به نام حضرت دوست که هرچه داریم از اوست

آمار استنباطی

استاد مربوطه: جناب آقای دکتر رستگار

تهیه و تنظیم: معصومه روا

سال تحصیل ۱۳۹۳

بسم الله الرحمن الرحيم

فهرست مطالب

بخش اول

بخش دوم

بخش سوم

بخش چهارم

جلسه پنجم

جلسه ششم

جلسه هفتم

جلسه هشتم

بخش نهم

جلسه اول

۹۲/۱۲/۱۵

آمار: به مجموعه ای از اصول و وش های برای تلخیص – طبقه بندی – سازمان دادن و تعمیم نتایج به کار می رود.

۲ شاخه در آمار وجود دارد

۱- **آمار توصیفی** : سازمان دادن – طبقه بندی – خلاصه

۲- **آمار استنباطی** : تعمیم نتایج

فواید آمار : کاربرد پژوهشی

بحث اندازه گیری : اندازه گیری توصیفی کمی یک پدیده یا فرآیند نسبت دادن اعداد و ارقام به جنبه های مختلف زندگی

خطاهای اندازه گیری

۱- خطای فردی، که اندازه گیری می کند (نداشتن دقت کافی – ناشی بودن فرد اندازه گیر)

۲- ابزارهای اندازه گیری (خطاهای ناشی از اندازه گیری – برای خلاقیت ابزار من نداشته باشیم)

در علوم انسانی از حدود واقعی استفاده می کنیم به عنوان مثال ۱۵/۵-۱۵-۱۴/۵

مهم : مقیاس ها یا سطوح اندازه گیری

نکته : از مقیاس اسمی عمل ریاضی انجام نمی شود

۱- **مقیاس اسمی** : مقیاس هایی هستند که برای شناسایی – طبقه بندی – نامگذاری استفاده می شود . معنای عددی ندارد (کمی نیستند)

مثال : جنسیت – رشته های تحصیلی – شماره پشت پیراهن ورزشکار – شغل – نژاد – شماره دانشجویی

۲- **مقیاس رتبه ای (ترتیبی)** : برای تعیین جایگاه یک فرد یا یک شئی در یک سلسله مراتب

مثال : آقای x از لحاظ خلاقیت پنجم است

نکته : اعمال ریاضی انجام نمی شود (به نوعی از مقیاس عددی استفاده می شود ولی فاصله ها دقیقا مشخص نیستند)

مقیاس اسمی برای مقایسه به کار نمی رود

مقیاس ترتیبی می توان تنها برای مقایسه به کار برد.

۱- ۳- **مقیاس فاصله ای** : مقیاس هایی هستند که فاصله بین سطوح آن شخص است و معنای عددی دارند دارای صفر قرار دادی هستند

مثال : دماهی هوا صفر درجه (صفر قراردادی)

۲- **مقیاس نسبی** : همه ویژگیهای مقیاس فاصله ای را دارند با این تفاوت که دارای صفر مطلق می باشد . مثال : موجودی من در بانک صفر است .

تقسیم بندی متغیرها : ۱- کمی : کم و زیاد بودن را مشخص می کند ← نمره هوش خلاق

۲-کیفی : معنای عددی ندارند ← جنسیت رنگها . وضعیت تاهل و مجرد

متغیرهای پیوسته : متغیرهایی هستند که پیوستار ارزش های مختلفی را به خود اختصاص می دهند . هر نقطه دارای معنای عددی خاص است .

متغیرهای گسسته : متغیرهایی هستند که فقط ارزشهای معین به خود اختصاص می دهد

×× تقسیم بندی های مستقل از هم هستند ← کمی گسسته درست نیست

- مهم ترین تقسیم بندی که مبنای همه کارهای پژوهشی است تقسیم بندی متغیرها به لحاظ نقشی است که در یک پژوهش به عهده دارد

انواع متغیرهایی که در پژوهش وجود دارند.

۱- متغیر مستقل : از هیچ تاثیر نمی پذیرد بلکه بر متغیر وابسته تاثیر گذار است یا به عبارتی در پژوهش ها آن را ابتکاری می کنیم .

۲- وابسته : تغییرات آن تحت کنترل متغیرهای دیگر است

۳- تعدیل کننده : متغیری است که رابطه بین متغیر مستقل و وابسته را تحت تاثیر قرار می دهد .

مثال : مقایسه اضطراب دانشجویان ریاضی و علوم تربیتی آزاد و پیام نور

۴- متغیر مزاحم : متغیری است که ما به دنبال حذف آن هستیم

هیچ وقت توی عنوان دنبال متغیر مزاحم نگردیم

۵- متغیر کنترل : رابطه بین خلاقیت و پیشرفت تحصیلی دانش آموزان دختر ← متغیر کنترل حذف نمی شوند .

مقایسه اضطراب دانشجویان روانشناسی و علوم تربیتی دانشگاه آزاد

اضطراب : متغیر وابسته دانشجویان روانشناسی : رشته تحصیلی متغیر مستقل دانشگاه آزاد : متغیر (کنترل)

کار محقق آن است که متغیرهای کنترل را بشناسد و متغیرهای فراهم را حذف کند

متغیر مزاحم یک چند پنهان است به تحقیق پژوهش وابسته است .

متغیرهای خصیصه ای و فعال

(هویتی – ذاتی) خصیصه ای : متغیرهایی هستند که تغییر آن در اختیار ما نیست ← مثل جنسیت- انتخاب رشته

فعال : در اختیار محقق می باشد ← اعمال نفوذ در آن هست ← دستکاری می شود (روش تدریس را استاد می تواند عوض کند)

آمار توصیفی : خلاصه کردن – طبقه بندی کردن – سازمان دادن

جدول توزیع فراوانی : طبقه بندی شده – طبقه بندی نشد

ایراد : درک آن زمان بر است

نمودار توزیع فراوانی ← بیان هندسی یا تصوری جدول توزیع فراوانی است

شاخص های مرکزی : شاخص هایی هستند که گرایش به مرکز توزیع را نشان می دهند

۳- میانه : نقطه ای است که توزیع را به دو قسمت مساوی تقسیم می کند (نقطه ثقل تعادل) الاکلنگ . مقایس رتبه ای در میانه کاربرد دارد ۵۰ درصد بالا و ۵۰ درصد پائین

۴- میانگین : برای مقیاس فاصله ای و نسبی کاربرد دارد . معدل مجموعه اعداد تقسیم بر حجم اعداد

انواع میانگین : ۱- هندسی ۲- همساز ۳- حسابی ۴- میانگین مرکب

۵- مد (نما) : مقیاس اندازه گیری متغیر مورد نظر اسمی باشد از مد استفاده می کنیم داده ای که بیشترین فراوانی را دارد . مد با مقدار مد از یکدیگر مستقل هستند .

- بی ثبات ترین و ضعیف ترین شاخص مرکزی مد یا همان نماست ، چون مقیاس اسمب است و هیچ عمل ریاضی با آن انجام نمی شود .
- با ثبات ترین شاخص مرکزی میانگین است .
- شاخص مرکزی به ما کمک می کند که گزارش نتایج ها را بررسی کنیم

لزوم استفاده از شاخص های پراکندگی :

میانگین به تنهایی وضعیت دانشجویان را کفایت نمی کند .

شاخص های پراکندگی

۱- دامنه تغییرات : زمانی استفاده می کنیم که مقیاس اندازه گیری فاصله ای نسبی باشد .

ویژگی محاسبه آن ساده است

معایب :

- در محاسبه نقطه از دو عدد (ک و ب) استفاده می شود
- در بعضی مواقع با حذف یک عدد مقدار آن به شدت تغییر می کند .

۲- انحراف متوسط : زمانی استفاده می شود که مقیاس اندازه گیری مورد نظر فاصله ای و نسبی باشد میانگین مجموع قدر مطلق انحراف از میانگین توزیع که به آن انحراف متوسط می گویند .

عیب : نمی توان از ارقام ریاضی استفاده کرد به خاطر وجود قدر مطلق

+ و — نادیده فاصله واقعی بین ارقام به حساب نمی آید

۳- واریانس (پراش) : زمانی از واریانس به عنوان شاخص پراکندگی استفاده می کنیم که مقیاس اندازه گیری مورد نظر فاصله ای و نسبی باشد

تعریف : میانگین مجموع مجذورات انحراف از میانگین توزیعی است

نکته : اگر نمونه بود $N = n - 1$ $\delta^2 = \frac{\sum(x-\bar{x})^2}{N}$

عیب واریانس : زمانی که فاصله اعداد با میانگین را به توان ۲ می رسانیم واحد اندازه گیری تغییر می کند مثال $m \leftarrow m^2$ تبدیل می شود .

(واحد اندازه گیری به کار رفته در محاسبات به دلیل مجذور شدن با واحد اندازه گیری واقعی متفاوت می شود)

انحراف معیار (انحراف استاندارد)

زمانی از این شاخص استفاده می کنیم که مقیاس اندازه گیری متغیر مورد نظر ما فاصله ای و نسبی باشد

تعریف : ریشه دوم میانگین مجموع مجذور انحراف از میانگین

$$\delta = \sqrt{\sum \delta^2} = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

- با ثبات ترین شاخص پراکندگی انحراف معیار (انحراف استاندارد) است .

انحراف چارکی : زمانی از انحراف چارکی به عنوان شاخص پراکندگی استفاده می کنیم که مقیاس اندازه گیری متغیر مورد نظر ما رتبه ای باشد .

تعریف چارک ها : نقاطی بر روی توزیع هستند که توزیع را به ۴ قسمت مساوی تقسیم می کند .

چارک اول : نقطه ۰/۲۵ چارک دوم : نقطه ۰/۵۰ چارک سوم ۰/۷۵

تاثیر انجام اعمال ریاضی با اعداد توزیع بر مقدار میانگین – واریانس – انحراف استاندارد

- هر عمل ریاضی که با اعداد توزیع شود با میانگین نیز انجام می شود .
- جمع و تفریق با اعداد توزیع بر مقدار واریانس تاثیری ندارد اما اگر تمامی اعداد یک توزیع را در یک عدد ثابت ضرب یا تقسیم کنیم واریانس در توان دوم آن عدد ضرب یا تقسیم می شود .
- توزیع $\times k \rightarrow \delta^2 + k^2$
- اگر تمامی اعداد یک توزیع با اعداد ثابتی جمع یا تفریق شود انحراف استاندارد تغییری نمی کند ولی اگر تمامی اعداد توزیع در عدد ثابتی ضرب یا تقسیم شود انحراف استاندارد نیز در همان عدد ضرب یا تقسیم می شود

مثال : تمامی اعداد یک توزیع با میانگین ۴ و انحراف استاندارد ۲ را در عدد ثابت ۳ ضرب نموده ایم میانگین و انحراف استاندارد توزیع جدید را محاسبه کنید .

$\bar{x} = 4 \qquad \delta = 2 \quad \delta^2 = 4$

$\bar{x} = 4 \times 3 = 12 \qquad \delta^2 = 9 \times 4 = 36 \qquad \delta = 6$

مثال : اگر میانگین یک توزیع به صورت x_1وx_2و ... ... وx_{10} عدد ۱۵ باشد میانگین توزیع جدید x_1و ... ... و$x_{10}, 14, 15$ را محاسبه کنید .

$\bar{x}_{10} = \frac{150}{10} = 15 \qquad \bar{x}_{12} = \frac{150+14+15}{12}$

نمرات درصدی :

نمراتی هسند که جایگاه یا موقعیت یک فرد یا یک عدد را در یک توزیع با توجه به مقیاس ۱۰۰ یا درصد مشخص می کند .

الف) رتبه درصدی (P_R)

ب) نقطه درصدی (P_X)

نمرات استاندارد : نمراتی هستند که جایگاه یک فرد یا یک عدد را در توزیع بر اساس دو پارامتر میانگین و انحراف استاندارد تعیین می کند

$(\bar{X} = 0 , \delta = 1)$

انواع مختلف نمرات استاندارد

۱- **نمره استاندارد** Z **:** در همه حال و شرایط میانگین آن صفر و انحراف استاندارد ۱ است .

مثال : در توزیع با $\bar{X} = 15$ و $\delta^2 = 9$ جایگاه فردی که نمره ۱۲ کسب نموده است را بر اساس نمره استاندارد تعیین کنید .

$$Z = \frac{X - \bar{X}}{s} = \frac{12 - 15}{3} = -1$$

(یک واحد از میانگین استاندارد پائین تر است)

۲- نمره استاندارد T :

$$T = \delta z + \bar{x} \qquad T = 10z + 50 \qquad \leftarrow \delta = 10 \ , \bar{x} = 50$$

۳- نمره استاندارد چند گانه :

عیب : تفاوت بین نمرات درصدی بیان کننده تفاوت بین نمرات استاندارد نیست .

