آزمون میانگین یک جامعه

حالت دوم: انحراف معیار متغیر مورد بحث در جمعیت معلوم نیست (σ معلوم نیست)

وقتی انحراف معیار جمعیت معلوم نباشد، محقق باید انحراف معیار جمعیت را با استفاده از نمونه n تایی و فرمول $S = \sqrt{S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$ برآورد کند.
در این حالت شاخص آماری آزمون عبارت خواهد بود $\frac{\bar{x} - \mu_{H_0}}{S/\sqrt{n}}$ که توزیع آماری این شاخص t می‌باشد.

در اینجا قبل از ادامه بحث توجه شما را به توزیع t و خواص آن که در صفحات قبل بیان شده است جلب می‌کنم و بخصوص این نکته که وقتی $n \to \infty$ آنگاه $t \to z$ خواهد بود و مرز بزرگ بودن و کوچک بودن تعداد نمونه نیز تعداد نمونه کمتر یا مساوی ۳۰ و بیشتر از ۳۰ در نظر گرفته شده است.
بنابراین علی‌رغم اینکه توزیع شاخص آماری آزمون، t با n-1 درجه آزادی می‌باشد برای نمونه‌های بزرگتر از ۳۰ می‌توان از جدول توزیع نرمال استاندارد برای تعیین عدد بحرانی و به دنبال آن ناحیه بحرانی (ناحیه رد H_0) استفاده کرد.
طبیعی است که در حالتی که تعداد نمونه کوچکتر یا مساوی ۳۰ است، برای تعیین عدد بحرانی باید از جدول توزیع t با توجه به درجه آزادی مربوطه که برابر با $df = n-1$ می‌باشد استفاده کرد.

مثال:
مدیریت نیروی انسانی یک سازمان دولتی اعلام می‌کند بررسی نمونه‌ای ۹ تایی از کارکنان آن سازمان در رابطه با تعداد روزهای مرخصی استعلاجی حاکی از آن است که توزیع تعداد روزهای استفاده از مرخصی استعلاجی نرمال با میانگین ۹ روز در سال و انحراف معیار ۳/۵ روز می‌باشد. آیا براساس این نتایج می‌توان در سطح خطای ۱٪ ادعا کرد، میانگین تعداد روزهای مرخصی استعلاجی کارکنان

آن است که حداقل ۸ روز در سال است؛ نتیجه آزمون در سطوح خطای ۵٪ و ۱۰٪ چیست؟

در این مسئله با توجه به ادعای مطرح شده فرضیه‌های صفر و یک عبارتند از:

$H_0: \mu \leq 8$
$H_1: \mu > 8$

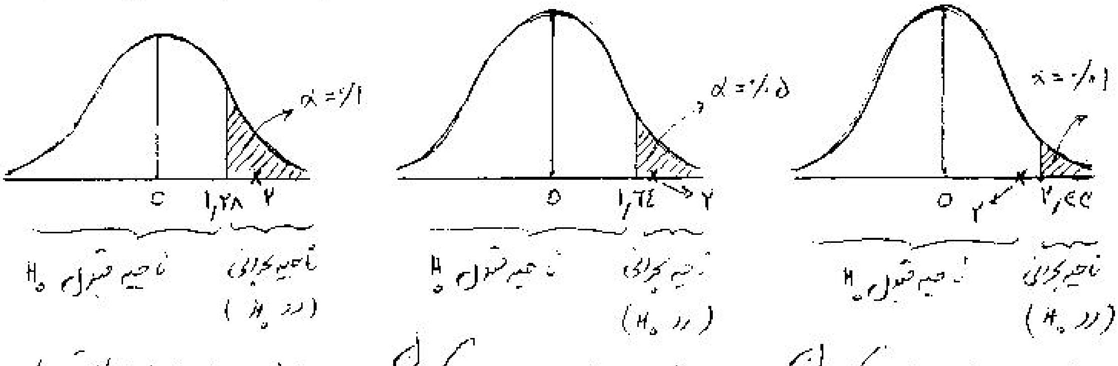
از آنجایی که تعداد نمونه بزرگتر از ۳۰ می‌باشد، علی‌رغم اینکه شاخص آماری آزمون از نظر تئوری t با $n=49-1=48$ درجه آزادی می‌باشد، اما به خاطر نرمال بودن توزیع t، می‌توان از سطوح معنی‌دار در جدول نرمال استاندارد نیز استفاده کرد.

