$ خط دگرسیون Y روی x می‌گویند.یعنی Y از روی x بدست آمده است.

۱-مثال۱-۵ مثال ۱: در یک بررسی میزان قد پدران و فرزندان آنها بصورت جدول زیر بدست آمده است:

قد پدران X	۱۶۵	۱۶۰	۱۷۰	۱۶۳	۱۷۳	۱۵۷	۱۷۸	۱۶۸	۱۷۳	۱۷۰	۱۷۵	۱۸۰
قد فرزندان Y	۱۷۳	۱۶۸	۱۷۳	۱۶۵	۱۷۵	۱۶۸	۱۷۳	۱۶۵	۱۸۰	۱۷۰	۱۷۳	۱۷۸

مطلوبست:

الف) بردازش خطی داده‌ها

ب) رسم نمودار داده‌ها و خط دگرسیون.

ج) با توجه به معادله دگرسیون پیش بینی کنید که اگر قد پدری cm۱۸۵ باشد قد فرزند او چقدر خواهد بود؟

حل: برای بدست آوردن a و b در معادله $y=ax+b$ جدول زیر را تشکیل می‌دهیم:

x	y	$x^۲$	xy	$y^۲$
۱۶۵	۱۷۳	۲۷۲۲۵	۲۸۵۴۵	۲۹۹۲۹
۱۶۰	۱۶۸	۲۵۶۰۰	۲۶۸۸۰	۲۸۲۲۴
۱۷۰	۱۷۳	۲۸۹۰۰	۲۹۴۱۰	۲۹۹۲۹
۱۶۳	۱۶۵	۲۶۵۶۹	۲۶۸۹۵	۲۷۲۲۵
۱۷۳	۱۷۵	۲۹۹۲۹	۳۰۲۷۵	۳۰۶۲۵
۱۵۸	۱۶۸	۲۴۹۶۴	۲۶۵۴۴	۲۸۲۲۴
۱۷۸	۱۷۳	۳۱۶۸۴	۳۰۷۹۴	۲۹۹۲۹
۱۶۸	۱۶۵	۲۸۲۲۴	۲۷۷۲۰	۲۷۲۲۵
۱۷۳	۱۸۰	۲۹۹۲۹	۳۱۱۴۰	۳۲۴۰۰
۱۷۰	۱۷۰	۲۸۹۰۰	۲۸۹۰۰	۲۸۹۰۰
۱۷۵	۱۷۳	۳۰۶۲۵	۳۰۲۷۵	۲۹۹۲۹
۱۸۰	۱۷۸	۳۲۴۰۰	۳۲۰۴۰	۳۱۶۸۴
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

۲- مثال ۱-۶

$$\overline{X}=\frac{۱}{n}\sum x_i=\frac{۱}{۱۲}\,۲۰۳۳=۱۶۹/۴۱$$

$$\overline{Y}=\frac{۱}{n}\sum y_i=\frac{۱}{۱۲}\,۲۰۶۱=۱۷۱/۷۵$$

$$S_{xy}=\sum_{i=۱}^{n} x_i\,y_i-n\,\overline{x}\,\overline{y}=۳۴۹۴۱۸-۱۲(۱۶۹/۴۱)\,(۱۷۱/۷۵)=۲۵۰/۲۵$$

$$S_{xx}=\sum x_i^{۲}-n\,\overline{x}^{۲}=۳۴۴۹۴۹-۱۲\,(۱۶۹/۴۱)^{۲}=۵۲۴/۹۱$$

$$a=\frac{S_{xy}}{S_{xx}}=\frac{۲۵۰/۲۵}{۵۲۴/۹۱}=۰/۴۷۷$$

$$b=\overline{Y}-a\,\overline{X}=۱۷۱/۷۵-۰/۴۷۷\,(۱۶۹/۴۱)=۹۰/۹$$

بنابراین معادله خط دگرسیون عبارتست از:

$$y=۰/۴۷۷\,x+۹۰/۹$$

ب) نمودار داده‌ها و خط دگرسیون عبارتست از:

ج) با توجه به خط دگرسیون می‌توان قد فرزند یک پدر با قد ۱۸۵ را به فرم زیر بدست آورد.

$$y=۰/۴۷۷\times۱۸۵+۹۰/۹=۱۷۹/۱۴۵$$

۳-S ۱۲. ۲ ضریب همبستگی و دگرسیون

۷-۱

قبلاً ضریب همبستگی دو متغیر X و Y را بصورت زیر تعریف نمودیم:

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\,\sigma_Y}$$

می‌توان با ساده نمودن معادله دگرسیون مقدار ضریب همبستگی را وارد معادله نمود:

$$y = a x + b \qquad a = \frac{S_{xy}}{S_{xx}} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}\left(\frac{\sigma_y}{\sigma_x}\right) = \rho_{xy}\frac{\sigma_y}{\sigma_x}$$

توجه کنید که در اینجا در محاسبه ρ_{xy} و σ_x و σ_y از مقادیر نمونه‌های مشاهده شده استفاده می‌شود.

به این ترتیب معادله خط دگرسیون بصورت زیر بدست می‌آید:

$$y - \bar{y} = \rho_{xy}\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

توجه کنید که همواره برای داده‌های مشاهده شده (x_i , y_i) دو معادله دگرسیون وجود دارد یک معادله بر حسب y نسبت به x می‌باشد و معادله دیگر بر حسب x نسبت به y می‌باشد معادله خط دگرسیون x روی y را می‌توان بصورت زیر بدست آورد:

$$x = \alpha y + \beta$$

$$\alpha = \frac{S_{xy}}{S_{yy}} \qquad \beta = \bar{x} - \alpha \bar{y}$$

$$S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i - n\bar{y}^2$$

۲–۸ با نوشتن معادله دگرسیون x روی y و استفاده از ضریب همبستگی بدست می‌آوریم:

$$x - \bar{x} = \rho_{xy}\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

بنابراین در حالت کلی دو خط دگرسیون y روی x و x روی y خواهیم داشت که عبارتند از:

$$y - \bar{y} = \rho_{xy}\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$x - \bar{x} = \rho_{xy}\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

با برابر قرار دادن $x = y$ در معادلات فوق محل تلاقی دو خط دگرسیون نقطه $\begin{cases} x = \bar{x} \\ y = \bar{y} \end{cases}$ بدست می‌آید. همچنین توجه کنید که:

$$a = \rho_{xy}\frac{\sigma_y}{\sigma_x}$$

$$\Rightarrow a\alpha = \rho_{xy}\frac{\sigma_y}{\sigma_x}\,\rho_{xy}\frac{\sigma_x}{\sigma_y} = \rho_{xy}^2$$

$$\alpha = \rho_{xy}\frac{\sigma_x}{\sigma_y}$$

بنابراین $a\alpha = \rho_{xy}^2$ و می‌بایستی مقداری در بازه $[0, 1]$ داشته باشد.

۲–۹ مثال ۲: در مثال ۱ مطلوبست:

الف) محاسبه ضریب همبستگی نمونه با توجه به مقادیر محاسبه شده a و b.

ب) معادله دگرسیون x روی y از روی ضریب همبستگی.

ج) رسم نمودار داده‌ها و دو خط دگرسیون y روی x و x روی y.

د) پیش بینی قد پدر اگر قد فرزند وی cm۱۷۹ باشد.