منـــابع : آمـار توصیفی و استنباطی دکتر علی دلاور . دکتر هـومن – دکتـر پاشا شریفی – دکتـر کیامنش

جلسه دوم

۹۲/۱۲/۲۳

توزیع نرمال طبیعی : اگر به عنوان مثال پیشرفت تحصیلی تعداد زیادی از دانشجویان را اندازه گیری کرده و نمودار توزیع فراوانی را برای ان نمرات ترسیم نمائیم خواهیم دید که شکل نمودار (چند ضلعی) به یک منحنی متقارن و زنگوله ای شکل شباهت دارد . به این نوع توزیع ، توزیع طبیعی می گوئیم .

شکل توزیع طبیعی تابع دو پارامتر است

۱- میانگین $\bar{x}$

۲- انحراف استاندارد δ

با توجه به اینکه شکل توزیع طبیعی تابع دو پارامتر میانگین و انحراف استاندارد می باشد و مقدار میانگین و انحراف استاندارد نیز برای نمونه های مختلف و موقعیت های مختلف متفاوت است احتمالا ما با خانواده ای از توزیع های طبیعی سر و کار خواهیم داشت حال برای اینکه بودن توجه به مقدار این دو پارامتر برای تمام موقعیت ها و متغیرها دارای یک توزیع واحد باشیم می توانیم آنها را به نمرات Z تبدیل کرده تا به این ترتیب یک توزیع استاندارد با میانگین ۵ و انحراف استاندارد ۱ می باشد . $Z = \frac{x-\bar{x}}{5}$

ویژگی های توزیع طبیعی استاندارد :

۱- شاخص های گرایش مرکزی در توزیع طبیعی استاندارد بر روی یکدیگر قرار می گیرند

نکته دوم : اینکه دو سوی توزیع هیچ وقت محور افقی را قطع نخواهد گرد.

در دو سوی میانگین نرمال دارای دو نقطه عطف (اوج) می باشد

بیشترین ارتفاع توزیع در مرکز توزیع می باشد .

سطوح زیر توزیع نرمال تابع احتمالات بوده و بر اساس درصد بیان می شود و در فواصل مختلف همیشه ثابت است .

برای تمام متغیرها $34/14$ در $[0,1]$ قرار دارد .

$13/59$در $[1,2]$ قرار دارد

$2/14$در $[2,3]$ قرار دارد

جدول توزیع نرمال :

از آنجایی که توزیع طبیعی استاندارد یک توزیع احتمالاتی است بر اساس آن امکان تهیه جدول توزیع نرمال فراهم شده است .

ستون اول ⟵نمرات Z

ستون دوم ⟵ فاصله هر نمره Z تا میانگین

ستون سوم ⟵ سطح بزرگتر (یا سطحی که پایین تر از نمره Z قرار می گیرد).

ستون چهارم ⟵ سطح کوچکتر (یا سطحی که بالا تر از نمره Z قرار می گیرد).

ستون پنجم ⟵ ارتفاع توزیع را در هر نقطه نشان می دهد .

اگر شکل توزیع از حالت متقارن بودن تغییر کند چوگلی یا کجی پیش می آید .

$SK = \frac{\bar{X}-Mod}{s}$ فرمول کجی

اگر کجی نمودار بین ۱+ و ۱- باشد می توان گفت توزیع تقریبا نرمال است .

موارد استفاده از جدول توزیع نرمال

الف: پیدا کردن رتبه در صورتی معادل یک نمره خام

مثال : در توزیعی با $\bar{x} = 15$ و $\delta = 2$ رتبه درصدی دانشجویی که نمره ۱۳ را کسب نموده است را محاسبه کنید .

$$Z = \frac{x-\bar{x}}{s} = \frac{13-15}{2} = -1$$

رتبه دانشجو از ۱۵۸۷٪ دانشجویان دیگر بالاتر است (نوبت ۱۶٪)

ب : مراجعه به جدول توزیع Σ و پیدا کردن Z مورد نظر

اگر Z مثبت ستون سوم

اگر Z منفی ستون چهارم

ج : پیدا کردن نمره خام معادل یک رتبه درصدی

د: پیدا کردن سطحی از توزیع که بین دو نمره خام یااستاندارد قرار می گیرند .

مثال : در توزیع با $\bar{x} = 16$ و $S = 3$ درصد افرادی را که نمراتی بین ۱۴ و ۱۸

$$\begin{cases} z_1 = -\frac{2}{3} \simeq -0/66 \rightarrow \text{جدول در ستون دوم} \rightarrow 0/2454 \\ z_2 = \frac{2}{3} \simeq 0/66 \rightarrow \text{جدول در ستون دوم} \rightarrow 0/2454 \end{cases}$$

نکته : اگر z_1 و z_2 هم جهت نبودند با هم جمع می کنیم ۴۹٪ افراد نمراتی بین ۱۴ و ۱۸ کسب نموده اند.

نکته : اگر z_1 و z_2 هم جهت بودند سطح بزرگتر − سطح کوچکتر

تعریف همبستگی : هرگاه بین دو یا چند متغیر رابطه وجود داشته باشد می گوئیم بین آنها همبستگی وجود دارد به عبارت دیگر هم تغییری یا هماهنگی تغییرات دو یا چند متغیره با یکدیگر را همبستگی می نامیم .

ضریب همبستگی : بیان مقداری یا عددی رابطه بین متغیرها را ضریب همبستگی می گوئیم .

مثال : ضریب همبستگی بین هوش و پیشرفت تحصیلی ۰/۶ است .

نمی توان بین دو متغیر استنباط علی و معنوی وجود ندارد هوش بالا منجر به پیشرفت نمی شود .

شدت همبستگی : به لحاظ شدت دامنه ضریب همبستگی ۰ تا ۱ است

جهت همبستگی : به لحاظ جهت دامنه ضریب همبستگی بین ۱+ تا ۱- است .

برای نشان دادن همبستگی بین دو متغیر یا چند متغیر با استفاده از نمودار پراکندگی استفاده می کنیم .

مثال : فرض کنید با دو متغیر X و y

نمودار پراکندگی

انواع ضریب همبستگی : ضریب همبستگی پیرسون – ضریب همبستگی اسپیرمن – ضریب همبستگی فای $\emptyset$ - ضریب همبستگی پای سرنان (r_b) – ضریب همبستگی یونیب بانی سرنال (r_{pb})- ضریب همبستگی انا ($\$$)← زمانی استفاده می شود که رابط بین دو متغیر خطی نباشد .

ضریب همبستگی پیرسون :

زمانی از ضریب همبستگی پیرسون استفاده می کنیم که مقیاس اندازه گیری متغیرهای مورد نظر ما فاصله ای و نسبی باشد .

مفروضه های ستفاده از ضریب همبستگی پیرسون :

۱. خطی بودن
۲. نرمال بودن (توزیع متغیرهای مورد نظر نرمال باشد)
۳. همگنی واریانس (پراکندگی بین دو متغیر نزدیک به هم باشد)

مثال : اطلاعات زیر مربوط به نمرات هوش و پیشرفت تحصیلی تعدادی از دانشجویان می باشد ضریب همبستگی بین این نمرات را محاسبه نمائید .

xy	پیشرفت y	هوش x
۸	۴	۲
۸	۲	۴
۶۰	۱۰	۶
۶۴	۸	۸
۶۰	۶	۱۰
$\sum xy = 200$	$\sum 6 = 30$	$\sum X = 30$
$\sum 6^2 = 220$		$\sum X^2 = 220$

$$\delta_{XY}=\frac{N(\sum xy)-\sum x\sum y}{\sqrt{[N\sum x^2-(\sum x)^2][N\sum y^2-(\sum y)^2]}}$$

$$=\frac{5(200)-(30)(30)}{\sqrt{(5(220)-900)(5(9001-220)}}=\frac{1000-900}{\sqrt{(1100-900)(1100-900)}}$$

$$=\frac{100}{\sqrt{200\times 200}}=\frac{100}{\sqrt{(200)^2}}=\frac{100}{200}=\frac{1}{2}=0/5$$

ضریب تعیین R^2

$$R^2=r_{x6}^2\times 100=(0/5)^2\times 100=25$$

با استفاده از ضریب تعیین می توانیم مقداری از واریانس (پراکندگی یا تغییرات) متغیر ۶ را که توسط متغیر X تعیین می شود مشخص نمائیم به عبارتی ضریب تعیین بیانگر واریانس مشترک بین دو متغیر می باشد.

ضریب همبستگی اسپیرمن : زمانی از ضریب همبستگی اسپیرمن استفاده می کنیم که مقیاس اندازه گیری متغیر مورد نظر ما رتبه ای باشد .

مثال : مدیر گروه مدیریت آموزش به منظور تعیین منبع درسی آمار استنباطی از ۲ استاد خواست تا ۸ کتاب درسی آمار را رتبه بندی کند با توجه به اطلاعات به دست آمده ضریب همبستگی مناسب را محاسبه نمائد .

فرمول ضریب همبستگی اسپیرمن

$$1-\frac{6\sum(D^2)}{n(n^2-1)}=1-\frac{6(24)}{8(63)}=0/71$$

$n=$تعداد مشاهدات

$D=$تفاوت بین رتبه ها

۰/۷۱ یک نوع ضریب توافق بین دو استاد

کتاب	استاد A	استاد B	D	D^2
۱	۳	۶	-۳	۹
۲	۴	۴	۰	۰
۳	۶	۸	-۲	۴
۴	۸	۷	۱	۱
۵	۷	۵	۲	۴
۶	۱	۲	-۱	۱
۷	۲	۱	۱	۱
۸	۵	۳	۲	۴
				$\sum D^2 = 24$

چند نکته در مورد ضرایب همبستگی :

۱- تمامی متغیرهای مورد استفاده در مطالعات همبستگی می باست مربوط به هر فرد به صورت جداگانه ثبت گردد .

۲- از ضرایب همبستگی ضرایب علی و معلولی استنباط نمی شود

۳- با افزایش حجم نمونه همبستگی بین متغیرها تا حدودی افزایش پیدا می کند

۴- رابطه بین دو متغیر گاه ممکن است ناشی از همبستگی آنها با یک متغیر سوم باشد

مثل: روز بارانی الف) پالتو ب) چتر روز عادی الف) پالتو ب) چتر

رابطه بین چتر دست گرفتن با پالتو پوشیدن بستگی به متغیر سوم (باران) می باشد .

«مدیریت و پژوهش در ایران مثل حروف یرملون است نوشته می شود خوانده نمی شود»

جلسه سوم

۹۳/۱/۲۲

رگرسیون : هرگاه بین دو متغیر هم بستگی وجود داشته باشد می توانیم آن را به کمک دیگری پیش بینی نمائیم هر چه میزان همبستگی بیشتر باشد طبیعتا قدرت پیش بینی نیز افزایش می یابد .

تعریف خط رگرسیون :

۱- خطی است که با بهترین پردازش ممکن از میان نقاط نمودار پراکندگی عبور کرده به طوریکه کمترین فاصله ای ممکن را با آن نقاط دارا می باشد

۲- خطی است که بر اساس ملاک حداقل مجذورات خطا ترسیم می شود

۳- خطی است که بر اساس نقاط پیش بینی شده ترسیم می شود

مثال : اطلاعات زیر مربوط به ۲ متغیر خودکار آمدی تحصیلی و تعلل ورزی تحصیلی می باشد .

براساس این اطلاعات

الف: معادله پیش بینی تعلل ورزی تحصیلی را تشکیل دهید

ب: تک تک نمرات تعلل ورزی تحصیلی را بر اساس خودکار آمدی پیش بینی نمائید

ج: خط رگرسیون پیش بینی تعلل ورزی را بر اساس خودکار آمدی را ترسیم کنید

$$b = \frac{\sum xy - \sum x \sum y}{N(\sum y^2) - (\sum y)^2} \qquad a = \frac{\sum y - b(\sum x)}{N}$$

مقدار b به ما می گوید که با یک واحد افزایش یا کاهش در متغیر y ، متغیر X چه مقدار تغییر می کند .

y'	y^2	xy	تعلل ورزی y	خودکارآمدی X
۶/۵۶۶	۳۶	۱۲	۶	۲
۶/۲۸۲	۶۴	۳۲	۸	۴
۵/۹۹۸	۱۰۰	۶۰	۱۰	۶
۵/۷۱۴	۴	۱۶	۲	۸
۵/۴۳	۱۶	۴۰	۴	۱۰
	$\sum y^2 = 320$	$\sum xy = 160$	$\sum y = 30$	$\sum X = 30$

$$y_1' = 6/85 + (-0/142)(2) = 6/566$$

$$y'_2 = 6/85 + (-0/142)(4) = 6/282$$
$$y'_3 = 6/85 + (-0/142)(6) = 5/998$$
$$y'_4 = 6/85 + (-0/142)(8) = 5/714$$
$$y'_5 = 6/85 + (-0/142)(10) = 5/43$$

در این جدول برای کشیدن خط رگرسیون باید نقاطی را که به دست آورده ایم را پیدا کنیم و به هم وصل کنیم همیشه ما برای پیش بینی متغیر خود به مقدار خطا می کنیم .