$n = 49 \qquad \bar{x} = 9 \qquad S = 3.5$

بنابراین مقدار شاخص آماری آزمون برابر است با:

$$t = \frac{\bar{x} - \mu_{H_0}}{S/\sqrt{n}} = \frac{9-8}{3.5/\sqrt{49}} = 2$$

با توجه به خطاهای ۱٪، ۵٪ و ۱۰٪ یا به عبارت دیگر $\alpha = 0.01$، $\alpha = 0.05$ و $\alpha = 0.1$ و با توجه به سطوح معنی‌دار در جدول نرمال استاندارد، در هر کدام از حالات فوق عدد بحرانی و به دنبال آن ناحیه بحرانی به صورت زیر خواهد بود.



چون ۲>۱٫۲۸ است پس آزمون در سطح خطای $\alpha = 0.1$ معنی‌دار است. یعنی فرض H_0 رد و H_1 یا ادعای محقق تأیید می‌شود.

چون ۲>۱٫۶۴، پس آزمون معنی‌دار است یعنی در سطح خطای $\alpha = 0.05$ فرض H_0 رد و H_1 یعنی ادعای محقق تأیید می‌شود.

چون ۲<۲٫۳۳، پس آزمون در سطح خطای $\alpha = 0.01$ معنی‌دار نیست. به عبارت دیگر در سطح خطای ۱٪ فرض صفر رد نمی‌شود، یا فرض H_1 یعنی ادعای محقق تأیید نمی‌شود.

دنباله حل مثال:

این مثال در سه سطح خطای $\alpha=/1$ ، $\alpha=/05$ ، $\alpha=/01$ حل و نتایج در هر مرحله بیان شد. در اینجا لازم است با توجه به نتایج بدست آمده، به چند نکته مهم که در همه آزمونهای آماری برقرار است اشاره شود:

- اگر آزمونی در سطح خطای α معنی‌دار نبود، ممکن است در سطح خطای بیشتری معنی‌دار شود، همچنانکه در مثال فوق آزمون در سطح خطای $\alpha=/01$ معنی‌دار نبود، امّا در سطوح خطای بزرگتر یعنی $\alpha=/05$ و $\alpha=/1$ معنی‌دار است.
- اگر آزمونی در سطح خطای α معنی‌دار باشد، در هر سطح خطای بیشتر نیز معنی‌دار خواهد بود. همچنانکه آزمون در سطح خطای $\alpha=/05$ معنی‌دار است، پس در سطح خطای $\alpha=/1$ نیز معنی‌دار خواهد بود.
- اگر آزمونی در سطح خطای α معنی‌دار نباشد، در هر سطح خطای کمتر نیز معنی‌دار نخواهد بود. همچنانکه در مثال فوق آزمون در سطح خطای $\alpha=/01$ معنی‌دار نبود، در هر سطح خطای کمتر از $/01$ نیز معنی‌دار نخواهد بود.
- اگر آزمونی در سطح خطای α معنی‌دار باشد، ممکن است در سطح خطای کمتر معنی‌دار نباشد. همچنانکه در مثال فوق آزمون در سطح خطای $\alpha=/05$ معنی‌دار است ولی در سطح خطای کمتر یعنی $\alpha=/01$ معنی‌دار نیست.