حل: الف) ضریب همبستگی با توجه به روابط ارایه شده عبارتست از:

$$a=\rho_{xy}\frac{\sigma_y}{\sigma_x}=۰/۴۷۷$$

$$S_{yy}=\sum y_y^{۲}-n\,\overline{y}^{۲}=۳۵۴۲۲۳-۱۲(۱۷۱/۷۵)^{۲}=۲۴۶/۲۵$$

$$\sigma_y=\sqrt{S_{yy}}=\sqrt{۲۴۶/۲۵}=۱۵/۶۹$$

$$\sigma_x=\sqrt{S_{xx}}=\sqrt{۵۲۴/۹۱}=۲۲/۹۱$$

$$\Rightarrow a=۰/۴۷۷=\rho_{xy}\,\frac{۱۵/۶۹}{۲۲/۹۱}\quad\Rightarrow\quad \rho_{xy}=۰/۶۹۶$$

ب) معادله دگرسیون x روی y با توجه به ρ_{xy} برابر است با:

$$x=\alpha y+\beta=\rho_{xy}\,\frac{\sigma_x}{\sigma_y}(y-\overline{y})+\overline{x}=۰/۶۹۶(\frac{۲۲/۹۱}{۱۵/۶۹})(y-۱۷۱/۷۵)+۱۶۹/۴۱$$

$$\Rightarrow\quad x=۱/۰۱۶\,y-۵/۱۲$$

ج) نمودار داده‌ها و دو خط دگرسیون و روی x و x روی y عبارتست از:

د) با استفاده از خط دگرسیون x روی y مقدار x را برای $y=۱۷۹$ پیش بینی می‌کنیم:

$$x=۱/۰۱۶\,y-۵/۱۲=۱/۰۱۶\,(۱۷۹)-۵/۱۲=۱۷۶/۷۴۴$$

۱۲. ۳ دگرسیون منحنی‌های چند جمله‌ای

با تعمیم روشی که برای برازش داده‌ها بصورت خطی ارایه شد به سادگی می‌توان به داده‌ها یک منحنی چند جمله‌ای بفرم $y=a_n x^n+a_{n-۱}x^{n-۱}+\cdots+a_۱ x+a_۰$ برازش داد. این کار را برای یک منحنی درجه دوم نشان می‌دهیم:

$$Y=ax^{۲}+bx+c$$

$$E=\sum_{i=۱}^{n}(y_i-Y_i)^{۲}=\sum_{i=۱}^{n}(ax_i{}^{۲}+bx_i+c-y_i)^{۲}$$

حال مجدداً با مشتق‌گرفتن نسبت به مقادیر مجهول a و b و c و برابر صفر قرار دادن معادلات یک دستگاه بدست می‌آوریم که با حل آن a و b و c بدست می‌آیند:

$$\frac{\partial E}{\partial a}=\sum_{i=1}^{n} 2x_i^2(ax_i^2+bx_i+c-y_i)=0$$

$$\frac{\partial E}{\partial b}=\sum_{i=1}^{n} 2x_i(ax_i^2+bx_i+c-y_i)=0$$

$$\frac{\partial E}{\partial c}=\sum 2(ax_i^2+bx_i+c-y_i)=0$$

با داشتن مقادیر نمونه‌های (x_i, y_i) به سادگی می‌توان دستگاه فوق را حل نموده و مقادیر مجهول a و b و c را بدست آورد.

۱۱ – estenbate amari

استنباط آماری بر روی ضرایب دگرسیونی

در بخش قبل بر اسا نمونه $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ مدل ساده دگرسیونی را معرفی نمودیم و به کمک روش کمترین مربعات برای $y_i = a + bx_i + E_i$ برآوردگرهای a و b $\hat{a}=\bar{y}-\hat{b}\bar{x}$, $\hat{b}=\frac{S_{xy}}{S_{xx}}$ را معرفی نمودیم.

در این بخش می‌خواهیم مدل احتمالی دگرسیون را معرفی کرده و بر اساس آن برخی از فرمهای آماری را بر روی a و b آزمون کنیم.