خطای پیش بینی $e = \sum(y - y')^2$ حداقل مجذورات

خط رگرسیون : خطی است که کمترین خطای ممکن را مرتکب می شود .

خطای استاندارد برآورد : برای تعیین دقت پیش بینی نمرات x بر اساس نمرات y از شاخص به نام خطای استاندارد برآورد استفاده می کنیم و با S_{xy} نشان می دهیم .

$$S_{xy} = \sqrt{\frac{\Sigma(y-y\prime)^2}{N}} \quad \text{یا} \quad S_{xy} = S_x\sqrt{1-\left(r_{xy}\right)^2}$$

آمار استنباطی : گاهی اوقات برای انجام پژوهش ها ما امکان دسترسی به تمام اعضای جامعه آماری را نداشته و یا دسترسی به تمام اعضاء از لحاظ اقتصادی مقرون به صرفه نیست لذا جهت تجزیه و تحلیل نتایج چنین پژوهش هایی از شاخص های آمار استنباطی استفاده می کنیم .

کار آمار استنباطی تقسیم نتایج از نمونه حرکت از نمونه به سمت جامعه – جرکت از گروه کوچک به بزرگ

تعریف جامعه آماری : به گروهی از افراد و یا اشیا اطلاق می گردد که حداقل دارای یک ویژگی مشترک بیشتر باشد جامعه آماری محدود تر می شود (جامعه اماری دل به خواهی نیست)

نمونه آماری :

به تعدادی از افراد و اشیاء گفته می شود که در یک جامعه تعریف شده انتخاب می شود به طوریکه معرف و نماینده واقعی جامعه باشد .

برای اینکه نمونه آماری معرف یا نماینده واقعی جامعه باشد می بایست از روشهای نمونه گیری مناسب استفاده نمائیم .

روشهای احتمالاتی : شانس تمام افراد جامعه برای انتخاب به عنوان عضوی از نمونه از صفر بزرگتر می باشد .

روش نمونه ای تصادفی ساده : زمانی از این روش استفاده می کنیم که ۱- لیست قبلی کامل از افراد جامعه آماری را در اختیار داشته باشیم البته توجه داشته باشیم که در این لیست می بایست افراد بر اساس نظم خاصی چیدمان شده باشند ..

معمولا در روش نمونه ای تصادفی ساده حجم نمونه و جامعه کم می باشد .

۲- در روش نمونه ای تصادفی ساده واحد نمونه گیری فرد می باشد و اگر درست اجرا شود همه افراد جامعه دارای شانس مساوی و مستقل هستند .

روش نمونه ای منظم (سیستماتیک) : در این روش اعضای نمونه از یک لیست تهیه شده که بر اساس نظم معین نمی باشد انتخاب می کنیم در روش منظم افراد شانس مستقل برای انتخاب شدن ندارند .

روش نمونه گیری خوشه ای :

در این روش حجم جامعه آماری معمولا بزرگ بوده و در یک گسترده وسیع جغرافیایی افراد جامعه پراکنده می شوند لذا در این موارد از روش خوشه ای استفاده می کنیم در روش خوشه ای چون حجم بزرگ است واحد نمونه گیری گروه است دیگه یکی یکی انتخاب نمی کنیم مثلا دانشگاه مدرسه کلاس

روش نمونه گیری طبقه ای : زمانی که جامعه آماری پژوهش از لحاظ یک متغیر تاثیر گذار در نتایج پژوهش نامتناجنس باشد از این روش استفاده می کنیم توجه داشته باشیم که در این روش افراد می بایست با همان نسبتی که در جامعه حضور دارند در نمونه نیز حضور داشته باشند یعنی باید نسبت ها رعایت شوند مثلا ۸۰٪ دختر ۲۰٪ پسر نمونه باید $\frac{80}{20}$ رعایت شود .

روشهای نمونه گیری غیر احتمالاتی :

روشهایی هستند که در آن شانس برخی از افراد برای انتخاب شدن برابر با صفر می باشد مثل نمونه گیری داوطلبانه یا نمونه گیری در دسترس (برای سهولت کار افرادی که در دسترس هستند) نمونه گیری گلوله برفی (یک مساله نادر را بررسی می کنیم) هدفمند به سمت یک نفر می رویم .

این روشها ضعیف هستند و سعی کنیم که از این روشها انتخاب نکنیم ولی در مواردی به ناچار مجبوریم از این روشها استفاده کنیم مانند گلوله برفی .

اشتباه نمونه گیری :

از آنجائیکه حجم نمونه معمولا کوچکتر از جامعه است احتمال پارامتر برآورده شده با پارامتر جامعه بسیار کم می باشد این امر مربوط به عوامل شانس و تصادف است به همین دلیل معمولا نه تنها پارامتر برآورد شده با پارامتر جامعه برابر نیست بلکه مقذدار میانگین یا بین سایر پارامترهای مربوط به نمونه های مختلف که از یک جامعه انتخاب می شوند نیز تفاوت وجود دارد .

جامعه ⟵ حجم وسیع　　　　　　　　نمونه ⟵ حجم کوچک

مرتکب خطا ⟵ شانس و تصادف ⟵ نمونه برآورد با جامعه برابر نیست

$\bar{x} = 52$　　　　　　　　$\bar{x} = 50$

ویژگی های آماری جامعه را پارامتر گویند.

آمار ⟵ ویژگی های نمونه

میانگین مربوط به نمونه ⟵ آماری

میانگین مربوط به جامعه ⟵ پارامتر

به شاخص که برای برآورد پارامتر جامعه به کار می رود ⟵ برآورد کننده

به عنوان مثال : اگر از میانگین نمونه به عنوان میانگین جامعه استفاده کنیم بـه نمونـه ⟵ بـرآورد کننده در آمار اشتباه نمونه گیری را می توانیم به کمک قضیه حد مرکزی توضیح دهیم .

جلسه چهارم

۹۳/۱/۲۹

قضیه حد مرکزی :

اگر از یک جامعه با حجم N نمونه های متعددی با حجم n را انتخاب نمائیم و میانگین این نمونه را حساب کنیم خواهیم دید که توزیع میانگین این نمونه ها نرمال و طبیعی خواهد بود و میانگین میانگین ها برابر با میانین جامعه می باشد همچنین توزیع نمونه ای میانگین ها دارای انحراف استانداردی است که به آن خطای معیار میانگین وی گویند . $(\delta \bar{x})$

نه تنها میانگین میانگین نمونه ها با هم یکی است بلکه میانگین میانگین قضیه نمونه ها نیز متفاوت است .

سوال : در توزیع با این مشخصات اگر نمونه ای را انتخاب کنیم با احتمال ۰/۹۵ با چه فاصله ای این نمونه از میانگین جامعه قرار می گیرد .

$$\delta_{\bar{x}} = \frac{\delta}{\sqrt{n}}$$

سوال : میانگین پیشرفت ریاضی جامعه دانش آموزان در درس آمار برابر با ۲۰ و خطای استاندارد میانگین برابر با ۲ می باشد احتمال مشاهده یک میانگین نمونه برابر با بیشتر از ۲۶ و برابر یا کمتر از ۱۶ چقدر است .

(از قضیه حد مرکزی استفاده می کنیم)

$$Z_{\bar{x}} = \frac{x - \bar{x}}{s} = \frac{26 - 20}{2} = 3 \quad \rightarrow 0/0013$$

انحراف استاندارد (خطای معیار میانگین)

$$Z_{\bar{x}} = \frac{16 - 20}{2} = -2 \quad \rightarrow 0/0228$$

کمتر از ۱۶ بیشتر بود چون توزیع نرمال چون توزیع متقارن است دیگه نمرات منفی قید نشده است معادل همان ۲ را در نقطه در نظر می گیریم .

برآورد کننده : شاخص آماری که برای برآورد پارامتر جامعه مورد استفاده قرار می گیرد برآورد کننده نام دارد.

ویژگی :

۱- بدون اریب بودن (یک طرفه نبودن – غیر سودار)

یعنی اگر از یک جامعه نمونه های متعددی را انتخاب کنیم و میانگین شاخص آماری این نمونه ها را محاسبه نمائیم میانگین شاخص آماری در نمونه های مختلف باید برابر با پارامتر جامعه باشد به عنوان مثال میانگین میانگین ها برابر با میانگین جامعه با پارامتر جامعه باشد به عنوان مثال میانگین میانگین ها برابر با میانگین جامعه

تعریف شاخص بدون اریب : به عبارتی در شاخص آماری محاسبه شده گرایش منظمی جهت انحراف از پارامتر جامعه وجود نداشته ، باشد می گوئیم آن شاخص آماری غیر سودار است .

۲- ثبات براورد (سازگاری یا یکنواخت بودن)

اگر با افزایش حجم نمونه مقدار شاخص آماری برآورده شده به پارامتر جامعه نزدیک تر شود می گوئیم آن شاخص آماری یک برآورد کننده با ثبات است .

۱- کارایی : برآورده کننده ای دارای کارایی است که حداکثر دقت را برای برآورد پارامتر جامعه داشته باشد به عبارتی میزان دقت یک برآورده کننده برای برآورد پارامتر جامعه

۲- مکفی بودن : اگر واریانس نمونه گیری یک برآورد کننده مانند میانگین کوچکتر از شاخص باشد آن را یک برآورد کننده مکفی می گویند .

$$S_{md}^2 = \frac{1/57 S_x^2}{n}$$

مثلا: میانگین کوچکتر از واریانس نمونه گیری میانه باشد .

$$S_{\bar{x}}^2 = \frac{S_x^2}{n}$$

تعریف فرضیه : برای انجام یک پژوهش محقق می بایست پس از انتخاب موضوع و مطالعه پیشینه نظری و تجربی پژوهش ، گزاره های پژوهشی خود را به شکل فرضیه یا سوال بیان نماید.

فرضیه حدس یا گمان خردمندانه ای است در مورد روابط متغیرها ؛ تفاوت بین گروه ها و تاثیر متغیرها بر یکدیگر ؛ که به شکل یک جمله خبری نوشته می شود .

گاهی اوقات در پژوهش مصداق باید وجود ندارد به جای فرضیه از سوال استفاده می کنیم .

آیا معرف فلان دارو برترک اعتبار موثر است؟

فرضیه صفر در مورد میانگین

$$H_0 = \bar{x}_1 = \bar{x}_2$$
$$H_1 = \bar{x}_1 \neq \bar{x}_2$$

ما معمولا فرضیه ها را از ۲ زاویه می توانیم تقسیم کنیم : الف) فرضیه صفر ب) فرضیه ۱ (حذف – مقابل- تحقیق- بدیل)

فرضیه صفر : بین x و y رابطه وجود ندارد – کلاس A و B با هم تفاوتی ندارد . در پژوهش فرضیه را با H_0 نشان می دهیم .

فرضیه خلاف : مقابل فرضیه صفر قرار دارد . در پژوهش در فصل ۱ و ۲ فرضیه حذف یا مقابل را می نویسم .

ولی فرضیه صفر اون فرضیه ای است که آزمون اماری با آن انجام می شود. اگر فرضیه صفر رد شود خلاف را در نظر می گیریم .

$$\begin{cases} H_0 = \bar{x}_1 = \bar{x}_2 \\ H_1 = \bar{x}_1 \neq \bar{x}_2 \end{cases}$$

فرضیه جهت دار – فرضیه بدون جهت

بین کلاس A و B تفاوت وجود دارد ← فرضیه بدون جهت

میانگین A بیشتر از کلاس B هست ← فرضیه جهت دار

بین x و y رابطه وجود دارد ← بی جهت

بین x و y رابطه منفی وجود دارد ← جهت دار

$$\begin{cases} H_0 = \bar{x}_1 = \bar{x}_2 \\ H_1 = \bar{x}_1 > \bar{x}_2 \end{cases}$$ آزمونهای دو متغیره جهت دار

$$\begin{cases} H_1 : \bar{x} = 50 \\ H_1 = \bar{x} \neq 50 \end{cases}$$ یک متغیره

$$\begin{cases} H_0 = \bar{x}_1 = 50 \\ H_1 : \bar{x}_1 > 50 \end{cases}$$ **حالت تک متغیره** **جهت دار**

آزمون آماری در دامنه : اگر پژوهشگر نتواند بر مبنای پیشینه نظریه تجربی حدس به خردانه ای را تنظم کند معمولا یا از فرضیه بدون جهت استفاده کرده و یا گزاره تحقیق خود را به شکل سوال بیان می کند در چنین مواردی از آزمونهای آماری دو دامنه استفاده می شود .

$$\begin{cases} H_0 = \mu_1 = \mu_2 \\ H_1 = \mu_1 \neq \mu_2 \end{cases}$$ دو متغیره

$$\begin{cases} H_0 = \mu_0 = 0 \\ H_1 = \mu_0 \neq 0 \end{cases}$$ یک متغیره

آزمون آماری یک دامنه : در این آزمون محقق به مبنای ... نظریه تجربی در مورد جهت اختلاف یا جهت تاثیر یا رابط متغیرها با یکدیگر تصمیم گیری می کند . یعنی جهت مشخص می شود .