از آنجائیکه همیشه سطح معنی‌دار یا خطای مورد پذیرش محقق قبل از انجام مراحل اجرایی آزمون مشخص نمی‌شود، برای جلوگیری از تغییر سطح معنی‌دار و یا بیان نتیجه آزمون در سطوح مختلف خطای نوع اول (α)، معمولاً نتیجه آزمون با مقداری به نام P-value بیان می‌شود. ذکر این نکته ضروری است که چون ادعای محقق در قالب فرض یک یا H_1 بیان می‌شود، محقق همیشه به دنبال تأیید H_1 یا به عبارت دیگر رد H_0 است

و این خود انگیزه‌ای است برای محقق که چنانچه در سطح خطای مشخص آزمون معنی‌دار نباشد یعنی فرض صفر رد نشود، با تغییر سطح معنی‌دار یا به عبارت دیگر با افزایش میزان خطا سعی در معنی‌دار کردن آزمون کند که این کار بطور کلی خالی از اشکال نیست.
بنابراین بهتر است بجای بیان نتیجه آزمون براساس یک مقدار خطای نوع اول مشخص، مثلاً $\alpha = 0.05$ یا هر مقدار دیگر α، نتیجه براساس یک سطح معنی‌دار بنام P-Value بیان شود:

: P-value

همچنانکه گفته شد، در تحقیقات علمی نتیجه آزمون آماری براساس یک مقدار بنام P-value بیان می‌شود. P-value فارغ از خطای نوع اول مورد پذیرش محقق و صرفاً براساس نوع فرضیه (یک دنباله یا دو دنباله) و براساس مقدار شاخص آماری آزمون که خود نیز براساس نتیجه حاصل از یک نمونه n تایی بدست آمده است، محاسبه می‌شود. در این بخش صرفاً P-value، در حالتی که شاخص آماری آزمون از توزیع نرمال استاندارد پیروی می‌کند در حالات مختلف فرضیه‌های صفر و یک دقیقاً برای تفهیم این شاخص بطور مبسوط شرح داده می‌شود.
از آنجائیکه P-value براساس مقدار شاخص آماری آزمون و نوع فرضیه‌ها محاسبه می‌شود، برای شرح این مسئله، فرضیه‌ها و مقدار شاخص آماری آزمون در حالتی که توزیع آن نرمال استاندارد است، بصورت فرضی در نظر گرفته می‌شوند.

حالت اول:

فرض کنید فرضیه‌های صفر و یک همچنین شاخص آماری آزمون به صورت زیر باشد

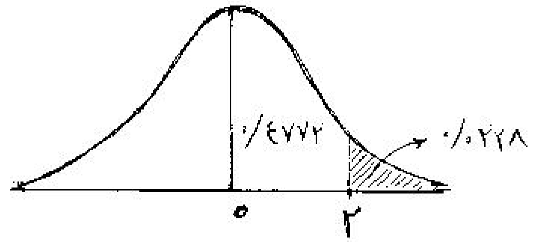
$H_0: \mu = a$

$H_1: \mu \neq a$

$$Z = \frac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}} = ۲$$

n > ۳۰ یا s در حالت

فرض کنید این مقدار را بر اساس داده‌های نمونه بدست آمده باشد. در این حالت با توجه به دو دنباله بودن فرضیه‌ها مقدار P برابر خواهد بود با:



۰/۵ − ۰/۴۷۷۲ = ۰/۰۲۲۸

مقدار ۰/۴۷۷۲ با استفاده از جدول توزیع نرمال استاندارد بدست می‌آید

با توجه به دو دنباله بودن فرضیه‌ها P-value = ۰/۰۲۸۸ × ۲ = ۰/۰۴۵۶

همچنانکه مشاهده می‌شود، در این حالت به جهت دنباله‌های راست و چپ که بر اساس مقدار شاخص آماری آزمون تعیین می‌شود در عدد ۲ ضرب می‌شود و مقدار P بدست می‌آید.