فرض کنید در مدل ساده دگرسیونی خطا یعنی E_i متغیر تصادفی با توزیع نرمال به فرم زیر باشد:

$$E_i \sim N(0,\sigma^2)$$

همچنین متغیرهای تصادفی E_i به ازای $i=1,2,\cdots,n$ را مستقل از هم در نظر می‌گیریم در این صورت چون Y_i یک ترکیب خطی از متغیر تصادفی E_i می‌باشد بنابراین:

$$Y_i \sim N(a+bx_i,\sigma^2)$$

۱۲ – estenbate amari

محاسبه برآوردگرهای a و b و σ^2

به ازای هر مقدار ثابت X_i تابع چگالی متغیر تصادفی Y_i برابر سات با:

$$f_{Y_i|X_i}(y_i|x_i)=\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{Y_i-(a+bx_i)}{\sigma}\right)^2} \quad -\infty<y_i<\infty$$

می‌توان به کمک روش ماکزیمم درستنمایی که روش دیگری است ۶ برآوردگرهای a و b و σ بذای بدست آوردن برآوردگرهای ماکزیمم درستنمایی پارامترهای a و b و σ به کمک نمونه تصادفی $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ ابتدا از تابع درستنمایی (یا لگاریتم آن که که ساده‌تر است) نسبت به a و b و σ مشتق می‌گیریم.

عبارتهای حاصل را برابر صفر قرار می‌دهیم، و سپس دستگاه معادلات حاصل را حل می‌کنیم.

بنابراین با مشتق گیری جزیی از:

$$\mathrm{Ln}(L)=-n\,\mathrm{Ln}(G)-\frac{n}{2}\mathrm{Ln}(2\pi)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[Y_i-(a+bx_i)\right]^2$$

نسبت به a و b و σ و برابر گذاشتن عبارت‌های حاصل ۶ به دست می‌آوریم:

$$\frac{\partial \operatorname{Ln}(L)}{\partial a}=\frac{۱}{\sigma^{۲}}+\sum_{i=۱}^{n}\left[Y_i-(a+b x_i)\right]=۰$$

$$\frac{\partial \operatorname{Ln}(L)}{\partial b}=\frac{۱}{\sigma^{۲}}+\sum_{i=۱}^{n} X_i\left[Y_i-(a+b x_i)\right]=۰$$

$$\frac{\partial \operatorname{Ln}(L)}{\partial \sigma}=\frac{-n}{\sigma}+\frac{۱}{\sigma^{۳}}+\sum_{i=۱}^{n}\left[Y_i-(a+b x_i)\right]^{۲}=۰$$

از دو معادله اول $\hat{a}$و $\hat{b}$ مشابه با برآوردگرهای به روش کمترین مربعات به صورت زیر به دست می‌آید.

$$\hat{b}=\frac{S_{xy}}{S_{xx}}$$

$$\hat{a}=\overline{Y}-\hat{b}\,\overline{X}$$

با قرار دادن $\hat{a}$و $\hat{b}$ در معادله سوم برآوردگر $\hat{\sigma}^{۲}$ برابر می‌شود با:

$$\hat{\sigma}^{۲}=\frac{S_{yy}-\hat{b}\,S_{xy}}{n}=\frac{SSE}{n}$$

اگر در براوردگر ماکزیمم درستنمایی $\hat{\sigma}^{۲}$، n را به $n-۲$ تبدیل کنیم آنگاه برآوردگر $S^{۲}=\frac{SSE}{n-۲}$ نااریب خواهد شد یعنی در این حالت:

$E\left[S^{۲}\right]=E\left[\frac{SSE}{n-۳}\right]=\sigma^{۲}$ می‌شود.

۱۳ mesal

مثال ۱۳: فرض کنیم یک کمپانی می‌خواهد تأثیر تبلیغات را در فروش کالاهای تولید شده بررسی کند بدین منظور داده‌های زیر را بعد از ۱۰ ماه بدست می‌آورد. این داده‌ها در جدول زیر مرتب شده است.

به کمک داده‌های بدست آمده برآورد a و b و $\sigma^{۲}$ را بدست اورید.