تصمیم گیری در مورد رد یا تاثیر فرضیه صفر

ما قالبا در قبلا فرضیه صفر دو بار تصمیم می گیریم .

۳- تصمیم گیری اول در مورد صحیح یا غیر صحیح بودن فرضیه صفر

۴- در مورد تایید یا رد فرضیه صفر می باشد .

نوع تصمیم ما در این دو موقعیت را می توانیم به کمک یک ماتریس نمایش دهیم .

بین ۰/۰۱ تا ۰/۰۵ فرضیه صفر را رد می کنی یعنی رابطه وجود دارد ، فرضیه خلاف را می پـذیری

آزمون های آماری وقتی گزارش می دهند .

سطح قابل قبول در پژوهش علوم انسانی ۰/۰۱ تا ۰/۰۵ است اگر ۰/۰۱ کارت نشد ناچاری ۰/۰۵ .

جلسه پنجم

۹۳/۲/۵

آزمون های آماری

درجه آزادی : به تعداد ارزشهایی اشاره دارد که پس از قرار دادن برخی محدودیت در داده ها آزادانه تغییر می کند .

دلیل استفاده از درجه آزادی : (df)

به خاطر کاهش خطای نمونه گیری و برآورد دقیق تری از واریانس یا پراکندگی متغیر مورد نظر در جامعه داشته باشیم . به تعداد اطلاعات مستقلی که بعد از قرار دادن برخی محدودیت در داده ها نمونه آماری بر اساس آن شکل می گیرد درجه آزادی گویند .

$$df = n - 1$$

توزیع نمونه ای z

اگر از یک جامعه ای با توزیع نرمال ، نمونه های مختلف انتهایی را انتخاب کنیم و مشخصه آماری (z) تک نمونه ای را محاسبه نمائیم و آنگاه نمودار توزیع فراوانی را ترسیم کنیم خواهیم دید که توزیع حاصل ، یک توزیع نرمال بوده و دارای انحراف استانداردی است که به آن خطای معیار میانگین می گوئیم ضمنا میانگین این توزیع برابر میانگین جامعه می باشد .

$$z = \frac{\bar{x} - \mu}{\delta_{\bar{x}}}$$

آزمون z تک متغره :

در مواقعی استفاده می شود که توزیع متغیر مورد نظر در جامعه نرمال باشد و همچنین نیز ما به انحراف استاندارد جامعه دسترسی داشته باشیم در ضمن مقیاس اندازه گیری متغیر مورد بررسی می بایست حداقل فاصله ای باشد با استفاده از این آزمون می توانیم ، میانگین یک نمونه ی انتخابی را با میانگین جامعه مقایسه نمائیم .

مثال : میانگین و انحراف استاندارد نمرات عزت نفس دانشجویان دختر دانشگاه آزاد به ترتیب برابر $\mu = 7/87$ و انحراف استاندارد $\delta = 1/1$ رئیس دانشگاه آزاد مرکز شیراز بر اساس مستنداتی که در اختیار دارد ، ادعا می کند که عزت نفس دانشجویان دختر این دانشگاه از

عزت نفس دانشجویان دختر دانشگاه های آزاد کشور متفاوت است با استفاده از اطلاعـات گـرد آوری شده زیر فرضیه مورد نظر رئیس دانشگاه را در سطح ۰/۰۵ آزمون نمائید .

گام اول : انتخاب آزمون آماری مناسب

متغیر : عزت نفس

چون می خواهیم میانگین نمونه با جامعه سنجیده شود و مقیاس اندازه گیری فاصله ای است .

گام دوم : بررسی مفروضه ها

معمولا حجم نمونه از ۳۰ تا بالاتر می رود حجم نمونه نرمال است

دوتا آزمون آماری داریم که یکی از آنها پرکارتر است (آزمون KSS) گالرگنواسمیرنوف (کجی و کشیدگی)

گام سوم : بیان فرضیه هاست

فرضیه صفر در این سوال $H_0 = \mu = 7/87$

فرضیه صفر می گوید برابر می باشد $H_1 = \mu \neq 7/87$

گام چهارم : تعیین معیار تصمیم گیری

فرضیه صفر را رد می کنیم اگر z مشاهده شده بزرگتر یا مساوی z بحرانی باشد . $z_{م} \geq z_{ب}$

فرضیه صفر را می پذیریم اگر z مشاهده شده کوچکتر از z بحرانی باشد . $z_{م} < z_{ب}$

$$\text{فرمول اصل}\begin{cases} 0/01 = 2/33 & \text{یک دامنه} \\ 0/05 = 1/65 & \text{یک دامنه} \end{cases}$$

$$Z\begin{cases} 2/58 & \text{دو دامنه} \\ 1/96 & \text{دو دامنه} \end{cases}$$

اگر از ۱/۹۶ حتی یک صدم پایین تر باشد فرضیه صفر تایید می شود .

گام پنجم : محاسبه مشخصه آماری آزمون z :

$$z_{\circ} = \frac{\bar{x} - \ell}{\delta_{\bar{x}}} = \frac{8/2 - 7/87}{0/11} = \frac{0/33}{0/11} = 3$$

$$\delta_{\bar{x}} = \frac{\delta}{\sqrt{v}} = \frac{1/1}{\sqrt{100}} = 0/11$$

$$z_{\circ} > z_{ب} \quad = 3 > 1/96$$

با توجه به اینکه z مشاهده شده از z بحرانی بزرگتر می باشد ، نتیجه می گیریم عزت نفس دانشجویان آزاد شیراز ، نسبت به دانشجویان آزاد کشور متفاوت است .

و با توجه به میانگین به دست آمده عزت نفس دانشجویان دختر شیراز از دانشجویان دختر آزاد کشور بیشتر است

برآورد فاصله اطمینان برای توزیع z تک متغیره

برآورد فاصله ای : به فاصله ای بر روی خط اعداد اشاره می کند که پارامتر جامعه در محدوده این فاصله قرار می گیرد .

برآورد نقطه ای : زمانی است که ما می خواهیم با یک یک ارزش پارامتر جامعه را دقیقا برآورد کنیم .

برآورد فاصله ای :

$$\bar{X} - \frac{Z_1 - \propto}{2} \quad (\delta_{\bar{x}}) \leq \mu \leq$$

$$\bar{X} + \frac{Z_1 - \propto}{2} \quad (\delta_{\bar{x}} \leq \mu \leq)$$

رئیس دانشگاه آزاد شیراز معتقد است عزت نفس دانشجویان دختر این دانشگاه از دانشجویان دانشگاه آزاد کشور بیشتر است با استفاده از اطلاعات زیر فاصله اطمینان ۰/۹۵ را برای برآورد میانگین جامعه انجام دهید .

$\sum 1 = 1/1$ میانگین نمونه $\bar{x} = 8/2$ $\alpha = 0/5$

$\mu = 7/8$ د حجم نمونه $= 100$ $z_1 = 1/96$

$$CI = 8/2 - 1/96(0/11) \leq \mu \leq 8/2 + 1/96(0/11)$$

$$\delta_{\bar{x}} = \frac{1/1}{\sqrt{n}} = 0/11$$

$CI = 7/98 \leq \mu \leq 8/42$ میانگین جامعه بین این دو عدد نیست

توزیع نمونه گیری z برای تفاوت بین میانگین ها : $(\bar{x}_1 - \bar{x}_2)$

اگر از یک جامعه ای با توزیع نرمال ، بارها و بارها ، دو نمونه مختلف زا انتخاب کنیم و تفاوت بین میانگین های نمونه های دو تایی را به دست آوریم و مشخصه آماری z برای تفاوت بین میانگین ها را محاسبه کنیم و سپس نمونه توزیع فراوانی را ترسیم کنیم ، خواهیم دید که شکل نمودار توزیع نرمال بوده و میانگین تفاوت بین میانگین ها برابر با تفاوت میانگین ها در جامعه می باشد همچنین این توزیع دارای انحراف استانداردی خواهد بود که به آن خطای معیار میانگین برای تفاوت میانگین ها گفته می شود

$$z_{,} = \frac{\bar{x}_1 - \bar{x}_2}{\delta_{\bar{x}_1 - \bar{x}_2}}$$

فرمول خطای معیار میانگین برای تفاوت میانگین ها :

$$\delta_{\bar{x}_1 - \bar{x}_2} = \sqrt{\delta^2_{\bar{x}_1} + \delta^2_{\bar{x}_2}}$$

آزمون z برای گروه های مستقل

زمانی از آزمون z برای گروه های مستقل استفاده می کنیم که قصد مقایسه دو گروه را در یک متغیر با مقاس فاصله ای و نسبی را داشته باشیم ضمنا در این شرایط انحراف استاندارد جامعه نیز برای هر دو گروه باید معلوم باشد

مفروضه مستقل بودن : با نمونه گیری حل می شود انتخاب ۱۰۰ نفر از بین همه به صورت تصادفی یعنی شانس هم مساوی بوده است

مفروضه نرمال بودن : زمانی که حجم از ۳۰ بالا می رود ، حجم نرمال است مفروضه های استفاده از z دو متغیره .

۱- مستقل بودن : (مستقل بودن افراد مستقل بودن گروه ها) با استفاده از روش نمونه گیری صحیح

۲- نرمال بودن : اگر حجم نمونه از ۳۰ بالاتر باشد حجم نرمال است اگر بخواهیم به جهت تجربی بررسی کنیم از کجی یا کشیدگی ۲+ و ۲- و یا (KSS) استفاده کنیم .

مثال : پژوهشگری معتقد است که خودکار آمدی تحصیلی دانشجویان رشته های علوم انسانی نسبت به دانشجویان علوم پایه متفاوت است . با توجه به اطلاعات گردآوری شده فرضیه مورد نظر این پژوهشگر را در سطح ۰/۵ آزمون نمائید.

خودکارآمدی : (نسبت به توانایی تحصیلی خود چقدر ایمان دارید)

گرو۲ پسر	گروه ۱ دختر
$V_2 = 36$	$v_1 = 26$
$\bar{x} = 14$	$\bar{x}_1 = 16$
$\delta = 4$	$S_n = 4$

گام اول : انتخاب آزمون : استفاده از z

گام دوم : بررسی مفروضه ، چون حجم نمونه بالاتر از ۳۰ است (۳۶ است) نرمال است

گام سوم : بیان فرضیات (علوم پایه و انسانی با هم تفاوتی ندارند)

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

گام چهارم : تعیین ملاک تصمیم گیری

$IF = z_{م} \geq z_{ب} \Rightarrow H_0$ رد می شود

$if\ z_{م} < z_{ب} \Rightarrow H_0$ تایید می شود

$$z_{م} = \frac{\bar{x}_1 - \bar{x}_2}{\delta_{\bar{x}_1 - \bar{x}_2}}$$

$$\delta_{\bar{x}_1 - \bar{x}_2} = \sqrt{\delta_{\bar{x}_1} + \delta^2_{\bar{x}_2}}$$

$$\delta_{\bar{x}_1} = \frac{\delta}{\sqrt{n_1}} = \frac{4}{\sqrt{36}} = \frac{4}{6} = 0/66$$

$$\delta_{\bar{x}_2} = \frac{4}{\sqrt{36}} = \frac{4}{6} = 0/66$$

گام آخر تصمیم گیری :

$Z_{ب} = 1/96$ $\quad Z_{م} = 2/16$ مشاهده شده

چون Z مشاهده از Z بحرانی بزرگتر است فرضیه صفر رد می شود در سطح ۰/۰۵ فرضیه H_0 رد می شود . $2/16 > 1/96$

اگر در سطح ۰/۰۱ باشد $2/16 < 2/58$ فرضیه صفر تاثیر می شود .