حالت دوم:

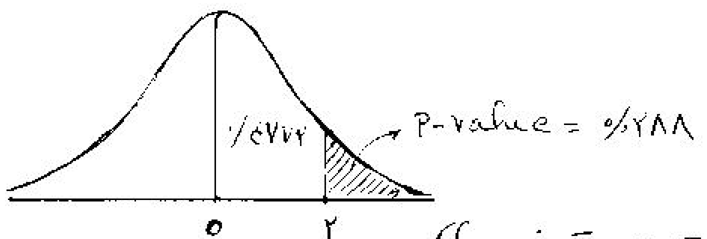
فرض کنید فرضیه‌های صفر و یک و همچنین مقدار شاخص آماری آزمون به صورت زیر باشد.

$H_0: \mu \leq a$

$H_1: \mu > a$

$$Z = \frac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}} = ۲$$

n ≥ ۳۰ برای s یا



در این حالت با توجه به جهت علامت نامساوی در فرض یک (>)، دنباله سمت راست یعنی سطح زیر منحنی از نقطه ۲ (مقدار شاخص آماری آزمون که در اینجا برای نمونه n تایی بدست آمده) تا ∞+، مقدار P را تشکیل می‌دهد

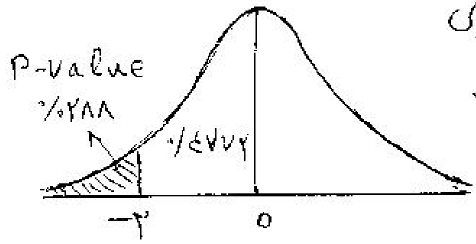
حالت دوم:

فرض کنید فرضیه‌های آماری یک طرفه و همچنین مقدار شاخص آماری آزمون به صورت زیر باشد:

$$H_0: \mu \geq a \qquad H_1: \mu < a$$

$$Z = \frac{\bar{x} - \mu_{H_0}}{\sigma / \sqrt{n}} = -۲$$

یا s برای $n>۳۰$



در این حالت نیز با توجه به جهت علامت نامساوی در فرض یک ($<$)، دنباله سمت چپ منحنی سطح زیر منحنی از نقطه ۲- (مقدار شاخص آماری آزمون که بر اساس نتیجه حاصل از نمونه n تایی محاسبه شده است) تا $-\infty$، مقدار P را تشکیل می‌دهد.

لازم به توضیح است که P-value می‌تواند از ۰ تا ۱+ تغییر کند. در آزمون‌های یک دنباله، بر اساس مقدار شاخص آماری آزمون و با توجه به جهت علامت نامساوی که به صورت "$<$" یا "$>$" می‌باشد، مساحت سمت راست زیر منحنی از مقدار شاخص آماری آزمون یا مساحت سمت چپ زیر منحنی از مقدار شاخص آماری آزمون، P-value را تشکیل می‌دهد.

حال محقق می‌تواند با توجه به خطای نوع اول مورد نظر خود و مقایسه آن با P-value در مورد آزمون نتیجه‌گیری کند که نتیجه‌گیری نهایی با توجه به قواعد زیر خواهد بود.

- اگر $P\text{-value} \geq \alpha$ باشد، در این صورت آزمون معنی‌دار نخواهد بود و فرضیه H_0 رد و H_1 تأیید می‌شود.
- اگر $P\text{-value} < \alpha$ باشد، در این صورت آزمون معنی‌دار است و ادعای محقق که در قالب H_1 بیان شده، مورد تأیید قرار می‌گیرد.

لازم به توضیح است که اگر آزمون آماری و مطالب کلی تجزیه و تحلیل آماری توسط یک نرم افزار آماری مانند SPSS انجام گیرد، در خروجی نرم افزار، نتایج آزمونهای انجام شده بوسیله P-value بیان می شود و تصمیم گیری محقق در خصوص رد یا قبول H_0 یا به عبارت دیگر پذیرش و یا عدم پذیرش H_1، با توجه به خطای نوع اول (α) به عهده شخص محقق خواهد بود.