حل:

$$S_{xx}=\sum_{i=۱}^{n} x_i^۲-\frac{(\sum_{i=۱}^{n} x_i)^۲}{n}=۹/۲۸-\frac{(۹/۴)^۲}{۱۰}=۰/۴۴۴$$

$$S_{xy}=\sum_{i=۱}^{n} x_i\,y_i-\frac{(\sum_{i=۱}^{n} x_i)(\sum_{i=۱}^{n} y_i)}{n}=۹۲۴/۸-\frac{(۹/۴)(۹۵۹)}{۱۰}=۲۳/۳۴$$

$$\overline{Y}=\frac{\sum_{i=۱}^{n} Y_i}{n}=\frac{۹۵۹}{۱۰}=۹۵/۹\quad,\quad \overline{X}=\frac{\sum_{i=۱}^{n} X_i}{n}=\frac{۹/۴}{۱۰}=۰/۹۴$$

پس:

$$\hat{b}=\frac{S_{xy}}{S_{xx}}=\frac{۲۳/۳۴}{۰/۴۴۴}=۵۲/۵۶۷۶\approx ۵۲/۵۷$$

$$\hat{a}=\overline{Y}-\hat{b}\,\overline{X}=۹۵/۹-(۵۲/۵۶۷۶)(۰/۹۴)\approx ۴۶/۴۹$$

بنابراین معادله خط دگرسیون به صورت: $y=۴۶/۴۹+۵۲/۵۷\,x$ می‌شود.

۱۴Mesal

برای محاسبه برآوردگر نااریب $\sigma^۲$ یعنی $S^۲$ به صورت زیر عمل می‌کنیم:

$$S_{yy}=\sum_{i=۱}^{n}(y_i-\overline{y})^۲=\sum_{i=۱}^{n} y_i^۲-\frac{(\sum_{i=۱}^{n} y_i)^۲}{n}=۹۳/۵۶۹-\frac{(۹۵۹)^۲}{۱۰}=۱۶۰۰/۹$$

با جایگذاری S_{yy} و $\hat{b}$ و S_{xy} در فرمول SSE خواهیم داشت:

$$SSE=S_{xy}-\hat{b}\,S_{xy}=۱۶۰۰/۹-(۵۲/۵۶۷۶)(۲۳/۳۴)=۳۷۰۰$$

پس:

$$S^۲=\frac{SSE}{n-۲}=\frac{۳۷۳/۹۷}{۸}=۴۶/۷۵۰$$

حال از برآوردگرهای ماکزیمم ددرستنمایی ضرایب دگرسیون ساده در آزمون فرضهایی درباره a و b و در ساختن فاصله‌های اطمینان برای این پارامترها استفاده می‌کنیم. بدین منظور نتایج زیر را بدون اثبات می‌پذیریم:

(۱) $$\hat{a}\sim N\left(a\,,\frac{\sigma^۲\sum_{i=۱}^{n} X_i^۲}{n\,S_{xx}}\right)$$

(۲) $$\hat{b}\sim N\left(b\,,\frac{\sigma^۲}{S_{xx}}\right)$$

(۳) اگر $S^۲=\frac{SSE}{n-۲}$ آنگاه $\frac{SSE}{\sigma^۲}=\frac{(n-۲)\,S^۲}{\sigma^۲}\sim X^۲_{(n-۲)}$

(۴) $S^۲,\hat{a}$ مستقل از هم هستند.

(۵) $S^۲,\hat{b}$ مستقل از هم هستند.

۱۵ - estenbate amari

استباط آماری بر روی b

رابطه خطی بین x و y اولین موضوعی است که مورد بررسی قرار می‌گیرد. بنابراین علاقمندیم بدانیم که آیا داده‌ها به اندازه کافی اطلاعات در رابطه با ارتباط خطی بودن y با x می‌دهد یا نه؟ آزمون فرضهای آماری که بر روی پارامتر b ساخته می‌شود ارتباط بین y و x را روشن می‌کند بعنوان مثال اگر y با افزایش یا کاهش x تغییر می‌کند می‌توان فرض $b = 0$ را در مقابل فرض $b \neq 0$ ازمون کرد. با توجه به اینکه برآوردگر $\hat{b}$ دارای توزیع $\hat{b} \sim N\left(b, \frac{\sigma^{1}}{S_{xx}}\right)$ می‌باشد می‌توان از تست آماری $T = \frac{\hat{b}-b_0}{\frac{\sigma}{S_{xx}}} \sim t(n-2)$.