- تشکیل فاصله اطمینان یا برآورد فاصله ای برای تفاوت بین میانگین ها در مواردی که انحراف استاندارد جامعه معلوم است .

$$C.I: (\bar{x}_1 - \bar{x}_2) - \left(z_1 - \frac{\propto}{2}\right)\left(\delta_{\bar{x}_1 - \bar{x}_2}\right) \leq \mu_1 - \mu_2 \leq$$

اطلاعات زیر مربوط به خودکار آمی تحصیلی دانشجویان رشته های علوم انسانی و علوم پایه می باشد فاصله اطمینان ۰/۹۵ را برای برآورد تفاوت بین این ۲ گروه در جامعه تشکیل دهید .

$\bar{x}_1 - \bar{x}_2 = 16 - 14 - 2$

$\frac{z_1 - \propto}{2} = 1/96$

$\delta_{\bar{x}_{1-\bar{x}_2}} = 0/924$

$C.I:\ (16-14) - (1/96)(0/924) \leq \mu_1 - \mu_2 = 0/19 \leq (16-14) +$

$(1/96)(0/924) > \mu_1 - \mu_2 = 3/81$

$0/19 \leq \mu_1 - \mu_2 \leq 3/81$

تفاوت بین میانگین های دانشجویان علوم انسانی و علوم پایه با احتمال ۰/۹۵ در ایـن فاصـله قـرار می گیرد . فرضیه صفر رد می شود.

جلسه ششم

۹۳/۲/۱۲

توزیع T منظور متغیر پژوهش (وابسته)

بعضی مواقع انحراف استاندارد جامعه برای محقق به لحاظ مقداری مشخص نمی باشد لذا در این شرایط امکان استفاده از توزیع نرمال یا توزیع Z جهت انجام آزمون فرض های آماری اکان پذیر نیست بنابراین می بایست از توزیع های دیگری غیر از توزیع Z استفاده نمود از جمله پر کاربردترین توزیع T می باشد اگر از جامعه ای با توزیع نرمال به دفعات متعدد نمونه هایی را انتخاب و هر بار مشخصه آماری T را برای حالت تک متغیری محاسبه نمائیم با استفاده از مقادیر t به دست آمده می توان نمودار توزیع فراوانی t را ترسیم کرد . که مقادیر t بر روی محور x ها و فراوانی یا احتمال مشاهده مقادیر t بر روی محور y ها قرار می گیرد به توزیع فراوانی حاصل توزیع نمونه ای t گفته می شود .

$$t_{\circ} = \frac{\bar{x} - \mu}{\delta_{\bar{x}}}$$

ویژگی های توزیع T

۱- توزیع T از خانواده توزیع نرمال می باشد و شکل آن نیز مانند توزیع نرمال تقریبا متقارن است به همین دلیل می توان از آن به عنوان آزمون فرض آماری استفاده نمود .

۲- برای هر مقدار تعیین حجم نمونه یک توزیع نظری T وجود دارد بنابراین با خانواده ای از توزیعهای T رو به رو می باشیم در حالی که بدون توجه به حجم نمونه فقط یک توزیع نرمال داریم .

۳- سطوح زیر منحنی در دو سوی توزیع t از توزیع نرمال بیشتر است دلیل آن این است که توزیع t تابع میانگین و انحراف استاندارد نمونه است در حالیکه توزیع Z تابع میانگین و انحراف استاندارد جامعه است .

۴- با افزایش حجم نمونه توزیع t به توزیع Z نزدیک تر می شود . (دلیل : هرچه حجم نمونه بالاتر رود میانگین و انحراف استامندارد نمونه به میانگین و انحراف استاندارد جامعه نزدیکتر می شود) .

۵- توزیع t بر حسب درجه آزادی تفسیر می شود . (دلیل : استفاده از انحراف معیار نمونه برای برآورد خطای معیار میانگین در توزیع t می باشد .

6- خطای معیار میانگین جامعه $\frac{\delta}{\sqrt{n}}$ نمونه $\frac{s}{\sqrt{n}}$

آزمون t تک متغیره :

زمانی از آزمون فوق استفاده می کنیم که بخواهیم میانگین یک گروه را با جامعه آماری آن مقایسه نمائیم و به انحراف استاندارد جامعه نیز دسترسی نداشته باشیم .

مفروضه های آزمون t تک متغیره :

1. مقیاس اندازه گیری متغیر مورد نظر باید فاصله سنی باشد
2. توزیع متغیر مورد نظر در جامعه می بایست نرمال باشد
3. مفروضه مستقل بودن نیز برقرار باشد . (یعنی آزمودنی ها از هم مستقل باشند نمره خانم × تاثیری در نمره های خانم Y نداشته باشد .

مثال :

رئیس دانشگاه آزاد شیراز معتقد است که میانگین آزمون استعداد تحصیلی دانشجویان سال اول این دانشگاه با میانگین داوطلبین ورود به دانشگاه های آزاد کل کشورهای متفاوت است در صورتی که میانگین نمرات استعداد تحصیلی دانشجویان کل کشور برابر با ۵۰۰ باشد فرضیه مورد نظر رئیس دانشگاه را در سطح ۰/۰۵ آزمون نمائید .

باید نمونه گیری کنیم ابتدا در مورد دانشگاه آزاد شیراز بعد میانگین و نمرات استاندارد استعداد تحصیلی را به دست می آوریم

یه نمونه انتخاب کردیم مثلا $\bar{x} = 600$ ، انحراف استاندارد $S = 80$ ، حجم نمونه $n = 20$

گام اول : انتخاب آزمون آماری (می خواهیم ۱ گروه رابا جامعه مقایسه کنیم)

گام دوم : بررسی مفروضه ها

۱- مهم ترین آن مستقل بودن

۲- نرمال بودن (بایستی بررسی کنیم $n > 30$ ، با گروه های مشابه بررسی کنیم $n < 30$)

۳- فاصبه ای و نسبی جزء مفروضه ها نیست جزء شرایط است و بهتر است در گام اول بررسی شود.

گام سوم : بیان فرضیه ها

$$\begin{cases} H_0 = \mu = 500 \\ H_1 = \mu \neq 500 \end{cases}$$

گام چهارم : نخستین ملاک تصمیم گیری (فرضیه صفر را رد می کنیم اگر Z مشاهده شده مساوی یا بزرگتر از Z بحرانی باشد) یعنی آقای رئیس درست مکی گوید و فرضیه صفر را می پذیریم در صورتی که Z مشاهده شده کوچکتر از Z بحرانی باشد .

گام پنجم : محاسبه آماری T

$$\text{مشاهده شده } t = \frac{\bar{x} - \mu}{S_{\bar{x}}} = \frac{600 - 500}{17/97} = \frac{100}{17/97} = 5/56$$

$$\text{خطا معیار میانگین } S_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{80}{\sqrt{20}} = 17/97$$

گام ششم : تصمیم گیری ($\propto = 0/05$ برای آزمون دو متغیره ۱/۹۶) بنابراین فرضیه صفر را رد می کنیم و نتیجه می گیریم $5/56 > 1/96$ که استعداد تحصیلی دانشجویان شیراز با کل کشور متفاوت است همچنین با توجه به مقادیر میانگین ها می کنیم که دانشجویان شیراز وضعیت بهتری را دارند.

برآورد تشکیل فاصله اطمینان

$$\bar{x} - Z_{1-\frac{\alpha}{2}}(S_{\bar{x}}) \leq \mu \leq \bar{x} + Z_{1-\frac{\alpha}{2}}(S_{\bar{x}})$$

مثال :

با توجه به اطلاعات مندرج در سوال قبل فاصله اطمینان ۹۹ درصدی رابرای برآورد میانگین جامعه تشکیل دهید .

$$600 - 2/58(17/97) \leq \mu \leq 600 + 2/58(17/97)$$

$$600 - 46/36 = 553/64 \qquad 553/64 \leq \mu \leq 646/3$$

توزیع نمونه ای t برای تفاوت بین میانگین ها :

اگر از جامعه ای نرمال بارها نمونه های دوتایی را انتخاب کنیم و هر بار مقدار t مشاهده شد . $t_{۴} = \frac{\bar{x}_1 - \bar{x}_2}{S\bar{x}_1 - \bar{x}_2}$ را محاسبه نمائیم . می توانیم نمودار توزیع فراوانی را برای این مقادیر ترسیم نمائیم که در این نمودار مقادیر t بر روی محور x ها و احتمال یا فراوانی مقادیر t بر روی محور y ها قرار می گیرد . در این حالت شکل توزیع به دست آمده شبیه توزیع نرمال بوده و به آن توزیع t برای مقایسه میانگین ها گفته می شود .

با استفاده از فرمول زیر می توان کلیه مسائل را حل کرد

$$\text{خطای معیار میانگین برای توزیع } S_{\bar{x}} = \sqrt{\frac{S_1^2(n_1-1)+S_2^2(n_2-1)}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

آزمون t برای گرو های مستقل :

زمانی از آزمون t برای گروه های مستقل استفاده می شود که قصد مقایسه میانگین های دو گروه را با یکدیگر داشته باشیم و ضمنا انحراف استاندارد جامعه نیز نا معلوم ، باشد و متغیر وابسته باید فاصله ای نداشته باشد .

مفروضه های آزمون t برای گروه های مستقل

۱- **مستقل بودن : ۲** حالت دارد الف) مستقل بودن افراد (یعنی انتخاب هیچ فردی بستگی به انتخاب فرد دیگری نداشته باشد. ب) مستقل بودن گروه ها

۲- **نرمال بودن :** توزیع متغیر مورد نظر در جامعه می بایست نرمال باشد . اگر حجم نمونه ای از ۱۵ به پایین باشد بایستی با گروه های مشابه بررسی کرد و اگر از ۱۵ به بالا باشد باید به صورت تجربی الف) آزمون کالموگروف ب) کشیدگی یا کجی **مورد بررسی قرار گیرد .**

۳- **همگنی واریانس ها :** برقراری همگن واریانس ها به صورت تجربی باید بررسی شود در عین حال اگر حجم گروه ها برابر باشد آزمون t در مقابل این مفروضه مقاوم است . یعنی بررسی هم نکنیم ایرادی ندارد .

مثال : محققی با هدف مقایسه اضطراب آمار دانشجویان دختر و پسر رشته مدیریت آموزشی ۲ نمونه را انتخاب کرد لازم به ذکر است که این محقق اعتقاد دارد که اضطراب آمار دانشجویان دختر و پسر متفاوت می باشد با توجه به اطلاعات به دست آمده از این دو نمونه فرضیه مورد نظر را در سطح ۰/۰۱ مورد آزمون قرار دهید .

گروه ۲	گروه ۱
پسر	**دختر**
$n_2 = 30$	$n_1 = 30$
$\bar{X} = 20$	$\bar{X}_1 = 35$
$S_2 = 2$	$S_1 = 4$

گام ۱ : انتخاب نوع آزمون

۲ گروه مستقل از هم داریم ثانیا متغیر اضطراب است یک متغیر فاصله ای است انحراف استاندارد جامعه ای هم نداریم بنابراین آزمون t مناسب است.

گام ۲ : بررسی مفروضه ها

۱- مستقل بودن : از طریق روش های نمونه گیری

۲- نرمال بودن : الف) گروه های مشابه ب) تجربی

۳- همبستگی واریانس ها : پراکندگی اضطراب دختر ها و پسرها خیلی با هم تفاوت نداشته باشد

گام ۳ : بیان فرضیه

$$H_0 = \mu_1 = \mu_2$$
$$H_1 = \mu_1 \neq \mu_2$$

گام ۴ : تصمیم گیری

فرضیه صفر را رد می کنیم اگر t مشاهده شده بزرگتر از t بحرانی باشد. $t \geq t$ بحرانی مشاهده نشد .

فرضیه صفر را رد می پذیریم اگر t مشاهده شده کوچکتر از t بحرانی باشد.

$$t_{م} \geq t_{ب}$$

گام ۵ : محاسبه مشخصه t

$$\text{مشاهده شده} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2(n_1-1)+S_2^2(n_2-1)}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$= \frac{25-20}{\sqrt{\frac{16(29)+4(29)}{58}\left(\frac{1}{30}+\frac{1}{30}\right)}} = \frac{15}{\sqrt{\frac{464+116}{58}\times 0/06}} = \frac{15}{\sqrt{0/06}}$$
$$= 62/5$$

گام ۶ : t مشاهده شده از t بحرانی بزرگتر است فرضیه صفر را رد می کنیم $62/5 > 2/58$

فرضیه صفر رد می شود .