در مثال صفحه ۱، با توجه به مقدار شاخص آماری آزمون و اینکه به علت بزرگ بودن تعداد نمونه می توان از جدول توزیع نرمال استاندارد استفاده کرد، $P\text{-value} = ۰/۰۲۲۸$ خواهد بود بنابراین هر مقدار α بزرگتر از ۰/۰۲۲۸، مثل $\alpha = ۰/۰۳$ یا $\alpha = ۰/۰۵$ آزمون معنی دار تلقی خواهد شد.

مثال:

نتایج حاصل از یک نمونه ۲۵ تایی از یک نوع عمل جراحی خاص در رابطه با هزینه های مربوط به آن نشان می دهد که توزیع هزینه ها نرمال با میانگین ۸۸۴ هزار تومان و انحراف معیار ۴۸ هزار تومان می باشد. آیا بر اساس این نتایج می توان در سطح خطای $\alpha = ۰/۰۵$ ادعا کرد میانگین هزینه های این عمل جراحی حداکثر ۹۰۰ هزار تومان است؟

در این مثال با توجه به اینکه تعداد نمونه کمتر از ۳۰ و انحراف معیار هزینه ها از نمونه ۲۵ تایی بدست آمده است، بنابراین شاخص آماری آزمون از توزیع t با $n-1$ درجه آزادی برخوردار است.

در این مثال فرضیه های صفر و یک به این صورت خواهند بود

$$H_0: \mu \geq ۹۰۰$$
$$H_1: \mu < ۹۰۰$$

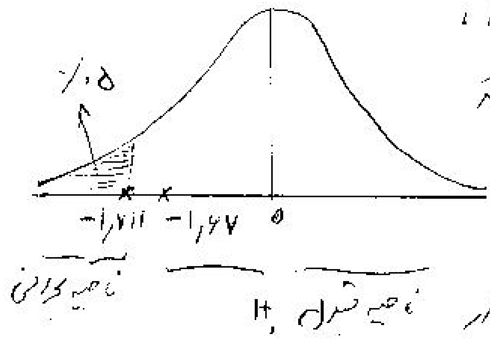
تعداد نمونه، میانگین و انحراف معیار حاصل از نمونه (شاخصهای آماری) برابر هستند با:

$$n = ۲۵, \quad \bar{x} = ۸۸۴, \quad S = ۴۸$$

شاخص آماری آزمون دارای توزیع t با $df = ۲۵ - ۱ = ۲۴$ برابر است با

$$t = \frac{\bar{x} - \mu_{H_0}}{S/\sqrt{n}} = \frac{۸۸۴ - ۹۰۰}{۴۸/\sqrt{۲۵}} = -۱/۶۷$$

حال با توجه به $\alpha = ۰/۰۵$ و جدول توزیع t با ۲۴ درجه آزادی می‌توان ناحیه بحرانی (ناحیه رد H_0) و ناحیه قبول H_0 را تعیین کرد.



با توجه به جهت علامت نامساوی در فرض H_1، نقطه t روی محور افقی را باید طوری پیدا کرد که مساحت زیر منحنی دنباله سمت چپ برابر با ۰/۰۵ % باشد که با استفاده از جدول توزیع t، این مقدار برابر خواهد بود با ۱/۷۱۱-

از آنجایی که ۱/۶۷- < ۱/۷۱۱- است یعنی در ناحیه بحرانی قرار ندارد، می‌توان نتیجه گرفت که بر اساس اطلاعات داده شده دلیل کافی برای رد H_0 وجود ندارد، یعنی آزمون معنی‌دار نیست و فرض محقق یعنی H_1 تأیید نمی‌شود. بنابراین نمی‌توان این ادعا را تأیید کرد که میانگین هزینه‌های این محل حداکثر ۹۰ هزار تومان است.