برای آزمون فرض‌های

$$(1)\quad \begin{array}{l} H_0: b = b_0 \\ H_1: b \neq b_0 \end{array} \qquad (2)\quad \begin{array}{l} H_0: b = b_0 \\ H_1: b > b_0 \end{array} \qquad (3)\quad \begin{array}{l} H_0: b = b_0 \\ H_1: b < b_0 \end{array}$$

استفاده نمود.

نواحی بحرانی برای آزمون فرض‌های بالا در سطح معنی‌دار α را می‌توان در جدول زیر خلاصه نمود.

هم‌چنین می‌توان با استفاده از تست آماری T با ضریب اطمینان $1 - \alpha$ فاصله اطمینان زیر را برای پارامتر b معرفی نمود.

$$P\left(\hat{b} - t^{(n-2)}_{(1-\frac{\alpha}{2})} + \frac{S}{\sqrt{S_{xx}}} < b < \hat{b} + t^{(n-2)}_{(1-\frac{\alpha}{2})} \frac{S}{\sqrt{S_{xx}}}\right) = 1-\alpha$$

۱۶-مثال ۱-۴

مثال ۴: با توجه به داده‌های مثال ۱ تعیین کنید که آیا پارامتر b اختلاف معنی‌داری از مقدار ۰ دارد یا نه (بوسیله استفاده از یک مدل خطی بین فروش ماهانه و مقدار تبلیغات)

حل: می‌خواهیم آزمون فرض زیر را انجام دهیم:

$$H_0: b = b_0$$
$$H_1: b \neq b_0$$

از تست آماری:

$$T = \frac{\hat{b}-0}{\frac{S}{\sqrt{S_{xx}}}} \sim t^{(n-2)}_{(1-\frac{\alpha}{2})} = t^{(8-2)}_{(1-\frac{\alpha}{2})} = t^{(6)}_{(1-\frac{\alpha}{2})}$$

استفاده می‌کنیم. با استفاده از جدول توزیع T به ازای $\alpha = ۰/۵$ $\;t^{(6)}_{۰/۹۷۰} = ۲/۳۰۶$ بنابراین چون:

$$T = \frac{\hat{b}}{\frac{S}{\sqrt{S_{xx}}}} = \frac{۵۲/۵۷}{\frac{۶/۸۴}{\sqrt{۰/۴۴۴}}} = ۵/۱۲ > ۲/۳۰۶$$

باشد پس فرض $H_\circ$ را رد می‌کنیم. بنابراین بر اساس این مشاهدات می‌توان گفت که هزینه‌های مربوط به تبلیغات تأثیر در پیش بینی فروش ماهانه کالا دارد. با توجه به اطلاعات داده شده می‌توان یک فاصله اطمینان با ضریب اطمینان $۱-\alpha = ۰/۹۵$ نیز برای یک پارامتر b بدست آورد.

جایگذاری مقادیر $\hat{b}$ و S و S_{xx} و $t^{(n-۲)}_{(۱-\frac{\alpha}{۲})}$ در $\hat{b} \pm t^{(n-۲)}_{(۱-\frac{\alpha}{۲})} \times \frac{S}{\sqrt{S_{xx}}}$ اطمینان بدست می‌آید.

$$۵۲/۵۷ \pm ۲/۳۰۶ \frac{۶/۸۴}{\sqrt{۰/۴۴۴}} \quad \Rightarrow \quad ۵۲/۵۷ \pm ۲۳/۶۷ \quad \Rightarrow \quad [۲۸\,,۹۰\,,۷۶]$$

این فاصله نشان می‌دهد که با احتمال ۰/۹۵ مقدار پارامتر b که بوسیله برآوردگر $\hat{b}$ برآورد می‌شود را دارا می‌باشد.