برآورد فاصله اطمینان برای تفاوت بین میانگین ها :

با توجه به اطلاعات سوال قبل فاصله اطمینان ۹۹ درصدی را برای برآورد تفاوت میانگین های دو گروه دختر و پسر در جامعه تشکیل دهید .

$$(\bar{X}_1 - \bar{X}_2) - Z_{1-\frac{\alpha}{2}}(S_{\bar{x}}) \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + Z_{1-\frac{\alpha}{2}}(S_{\bar{x}})$$

$$(35 - 20) - 2/58(0/24) \leq \mu_1 - \mu_2 \leq (35 - 20) + 2/58(0/24)$$

$$14/39 \leq \mu_1 - \mu_2 \leq 15/62$$

ازمون T برای گروه های وابسته جزء طرح های درون آزمونی (مستقل جزء آزمون های بین آزمونی) یک آزمون در ۲ موقعیت مورد اندازه گیری قرار گیرد

موارد استفاده

۱- هر یک از آزمونی ها در ۲ موقعیت مورد آزمون قرار می گیرد

۲- جفتهای هم تراز یا هم تاشده

۳- بر اساس یک متغیری که با متغیر وابسته همبستگی بالایی دارد جور می کنیم

۴- هوش با اضطراب آمار بالایی دارد

۵- آزمون هوش انتخاب ← اندازه گیری

۶- بر مبنای هوش گروه بندی می کنیم Ø به این گروه همتا می گویند .

مفروضه ها

۱- مستقل بودن (که یک جنبه دارد همان مستقل بودن افراد است)

۲- نرمال بودن (مانند قبل بررسی می شود)

۳- همگن واریانس ها

مثال : یک روانشناس معتقد است که استفاده از روش آزمون حل مساله تا حدود زیادی اعتماد به نفس دانشجویان درس آمار استنباطی را افزایش می دهد به همین دلیل یک نمونه ۸ نفره را انتخاب کرد تا فرضیه مورد نظر خود را آزمون نماید نمرات زیر مربوط به اعتماد به نفس این دانشجویان قبل و بعد از آزمون حل مساله می باشد . فرضیه مورد نظر محقق را در ۰/۰۱ آزمون نمائید .

جلسه
هفتم

۹۳/۲/۱۹

توزیع F : اگر از جامعه ای با توزیع نرمال نمونه های متعددی را انتخاب کرده و هر بار با هدف مقایسه میانگین این نمونه ها مقدار F را محاسبه نمائیم می توانیم با برقرار مقـادی F بـر روی محور x ها و فراوانی یا احتمال مشاهده مقادیر F بر روی محور y ها توزیع فراوانـی را ترسـیم نمائیم توزیع حاصل ، توزیع F می گوئیم.

$$F_{\text{مشاهده شده}} = \frac{M_{Sb}}{M_{SW}} \qquad F_{\circ} = \frac{M_{Sb}}{M_{SW}}$$

ویژگی های توزیع F :

۱- با افزایش حجم نمونه ، میانگین توزیع F بـه یـک نزدیـک مـی شـود (ولـی در توزیـع z ، میانگین به میانگین جامعه نزدیک می شود)

۲- توزیع F دارای ویژگی تک نمایی است

یک مقدارF هست که دارای بیشترین فراوانی است که این دو مقدار باشد امکن ندارد — در توزیع نرمال میانه ، انحراف روی هم قرار می گیرد ولی در F اینطور نیست .

۳- حد پایین توزیع F برابر با صفر است بنابراین توزیع دارای کجی مثبت است بنابراین مقادیر منفی نداریم

۴- توزیع F نیز مانند توزیع T تشکیل خانواده ای از توزیع های F می دهد چرا که این توزیع بر حسب درجات آزادی تشکیل می شود .

۵) بین توزیع T و F رابطه زیر برقرار است

$$F_{(1,d_{fw})} = t^2_{(dfw)}$$

دو درجه داریم : الف) آزادی بین گروهی : تعداد گروهها ب) آزادی

اگر تعداد گروه ها ۲ باشد مهای یک شده ← شده ۱

به جای تحلیل واریانس از آزمون T می توان استفاده کرد .

آزمون تحلیل واریانس

هرگاه قصد مقایسه میانگین های ۲ یا بیش از ۲ گروه را داشته باشیم از روش آماری تحلیل واریانس استفاده می کنیم .

توجه داشته باشیم که اگر دارای یک متغیر مستقل باشیم از تحلیل واریانس یک عاملی و اگر دارای دو متغیر مستقل باشیم از تحلیل واریانس دو عاملی استفاده می کنیم . همچنین اگر دارای بیش از یک متغیر وابسته باشیم در این شرایط از تحلیل واریانس چند متغیره (من و وا) استفاده می کنیم تحلیل واریانس یک متغیره (آن و وا) یا یک عاملی یک راهه

تاثیر روش های تدریس بر اعتماد به نفس و پیشرفت تحصیلی

تاثیر روش های تدریس دختر بر پیشرفت تحصیلی

مثال : تاثیر روش تدریس دختران بر پیشرفت تحصیلی و اعتماد به نفس

۱) تاثیر روش های تدریس بر پیشرفت تحصیلی دانش آموزان

چون یک متغیر وابسته بیش از دو سطح (تحلیل واریانس یک عاملی)

۲) تاثیر روش تدریسو جنسیت بر پیشرفت تحصیلی (آن ووای دوعاملی دو طرفه دوراهه)

۲ مستقل ۱ وابسته

۳) تحلیل متغیر

الف) ۲ عامل وابسته و ۱ عامل مستقل (من و وای تک عاملی)

ب) عامل وابسته و ۲ عامل مستقل (من و وای دو عاملی)

۱ متغیر مستقل و ۱ وابسته	آن و وای یک عاملی
۲ متغیر مستقل و ۱ وابسته	آن و وای دو عاملی
۱ مستقل و ۲ وابسته	من و وای یک عاملی
۲ مستقل و ۲ وابسته	من و وای دو عاملی

تحلیل واریانس آن وای یک طرفه

پیش نیازها :

۱) وجود یک متغیر مستقل با دو یا بیش از دو سطح (برای دو سطح می شود استفاده کرد ولی مرسوم نیست)

۲) سطوح متغیر مستقل به لحاظ کمی یا کیفی باید با هم مستقل باشند . (مثلا وقتی می گوئیم سه روش تدریس روشها تفاوت کیفی دارند) .

۳) هر فرد می تواند داخل یکی از پروه ها قرار بگیرد

۴) یک متغیر وابسته داشته باشیم که مقیاس آن فاصله ای و نسبی باشد .

مفروضه های استفاده از تحلیل واریانس یک عاملی

۱) **مستقل بودن :** مستقل بودن افراد و گروه ها (با روش نمونه گیری صحیح)

۲) **نرمال بودن توزیع نمره های متغیر وابسته در جامعه :** اصطلاحا اگر حجم گروه ها برابر باشند تحلیل واریانس به این مفروضه ها مقاوم است.

۳) **همگن واریانس ها :** واریانس های گروه A و B و C باید در جامعه به هم نزدیک باشد .

چرا می گوئیم در جامعه : بر اساس نمونه بررسی می کنیم و چون بر حسب تصادف انتخاب می کنیم پس می گوئیم جامعه

فرضیه های مورد آزمون در تحلیل واریانس یک عاملی :

در تحلیل واریانس آزمون یک دامنه نداریم فرضیه جهت دار نداریم

چون مقایسه ها با هم مستقل نیستند . $A = B = C = D$

- به دلیل مستقل نبودن مقایسه ها از آزمون آماری یک دامنه نمی توانیم استفاده کنیم .

- به این دلیل نمی توانیم از چند آزمون t مستقل به جای آزمون تحلیل واریانس استفاده کنـیم که به دلیل مستقل نبودن مقایسه ها مقدار خطای نوع اول بر اساس رابطـه زیـر افـزایش پیـدا بکند .

AB

BC $\qquad \propto = 1-(1-\propto)^c$

AC

$C=$ تعداد مقایسه ها

$1=$ همیشه عدد ثابت است

$\propto=$ احتمال ارتکاب خطای نوع اول

مثال : تعداد ۳ گروه و احتمال خطای نوع اول ۰/۰۵ باشد

$$\propto = 1-(1-0/05)^3 = 0/14$$

یعنی ۱۴ خطا را می پذیریم

اگر تعداد مقایسه ها بیشتر باشد این احتمال (ارتکاب خطا) بیشتر است .

- در تحلیل واریانس یک عاملی فقط یک فرضیه داریم

$H_0 = \mu_1 = \mu_2 \ldots\ldots \mu_k$ به تعداد گروهها ← بین گروه ها تفاوتی نیست

$H_1 = \mu_i \neq \mu_{i'}$ ← فرضیه مقابل می گوید

یعنی حداقل بین یک گروه با گروه های دیگر تفاوت وجود دارد.

تفکیک مجذورات در تحلیل واریانس

در آمار توصیفی مجموع مجذورات

$$S^2 = \frac{\sum(x-\bar{x})^2}{N-1}$$

مجموع مجذورات یعنی تحلیل واریانس

$$Ss_t = S_b S + SS_w$$

$SS_b =$ شانس و تصادف و خطا + تاثیر متغیر مستقل

$SS_w =$ خطا و شانس و تصادف

فرض من این است که بر اساس نمونه گیری با هم تفاوتی ندارند اگر با هم باشند نتیجه شانس است.

محاسبه درجات آزادی در تحلیل واریانس :

$$dF_b = k - 1$$

$$dF_w = N - K$$

$$df_t = N - 1$$

K← تعداد گروه ها -۱

N← تعداد کل افراد

K← تعداد گروه ها

$$df_t = dF_b + dF_w = k - 1 + N - K = N - 1$$

مثال : تعداد ۳ گروه و احتمال خطای نوع اول ۰/۰۵ باشد

$$\propto = 1 - \left(1 - \frac{0}{05}\right)^3 = 0/14$$

اگر تعداد مقایسه ها بیشتر باشد این احتمال (ارتکاب خطا) بیشتر است .

- در تحلیل واریانس یک عاملی فقط یک فرضیه داریم

$H_o = \mu_1 = \mu_2 \ldots\ldots \mu_k$ ← بین گروه ها تفاوتی نسبت

$$H_1 = \mu_2 \neq \mu_{i\prime}$$

بعضی حداقل سن یک گروه با گروه های دیگر تفاوت وجود دارد

محاسبه میانگین مجذورات یا میانگین

$MS_t = \frac{SS_t}{df_t}$ $MS_b = \frac{SS_b}{df_b}$ $MS_w = \frac{SS_W}{df_w}$

محاسبه مقدار F مشاهده شده :

$$F_{\circ} = \frac{M_{sb}}{M_{sw}}$$

جدول خلاصه تحلیل واریانس

منبع مجذورات SS_t	مجذورات SS	درجه آزادی df	میانگین مجذورات MS	مقدار F
SS_b SS_W	SS_b SS_W	$N - k$		

SS_t	SS_t	$N-1$	$\frac{SS_t}{N-1}$	$\frac{MS_b}{MS_w}$
SS_b	SS_b	$k-1$	$\frac{SS_b}{k-1}$	
SS_w	SS_w	$N-k$	$\frac{SS_w}{N-k}$	

محاسبه مجذورات در تحلیل واریانس :

$$SS_t = \sum x^2 - \frac{(\sum x)^2}{N}$$

$$SS_W = (SS_{W1} + SS_{W2} + \cdots .. + SS_{Wk})$$

درون هر گروه تک تک حساب کنیم بعد جمع بزنیم فرمول زیر (همان فرمول کلی است ولی ما برای گروه ها تک تک حساب می کنیم).

$$SS_{W1} = \sum x_1^2 - \frac{(\sum x_1)^2}{N_1}$$

$$SS_b = SS_t - SS_W$$

مقدار F	میانگین مجذورات MS	درجه آزادی df	مجذورات SS	منبع مجذورات SS_t
$\frac{MS_b}{MS_w}$	$\frac{SS_b}{k-1}$	$k-1$	SS_b	SS_b
	$\frac{SS_W}{\mathrm{N-k}}$	$\mathrm{N-k}$	SS_W	SS_W
	$\frac{SS_t}{N-1}$	$N-1$	SS_t	مجذورات کل

مثال : سوال امتحان

پژوهشگری قصد دارد تاثیر سه نوع روش تدریس سخنرانی ، حل مساله و بحث گروهی را بـر عـزت نفس دانش آموزان مورد بررسی قرار دهد برای این منظور وی ۳ کلاس ۵ نفره را انتخـاب و جهـت آموزش مطالب درسی از این سه روش استفاده نمود با توجه به اطلاعات به دست آمده فرضیه مورد نظر این محقق در ۰/۰۵ را مورد آزمون قرار دهید ؟

x^2	بحث گروهی	x^2	حل مساله	x^2	سخنرانی x
۲۲۵	۱۵	۱۰۰	۱۰	۳۶	۶
۱۹۶	۱۴	۶۴	۸	۱۶	۴
۱۴۴	۱۲	۸۱	۹	۴	۲
۱۶۹	۱۳	۱۴۴	۱۲	۲۵	۵
۱۰۰	۱۰	۱۰۰	۱۰	۹	۳
$\sum x^2 = 834$	$\sum x = 64$	$\sum x^2 = 489$	$\sum x = 49$	$\sum x^2 = 90$	$\sum x = 20$

$$SS_t = \sum x^2 - \frac{(\sum x)^2}{N}$$

	A	B	C		کل
$\sum x$	20	49	64	$\rightarrow$	133
$\sum x^2$	90	489	434	$\rightarrow$	1413

$$SS_t = 1413 - \frac{(133)^2}{3 \times 5} = 1413 - \frac{17689}{15} = 1413 - 1179 = 234$$

$SS_t = 234$ یا $233/7$

$$SS_{w_{\text{س}}} = 90 - \frac{(20)^2}{5} = 90 - \frac{400}{5} = 10$$

$$SS_{w_{\text{ص}}} = 489 - \frac{(49)^2}{5} = 8/8$$

$$SS_{w_{\text{محدث}}} = 834 - \frac{(64)^2}{5} = 14/8$$

$$SS_w = SS_{w_1} + SS_{w_2} + SS_{w_3} = 10 + 8/8 + 14/8 = 33/6$$

$$SS_b = SS_t - SS_w = 233/7 - 33/6 = 200/1$$

	SS	df	MS
بین گروهی SS_b	200/1	2	$\frac{200/1}{2} = 100/05$ $F_{\text{م}} = \frac{100/05}{2/8} = 35/75$
SS_w درون گروهی	33/6	12	$\frac{33/6}{12} = 2/8$
SS_tمجذورات کل	233/7	14	$\frac{233/7}{14} = 16/70$

F مشاهده شده را به دست آوردیم گام بعدی F جدول را حساب کنیم که در کتابها هست اول درجه آزادی صورن را در ستون پیدا کنیم درجه آزادی ۲ هست . بعد مخرج ۱۲ را پیدا می کنیم ۲ تا عدد هست اولی برای سطح ۰/۰۵ و دومی برای سطح ۰/۰۱ چون ما سطح ۰/۰۵ خواستیم پس طبق جدول ۱۹/۴۱ می شود

$F_{ب} = 19/41$ $\qquad$ $F_{م} = 35/75$

گام آخر : تصمیم گیری ۳۵/۷۵ بزرگتر از F بحرانی که ۱۹/۴۱ است پس فرضیه صفر را رد می کنیم و می گوئیم $35/75 > 19/41$ این سه روش تدریس بر عملکرد دانش آموزان تاثیر مثبت دارد .

۱) انتخاب آزمون

۲) بیان فرضیه ← فرضیه صفر

$$H_0 = \mu_1 = \mu_2 = \mu_3$$
$$H_0 = \mu_{س} = \mu_{ح} = \mu_{ب}$$
$$H_1 = \mu_i \neq \mu_{i'}$$

۳) بررسی مفروضات مستقل بودن همگن واریانس

۴) تعیین ملاک تصمیم گیری : اگر F مشاهده شده بزرگتر یا مساوی F بحرانی باشد فرضیه صفر را رد می کنیم اگر F مشاهده شده کوچکتر باشد فرضیه صفر را می پذیریم

۵) محاسبات

۶) تصمیم گیری مقایسه F مشاهده شده و F بحرانی .

چون می خواهیم ببینیم تاثیر و تفاوت گروه ها چقدر است پس از آزمون تعقیبی استفاده می کنیم .

جلسه هشتم

۹۳/۲/۲۶

آزمانهای تعقیبی (پس تجربی – پست هاک)

زمانی از مقایسه های پس تجربی استفاده می کنیم که مقدار F به دست آمده از تحلیل واریانس از نظر آماری معنی دار باشد همان طور که ذکر شد بر اساس معنی داری F تحلیل واریـانس فقـط می توانیم در مورد وجود یا عدم وجود رابطه خطی بین متغیر مستقل و وابسته قضـاوت کنـیم لـذا برای پی بردن به وجود تفاوت بین سطوح متغیر مستقل در زمینه متغیر وابسته از آزمون های پس تجربی استفاده می کنیم .

روشهای مختلفی برای مقایسه های پس تمرینی وجود دارد به عنوان مثال آزمون LSD یا دانت یـا نیومن گلز یا $Less$. (بون فورنی – شفه – توکی)

آزمون توکی یا HSD

مورد استفاده : زمانی که حجم نمونه در گروه ها برابر باشد و قصد مقایسه ها دو به دوی میانگین ها را داشته باشیم از آزمون تعقیبی توکی استفاده می کنیم همچنین در برخی منابع گفته شـده کـه در صورت مساوی نبودن حجم نمونه ها می توان ، تعدیل حجم نمونه ها برای مقایسه های دوتـایی از آزمون توکی استفاده نمود .

تحلیل واریانس دو عاملی اثرهای آمیخته :

اگر در تحلیل واریانس سطوح یکی از متغیرها تمامی سـطوح ممکـن یـا مـورد علاقـه پژوهشـگر را تشکیل دهد و سطوح متغیر دیگر به تصادف توسط پژوهشگر انتخاب شود . به آن تحلیـل واریـانس دو عاملی اثرهای آمیخته می گویند .

تحلیل واریانس دو عاملی

مورد استفاده : هرگاه بخواهیم اثر دو متغیر مستقل حداقل با ۲ سطح را بر یک متغیـر وابسـته کـه دارای مقیاس فاصله ای و نسبی است مورد آزمون قرار دهیم از آزمون تحلیـل واریـانس دو عـاملی استفاده می کنیم.

هدف اصلی از اجرای آزمون تحلیل واریانس دو عاملی پاسخ به این سوال است که آیا تفاوت بین میانگین ها در سطح عامل A و B معنی دار است یا خیر

انواع مدل های تحلیل واریانس دو عاملی :

۱- مدل تحلیل واریانس دو عاملی اثرهای ثابت : اگر سطوح متغیرهای مستقل در تحلیل واریانس تمامی سطوح ممکن یا مورد علاقه پژوهشگر را تشکیل دهد به آن تحلیل واریانس دو عاملی با اثرهای ثابت گفته می شود .

۲- تحلیل واریانس دو عاملی اثرهای تصادفی : اگر در تحلیل واریانس سطوح متغیرهای مستقل تمامی سطوح مورد علاقه پژوهشگر را تشکیل ندهد به آن تحلیل واریانس دو عاملی اثرهای تصادفی گفته می شود .

اگر در تحلیل واریانس سطوح متغیر مستقل تمامی سطوح مورد علاقه پژوهشگر با تمامی سطوح ممکن را تشکیل دهد به این نوع تحلیل واریانس ، تحلیل واریانس اثرهای ثابت گفته می شود .
اما در صورتی که در تحلیل واریانس سطوح متغیر مستقل تمامی سطوح مورد علاقه یا ممکن را تشکیل ندهد به آن تحلیل واریانس اثرهای تصادف گفته می شود .

- در تحلیل واریانس اثرهای تصادفی جهت تبعیین تاثیر متغیر مستقل بر متغیر وابسته از مجذور ρ استفاده می کنیم .

مثال : با فرض تصادفی بودن تحلیل واریانس در مثال قبل شاخص قدرت تاثیر روش تدریس بر اعتماد به نفس دانش آموزان را محاسبه نمایند .

۶۹ درصد واریانس عزت نفس توسط روش های تدریس تعیین می شود .

شاخص تحمل (قدرت تاثیر)

مقدار F مشاهده شده در تحلیل واریانس به ما می گوید که متغیر مستقل بر متغیر وتبسته تاثیر دارد ولی هیچ اطلاعاعی در مورد قدرت یا شدت ایت تاثیر در اختیار ما قرار نمی دهند . لذا برای

این منظور می بایست در تحلیل واریانس اثرهای ثابت از مجذور امگا برای تعیین قدرت تاثیر متغیر مستقل بر وابسته استفاده کنیم .

$$w^2 = \frac{SS_b - (k-1)MS_w}{SS_t + MS_w}$$

مثال : با توجه به مثال قبل در تحلیل واریانس میزان تاثیر روشهای تدریس بر اعتماد به نفس دانش آموزان را محاسبه نمائید .

$$SS_b = 200/1$$
$$SS_t = 233/7 \qquad w^2 = \frac{SS_b - (k-1)MS_w}{SS_t + MS_w} = \frac{200/1 - 2(2/8)}{233/7 + 2/8} = 0/82$$
$$k - 1 = 2$$
$$MS_w = 2/8$$

با توجه به مجذور امگا نتیجه می گیریم که ٪ ۸۲ از واریانس اعتماد به نفس توسط روش های تدریس تعیین می شود .

مثال : در مثال قبل $n_1 = 5$ ، $n_2 = 4$ ، $n_3 = 3$ با توجه به معنی دار شدن نتایج تحلیل واریانس آزمون تعقیبی شفه را برای مقایسه دو به روی میانگین ها به کار ببرید .

گام ۱ : بیان فرضیه

$$H_0 = \mu_1 = \mu_2$$
$$H_1 = \mu_1 \neq \mu_2$$

گام ۲ : تشکیل مقایسه : (وزن را هر عددی بدهیم اشکالی ندارد فقط عددی غیر از صفر و نتیجه صفر باشد)

فرض $w_1 = -2$ $\qquad w_2 = 1$ $\qquad w_3 = 1$

سخنرانی $\mu_1 = 4$ $\qquad$ حل مساله $\mu_2 = 9/8$ $\qquad$ بحث گروهی $\mu_3 = 12/8$

$$C = w_1(\mu_1) + w_2(\mu_2) + w_3(\mu_3) = -2(4) + 1(9/8) + 1(12/8) = 13/6$$

گام ۳ : محاسبه t مشاهده شده

$$t_{م} = \frac{C}{\sqrt{MS_w(\frac{w_1}{n_1}+\frac{w_2}{n_2}+\frac{w_3}{n_3})}} = \frac{13/6}{\sqrt{2/8(\frac{-2}{5}+\frac{1}{4}+\frac{1}{3})}} = 19/4$$

گام ۴ : محاسبه t بحرانی

$$t_{ب} = \sqrt{(k-1)f} = \sqrt{(3-1)(8/02)} = \sqrt{2\times 8/02} = \sqrt{16/4} = 4/2$$

گام ۵: مقایسه و تصمیم گیری

$$t_{م} \geq t_{ب} \Rightarrow \frac{19}{4} \geq \frac{4}{2} \rightarrow \text{فرضیه صفر رد می شود}$$

$$H_0 = \mu_{س} = \frac{\mu_{ب}+\mu_{ج}}{2}$$

آزمون شفه را فقط برای مقایسه ترکیبی استفاده کنید .

$$\mu_1 : \mu_{س} \neq \frac{\mu_{ب}+\mu_{ج}}{2}$$

تنها تفاوت ترکیبی یا دو تایی در بیان فرضیه است .

لازم به ذکر است که این وزن ها اولا می بایست که اعداد غیر صفر باشند و ثانیا مجموع آنها نیز برابر با صفر باشد .

$$C = W_1(\mu_1) + W_2(\mu_2) + \cdots + W_k(\mu_k)$$

در عمل میانگین جامعه (μ) را نداریم پس به جای آن $\bar{x}$ قرار می دهیم .

گام ۳ : محاسبه مقدار t مشاهده شده

$$T_{م} = \frac{C}{\sqrt{MS_w\left(\frac{w_1}{n_1}+\frac{w_2}{n_2}+\cdots,+\frac{w_k}{n_k}\right)}}$$

گام ۴ : محاسبه t بحرانی

$$T = \sqrt{(k-1)f_{ب(df_b,df_w,\alpha)}}$$

گام ۵ : مقایسه و تصمیم گیری

$$if \;\; t_{م} \geq t_{ب} \Rightarrow \text{تفاوت معنادار} \rightarrow \text{فرض صفر رد می شود}$$

$if\ \ t_{\text{م}} < t_{\text{ج}} \Rightarrow$ تفاوت بی معنی $\rightarrow$ فرض صفر را می پذیریم

با توجه به نتایج آزمون توکی روش حل مساله از نظر تاثیر گذاری بر عـزت نفـس نسـبن بـه روش سخنرانی روش توانمند تری می باشد همچنین با توجه به نتایج روش بحث گروهی نیـز نسبت بـه روش سخنرانی تاثیر بیشتر بر اعتماد به نفس دانشجویان دارد در حالی که که از نظر تاثیر گذاری بر اعتماد به نفس بین روش بحث گروهی و حل مساله تفاوتی وجود ندارد.