۱۷-مثال ۲-۴

بطور مشابه می‌توان برای پارامتر a فاصله اطمینان و آزمون فرض ساخت. با توجه به اینکه برآوردگر $\hat{a}$ دارای توزیع $\hat{a} \sim N\left(a, \frac{\sigma^۲ \sum x_i^۲}{n S_{xx}}\right)$ می‌باشد می‌توان از تست آماری $T = \frac{\hat{a}-a}{S\sqrt{\frac{\sum x_i^۲}{n S_{xx}}}} \sim t(n-۲)$ استفاده نمود و خواص بحرانی را برای فرض‌های آماری مختلف بر روی پارامتر b در سطح معنی‌دار α در جدول زیر ساخت. همچنین یک فاصله اطمینان با ضریب اطمینان $(۱-\alpha)$ ۱۰۰% برای a بصورت:

$$P\left(\hat{a} - t^{(n-۲)}_{(۱-\frac{\alpha}{۲})} S\sqrt{\frac{\sum x_i^۲}{n S_{xx}}} < a < \hat{a} + t^{(n-۲)}_{(۱-\frac{\alpha}{۲})} S\sqrt{\frac{\sum x_i^۲}{n S_{xx}}}\right) = ۱-\alpha = \circ$$

برآورد $E[Y|X]$ مقدار میانگین Y به شرط مقدار داده شده X.

برآورد میانگین Y به شرط مقدار داده شده X یکی از مسائل مهم است که می‌بایست مورد بررسی قرار گیرد. بعنوان مثال ممکن است مسئول حفاظت جان کارگران کارخانه علاقمند باشد که متوسط حوادثی که برای کارگران اتفاق می‌افتد، به ازای ساعات خاصی از آموزش که به کارگران داده می‌شود را برآورد کند.

فرض کنید که x و y یک رابطه خطی بر اساس مدل احتمالی دگرسیونی داشته باشد بطوریکه:

$$E[Y|X] = a + bx$$

دیدیم که تابع چگالی شرطی متغیر تصادفی Y به ازای مقدار داده شده X برابر است با:

$$f_{Y|X}(y|x) = \frac{۱}{\sigma\sqrt{۲\pi}} e^{-\frac{۱}{۲\sigma^۲}(y-(a+bx))^۲} \qquad -\infty < y < \infty$$

۱۸-مثال ۳-۴

حال می‌خواهیم با فرض خطی بودن رابطه بین x و y فاصله اطمینان و آزمون فرض برای پارامتر $E[Y|X] = a + bx$ به ازای مقدار داده شده x بدست آوریم. همانطور که قبلاً اشاره شد $a + bx_p$ میانگین شرطی Y به شرط $X = x_p$ است و از $\hat{Y} = \hat{a} + \hat{b}x_p$ به عنوان یک برآورد برای $a + bx_p$ یاد کردیم. حال می‌خواهیم فرض $H_\circ: E[Y|X=x_p] = E_\circ$ را در مقابل فرض $H_۱: E[Y|X=x_p] \neq E_\circ$ ازمون می‌کنیم می‌توانیم با توجه به تابع توزیع $\hat{Y}$ تست آماری زیر را برای آزمون فرض فوق معرفی نماییم.

$$T=\frac{\hat{Y}-E_{\circ}}{S_{\hat{Y}}}=\frac{\hat{Y}-E_{\circ}}{\sqrt{S^{۲}\left[\frac{۱}{n}+\frac{(x_p-\bar{x})^{۲}}{S_{xx}}\right]}}\sim t(n-۲)$$