آزمون شفه :

مورد استفاده: برای مقایسه دو به دوی میانگین ها (زمانی که حجم نمونه ها برابر نباشد) و مقایسـه ترکیبی میانگین ها در صورت معنادار شدن نتایج واریانس از آزمون تعقیبی شفه استفاده می کنیم.

$$\begin{cases} H_0 = \mu_1 = \mu_2 \\ H_1 = \mu_1 \neq \mu_2 \end{cases}$$

آمار روش سخنرانی حـداقل بـا $\frac{\mu_{\text{ب}}+\mu_{\text{ح}}}{2} = H_0 = \mu_{\text{س}}$ مقایسـه ترکیبـی بـا

متوسط حل مسالهو بحث گروهی $\frac{\mu_{\text{ب}}+\mu_{\text{ح}}}{2} = H_1 = \mu_{\text{س}}$ برابری می کند یا نه ؟

مراحل انجام آزمون

گام ۱ : بیان فرضیه

$$\begin{cases} H_0 = \mu_1 = \mu_2 \\ H_1 = \mu_1 \neq \mu_2 \end{cases}$$

گام ۲ : تشکیل مقایسه : برای این منظور می بایست ابتدا به میانگین گروه ها وزن دهیم .

مثال : در مثال قبلی در صورت معنادار بودن مقدار F آزمـون تـوکی را بـرای مقایسـه دو بـه دوی روشهای تدریس به کار ببرید .

بر اساس فرمول $n = \frac{k}{\frac{1}{n_1}+\frac{1}{n_2}+\cdots+\frac{1}{k}}$ می توان حجم نمونه را تعدیل کرد و از یک n واحـد اسـتفاده کنیم .

مراحل انجام آزمون توکی :

گام ۱ : میان فرضیه

$$\begin{cases} H_o: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

گام ۲ : ماتریسی را بر اساس تعداد سطوح متغیر مستقل و تفاوت بین میـانگین هـای ایـن سـطوح تشکیل می دهیم .

X : تفاوت بین میانگین ها به جای ضربدر قرار می گیرد.

	1	2	3
سخنرانی =۱	.		
میزگرد ۲	×	.	
حل مساله ۳	×	×	.

گام ۳ : محاسبه مقدار HSD :

$$HSD = q_{(d,df,k)}\sqrt{\frac{MS_w}{n}}$$

گام ۴ : مقایسه تفاوت بین میانگین ها با مقدار HSD

گام ۵ : تصمیم گیری : اگر مقدار HSD بزرگتر از تفاوت میانگین ها بود فرضیه صفر را می پذیریم.

اگر مقدار HSD کوچکتر یا مساوی تفاوت بین میانگین ها بود فرضیه صفر را رد می کنیم .

جلسه نهم

۹۳/۳/۲

یک روانشناس معتقد است که استفاده از روش آزمون حل مساله تا حدود زیادی اعتماد به نفس دانشجویان درس آمار استنباطی را افزایش می دهد به همین دلیل یک نمونه ۸ نفره را انتخاب کرده تا فرضیه مورد نظر خود را آزمون دهد . نمرات زیر مربوط به اعتماد به نفس این دانش جویان قبل و بعد از آزمون حل مسئله می باشد فرضیه مورد نظر محقق را در سطح ۱ آزمون نمائید .

- **تعیین آزمون** ← پیش آزمون و پس آزمون (آزمون T وابسته)
- **مفروضه** ← الف) همگنی واریانس ب) فاصله ای و .... ج) نرمال بودن – مستقل بودن
- **بیان فرضیه** $\begin{cases} H_0 = \mu_1 = \mu_2 \\ H_1 = \mu_1 \neq \mu_2 \end{cases}$
- **تعیین ملاک تصمیم گیری :**

الف) $T_م > T_ج$ ← فرضیه صفر رد می شود

ب) $T_م < T_ج$ ← فرضیه صفر تائید می شود

۵. محاسبه T مشاهده شده

ردیف	نمره پیش آزمون	نمره پس آزمون	D	$\overline{D}$	$(D - \overline{D})$	$(D - D')^2$
۱	۴	۶	۲	۱/۵	۰/۵	۰/۲۵
۲	۶	۷	۱	۱/۵	-۰/۵	۰/۲۵
۳	۲	۳	۱	۱/۵	-۰/۵	۰/۲۵
۴	۸	۱۰	۲	۱/۵	۰/۵	۰/۲۵
۵	۱۰	۸	۲-	۱/۵	-۳/۵	۱۲/۲۵
۶	۵	۶	۱	۱/۵	۰/۵	۰/۲۵
۷	۳	۵	۲	۱/۵	۰/۵	۰/۲۵
۸	۱	۴	۳	۱/۵	۱/۵	۲۵/۲
						$\sum(D - D')^2 = 16$

$$S_{\bar{D}} = \frac{0/57}{\sqrt{8}} = 0/20$$

$$t_{م} = \frac{\bar{D}}{S_D}$$

$$t_{م} = \frac{x_1 - x_2}{\frac{S_D}{\sqrt{8}}}$$

$$t_2 = \frac{x_1 - x_2}{\sqrt{S_{\bar{x}_1}^2 + S_{\bar{x}_2}^2 - 2(r_{xy})S_{\bar{x}_1}S_{\bar{x}_2}}}$$

انحراف استاندارد تفاوت ها

$$S_D' = \sqrt{\frac{\sum(D-\bar{D})^2}{n-1}} = \sqrt{\frac{16}{7}} = 0/57$$

با توجه به اینکه $T_{م} > T_{ب}$ است فرضیه صفر رد می شود پس نتیجه می گیریم آموزش دوره حل مسئله بر اعتماد به نفس تاثیر می گذارد .

تفاوت ازمون T برای گرو های مستقل و آزمون T برای گرو های وابسته :

۱- ازمون T برای گرو های مستقل یعنی میانگین دو گروه مستقل را بیان می کند

ازمون T برای گرو های وابسته یا اندازه های یک گروه در ۲ موقعیت

ازمون T برای گرو های وابسته یا ۲ گروه جفت شده ، همتا شده

در واقع اینجا یک گروه ۲ گروه مستقل

۲- درجه آزادی مستقل $d_f:\ n_1 + n_2 - 2 \rightarrow t_{\text{مستقل}}$

$d_f:\ n - 1 \rightarrow t_{\text{وابسته}}$

۳- مقدار خطای معیار

مقدار خطای معیار میانگین در آزمون t گرو های وابسته نسبت به t گرو های مستقل کمتر است یعنی آزمون قویتر است .

۴- تفاوت

در نظر گرفتن همبستگی بیش آزمون و پس آزمون در محاسبه خطای معیار میانگین گروه های وابسته

برآورد فاصله ای برای گروه های وابسته :

جهت برآورد فاصله ای گروه های وابسته دقیقا مانند تفاوت میانگین گروه های مستقل عمل می کند

$$CI = (\bar{x}_1 - \bar{x}_2) - t_{\propto,df}\left(S_{\bar{x}_1-\bar{x}_2} 1 < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) +\right) t_{\propto,df}\left(S_{\bar{x}_1-\bar{x}_2}\right)$$

مثال : فاصله اطمینان ۰/۹۹ برای داده های مثل قبل تشکیل دهید .

$$CI = (4/87 - 6/12) - (3/49)(0/2) < \mu_1 - \mu_2 < (4/87 - 6/12) + (3/49)(0/2) - 1/95 < \mu_1 - \mu_2 < -0/55$$

تفسیر : با ۹۹ درصد اطمینان ، فاصله اطمینان به دست آمده

تفاوت واقعی دو نمونه وابسته در جامعه را شامل می شود .

تفاوت اعتماد به نفس قبل از حل مساله و بعد از حل مساله

تعریف برخی اصطلاحات مرتبط با تحلیل واریانس دو عاملی

طرحهای کاملا متقاطع :

متغیر	مرد	زن
رشته	علوم انسانی	علوم پایه

اگر در تحلیل واریانس دو عاملی سطوح یک متغیر تمامی سطوح متغیر دیگر را قطع کند به آن طرح کاملا متقاطع گفته می شود .

طرح تحلیل واریانس نامتقاطع :

اگر در طرح واریانس دو عاملی سطوح یک متغیر تمامی سطوح متغیر دیگر را پوشش ندهد به آن طرح نامتقاطع گفته می شود .

اثر متقابل :

یکی از مزایای تحلیل واریانس دو عاملی نسبت به تحلیل واریانس یک عاملی امکان بررسی اثر متقابل عامل های A و B بر متغیر وابسته می باشد .

مرادی علوم انسانی با مرادی علوم پایه

پیشرفت تحصیلی چه طور می باشند .

نحوه نمایش اثر متقابل عامل های A و B بر متغیر وابسته به کمک نمودار

گام اول : ترسیم محور مختصات

گام دوم : مطرح کردن محور x بر اساس یکی از سطوح ذمتغیرهای مستقل

گام سوم : مدرج کردن محور y ها بر اساس مقادیر متغیر وابسته

گام چهارم : ترسیم نمودار بر اساس میانگین خانه های جدول برای سطح متغیر دیگر

اثر متقابل نامنظم :

اگر خطوط داخل نمودار موازی نبوده و یکدیگر را قطع کنند دو متغیر مستقل A و B دارای اثر متقابل نامنظم است .

یعنی جنسیت در ارتباط با رشته های تحصیلی برای علوم پایه جنسیت مرد باعث افزایش شده ولی زن ها عملکردشان پائین آمده ← پس سطوح متفاوت بوده

عدم وجود اثر متقابل :

اگر خطوط داخل نمودار موازی بوده و یکدیگر را قطع نکنند بیانگر عدم وجود اثر متقابل بین متغیر A و B در زمینه تاثیر گذاری بر متغیر وابسته می باشد .

اثر متقابل منظم : اگر خطوط داخل نمودار موازی نبوده و یکدیگر را قطع نکنند بیـانگر وجـود اثـر متقابل منظم بین متغیر A و B از نظر تاثیر گذاری بر متغیر وابسته می باشد .

انواع فرضیه های مورد آزمون در تحلیل واریانس دو عاملی :

$A \quad H_0: \mu_1 = \mu_1 \ldots\ldots\ldots \mu_k$
$H_1:: \ \mu_1 \neq \mu_{i'}$
$B \quad H_0: \mu_1 = \mu_2 \ldots\ldots\ldots \mu_k$
$H_1:: \ \mu_1 \neq \mu_{i'}$
$H_0:$ اثر $AB = 0$

$H_1:$ اثر $AB \neq 0$

افراض منبع مجذورات در تحلیل واریانس دو عاملی :

کل منبع مجذورات

$SS_t = A$عامل $+ B$ عامل $+ AB +$ گروهی درون $\rightarrow$ آزمونهایی به درون هر گروه

$$SS_t = SS_A + SS_B + SS_{AB} + SS_w$$

$SS_A =$ در واقع یعنی تغییرات در متغیر وابسته که نتیجه عامل A و شانس و تصادف می باشد که نتیجه عامل A و شانس و تصادف می باشد.

$SS_B =$ تاثیر B + شانس و تصادف

$SS_{AB} =$ تاثیر AB + شانس و تصادف

$SS_w =$ شانس و تصادف

نحوه محاسبه درجات آزادی در تحلیل واریانس دو عاملی :

$df_A: C - 1$ $\rightarrow$ 1-A تعداد سطوح عامل $\rightarrow$

$df_B = R - 1$ $\rightarrow$ 1-B تعداد سطوح عامل

$$df_{AB} = (C - 1)(R - 1)$$

$$df_W = CR(n - 1)$$

نحوه محاسبه میانگین مجذورات :

$$MS_A = \frac{SS_A}{df_A}$$

$$MS_B = \frac{SS_B}{df_B}$$

$$MS_{AB} = \frac{SS_{AB}}{df_{AB}}$$

$$MS_W = \frac{SS_W}{df_W}$$

نحوه محاسبه F

$$F_A = \frac{MS_A}{MS_W}$$

$$F_B = \frac{MS_B}{MS_W}$$

$$F_{A.B} = \frac{MS_{AB}}{MS_W}$$

سوال : اول اثر تعامل بایستی در تفسیر بررسی شود.

منبع مجذورات	مجموع مجذورات	df	Ms	F
مجذورات عامل A	SS_A	$c-1$	$\frac{SS_A}{c-1}$	$\frac{MS_A}{MS_W}$
مجذورات عامل B	SS_B	$R-1$	$\frac{SS_B}{R-1}$	$\frac{MS_B}{MS_W}$
مجذورات عامل AB	SS_{AB}	$(c-1)(R-1)$	$\frac{SS_{AB}}{(c-1)(R-1)}$	$\frac{MS_{AB}}{MS_W}$
مجذورات عامل درون گروهی	SS_w	$CR(n-1)$	$\frac{SS_W}{CR(n-1)}$	
مجذورات کل	SS_t	$CR_n(n-1)$		