متغیر تصادفی T دارای توزیع t با $(n-۲)$ درجه آزادی می‌باشد. بنابراین در سطح معنی‌دار α فرض $H_{\circ}$ زمانی رو می‌شود که $|T|>t_{(۱-\frac{\alpha}{۲})}^{(n-۲)}$ باشد. بطور مشابه می‌توان فرض $H_{\circ}$ داده شده را در مقابل فرض‌های $H_۱: E\left[Y|X=x_p\right]>E_{\circ}$ یا $H_۱: E\left[Y|X=x_p\right]<E_{\circ}$ آزمون نمود که در این صورت فرض $H_{\circ}$ به ترتیب زمانی رد خواهد شد که داسته باشیم:

$$T<-t_{(۱-\frac{\alpha}{۲})}^{(n-۲)} \quad , \quad T>-t_{(۱-\frac{\alpha}{۲})}^{(n-۲)}$$

از تست آماری فرض شده می‌توان برای پیدا کردن فاصله اطمینان برای $E\left[Y|X=x_p\right]$ نیز استفاده نمود.

$$P(\hat{Y}-t_{(۱-\frac{\alpha}{۲})}^{(n-۲)}\sqrt{S^{۲}\left[\frac{۱}{n}+\frac{(x_p-\bar{x})^{۲}}{S_{xx}}\right]}<a+bx_p(\hat{Y}+t_{(۱-\frac{\alpha}{۲})}^{(n-۲)})\sqrt{S^{۲}\left[\frac{۱}{n}+\frac{(x_p-\bar{x})^{۲}}{S_{xx}}\right]}$$

۱۹-مثال

مثال: از داده‌های مثال ۲ استفاده کرده و یک فاصله اطمینان با ضریب اطمینان ٪۹۵ برای $X=۱/۰$ پیدا کنید.

حل:

ابتدا $\hat{Y}$ را در نقطه $X_p=۱/۰$ تخمین می‌زنیم:

$$\hat{Y}=\hat{a}+\hat{b}x_p$$

در مثالهای قبل $\hat{a}=۴۶/۴۹$ و $\hat{b}=۵۲/۵۷$ را بدست آوردیم.

پس:

$$\hat{Y}=۴۶/۴۹+(۵۲/۵۷)(۱/۰)=۹۹/۰۶$$

فرمول مناسب برای فاصله اطمینان برابر بود با:

$$\hat{Y}\pm t_{(۱-\frac{\alpha}{۲})}^{(n-۲)}\sqrt{S^{۲}\left[\frac{۱}{n}+\frac{(x_p-\bar{x})^{۲}}{S_{xx}}\right]}$$

با جایگذاری مقادیر مناسب در عبارت فوق که قبلاض محاسبه شده است خواهیم داشت:

$$۹۹/۰۶\pm(۲/۳۰۶)\sqrt{۴۶/۷۵\left[\frac{(۱/۰-۰/۹۴)}{۰/۴۴}\right]}$$

بعد از محاسبات انجام شده حاصل می‌شود:

$$۹۹/۰۶\pm ۵/۱۸ \qquad \Rightarrow \qquad [۹۳/۸۸\ ,\ ۱۰/۲۴]$$

می‌توان از محاسبات فوق نتیجه گرفت که برای هر واحد هزینه تبلیغات $(۱۰/۰۰۰\ \$)$ متوسط فروش ماهانه با احتمال ٪۹۵ در فاصله ۹۳۸۸۰۰ \$ و ۱۰۴۲۴۰۰ \$ خواهد بود. در شکل زیر برای $E\left[Y|X\right]$ رسم شده است.

۲۰- Zarebe

ضریب همبستگی

در بسیاری از اوقات نیاز به شاخصی داریم که چگونگی ارتباط بین دو متغیر x و y را اندازه بگیرد. این شاخص ضریب همبستگی خطی بین دو متغیر x و y نامیده می‌شود.

ضریب همبستگی خطی نمونه‌ای برآورد نامناسب برای ضریب همبستگی خطی تعریف شده در فصل‌های قبل است این معیار میزان قوت ارتباط خطی بین متغیرهای x و y را اندازه می‌گیرد و اولین بار دانشمند معروفی انگلیسی به نام کارل پیرسون آنرا معرفی نمود از این جهت به آن ضریب همبستگی خطی پیرسون نیز می‌گویطم و به فرم زیر تعریف می‌شود:

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$