آزمونها

آزمون ۱ : آزمون z برای میانگین یک جامعه (واریانس معلوم)

موضوع : تعیین معنی‌دار بودن اختلاف بین میانگین مفروض یک جامعه μ_0 و میانگین یک نمونه $\bar{x}$

شرایط :

۱ـ معلوم بودن واریانس σ^2 ضروری است. (اگر σ^2 معلوم نیست به آزمون t برای میانگین یک جامعه مراجعه شود.)

۲ـ چنانچه جامعه نرمال توزیع شده باشد آزمون دقت لازم را دارد. ولی اگر جامعه نرمال نباشد باز هم آزمون راهنمای خوبی خواهد بود.

روش :

از جامعه‌ای با میانگین مفروض μ_0 و واریانس معلوم σ^2، نمونه‌ای تصادفی به حجم n گرفته و میانگین نمونه $\bar{x}$ را محاسبه می‌کنیم.

آماره آزمون عبارتست از:

$$Z = \frac{(\bar{x} - \mu_0)}{\sigma/\sqrt{n}}$$

که می‌تواند با توزیع نرمال استاندارد (به هر یک از دو صورت آزمون یک دامنه یا دو دامنه) با ناحیه بحرانی در سطح α مقایسه شود.

مثال عددی :

$\mu = ۴$ و $n = ۹$ و $\bar{x} = ۴/۶$ و $\sigma = ۱$ $\qquad z = ۱/۸$

با توجه به جدول شماره ۱ مقدار بحرانی برابر $z_{\alpha(۰/۰۵)} = ۱/۹۶$ می‌شود که با استفاده از آزمون دو دامنه ($H_0 : \mu = \mu_0$ و $H_1 : \mu \neq \mu_0$) فرضیه صفر H_0 رد نمی‌شود. اما با استفاده از آزمون یکدامنه ($H_0 : \mu = \mu_0$ و $H_1 : \mu > \mu_0$) این فرضیه رد می‌شود.

آزمون ۲: آزمون z برای میانگین دو جامعه (واریانسهای معلوم و مساوی)

موضوع: تعیین معنی‌دار بودن اختلاف میانگین دو جامعه

شرایط:

۱ـ هر دو جامعه باید واریانسهای مساوی داشته باشند و این واریانس برای بایستی معلوم باشد. (اگر σ^2 معلوم نیست به آزمون t میانگین‌های دو جامعه مراجعه شود.)

۲ـ چنانچه جامعه‌ها نرمال توزیع شده باشند آزمون دقت لازم را دارد. ولی اگر جامعه‌ها نرمال نباشند باز هم آزمون تقریباً مناسب خواهد بود.

روش:

دو جامعه با میانگینهای μ_1 و μ_2 در نظر گرفته و نمونه‌های تصادفی مستقل به حجم n_1 و n_2 انتخاب و سپس میانگینهای نمونه $\bar{x}_1$ و $\bar{x}_2$ را بدست می‌آوریم.
آماره آزمون عبارتست از:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_x \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

که می‌تواند با توزیع نرمال استاندارد (به هر یک از دو صورت آزمون یک دامنه یا دو دامنه) مقایسه شود.

مثال عددی:

$\sigma^2 =$ ۲/۰۷۵۰ و $\sigma =$ ۱/۴۴۰۵ و $\bar{x}_2 =$ ۱/۷ و $\bar{x}_1 =$ ۱/۲ و $n_2 =$ ۱۶ و $n_1 =$ ۹

$z =$ ۰/۸۳۳

با توجه به جدول شماره ۱ مقدار بحرانی برابر ۱/۹۶ = z_α (۰/۰۵) است که با استفاده از آزمون دو دامنه ($H_0: \mu_1 - \mu_2 = 0$ ؛ $H_1 = \mu_1 - \mu_2 \neq 0$) و نیز آزمون یکدامنه ($H_0: \mu_1 - \mu_2 = 0$ ؛ $H_1 = \mu_1 - \mu_2 > \mu_0$) فرضیه H_0 رد نمی‌شود.

آزمون ۳ : آزمون Z برای میانگین دو جامعه (واریانس‌های معلوم و نامساوی)

موضوع: تعیین معنی‌دار بودن اختلاف میانگین دو جامعه

شرایط:

۱ـ معلوم بودن واریانس‌های دو جامعه ضروری است.(اگر واریانسها معلوم نیستند به آزمون t برای میانگین دو جامعه مراجعه شود.)

۲ـ چنانچه جامعه‌ها نرمال توزیع شده باشد آزمون دقت لازم را دارد. ولی اگر جامعه‌ها نرمال نباشند باز هم آزمون تقریباً مناسب می‌باشد.

روش:

دو جامعه با میانگینهای $\mu_۱$ و $\mu_۲$ و واریانسهای $\sigma_۱^۲$ و $\sigma_۲^۲$ در نظر گرفته و نمونه‌های تصادفی مستقل به حجم $n_۱$ و $n_۲$ انتخاب و سپس میانگینهای نمونه $\bar{x}_۱$ و $\bar{x}_۲$ را محاسبه می‌کنیم. آماره آزمون عبارتست از:

$$Z=\frac{(\bar{x}_۱-\bar{x}_۲)-(\mu_۱-\mu_۲)}{\sqrt{\frac{\sigma_۱^۲}{n_۱}+\frac{\sigma_۲^۲}{n_۲}}}$$

مثال عددی:

$n_۱ = ۱۳$ و $n_۲ = ۸$ و $\bar{x}_۱ = ۸۰/۰۲$ و $\bar{x}_۲ = ۷۹/۹۸$ و $\sigma_۱^۲ = ۰/۰۰۰۵۷۶$ و

$\sigma_۲^۲ = ۰/۰۰۱۰۸۹$ و $z = ۲/۹۸$

با توجه به جدول شماره ۱ مقدار بحرانی $z_\alpha\ (۰/۰۵) = ۱/۹۶$ است. بنابراین فرضیه صفر رد می‌شود

آزمون ۴: آزمون z برای یک نسبت (توزیع دو جمله‌ای)

موضوع: تعیین معنی‌دار بودن اختلاف بین یک نسبت مفروض p_0 و یک نسبت مشاهده شده p.

شرایط:
این آزمون تقریبی است و چنین فرض می‌کند که تعداد مشاهدات نمونه به اندازه کافی بزرگ می‌باشد (یعنی $n = ۳۰$) تا بتواند تقریب نرمال را بجای توزیع دو جمله بکار برد.

روش:
یک نمونه تصادفی n تایی از جامعه‌ای، که در آن نسبت p_0 متعلق به یک گروه خاص باشد، را انتخاب و نسبت p در آن گروه را محاسبه می‌کنیم.
آماره آزمون عبارت است از:

$$Z = \frac{P - P_0}{\sqrt{\frac{P_0(1 - P_0)}{n}}}$$

مثال عددی:

$z = -۲$ و $p_0 = ۰/۵$ و $p = ۰/۴$ و $n = ۱۰۰$

با توجه به جدول شماره ۱ مقدار بحرانی برابر $z_{\alpha\,(۰/۰۵)} = \pm ۱/۹۶$ است.
بنابراین فرضیه صفر رد می‌شود.

توضیح:
با توجه به متقارن بودن توزیع نرمال، در مواردیکه آماره آزمون (z) محاسبه شده منفی باشد میتوان از قدر مطلق آن استفاده کرده و این قدر مطلق را با مقادیر مثبت جدول ۱ مقایسه نمود و در صورت بزرگتر بودن آماره آزمون، فرضیه صفر را رد نمود.

آزمون ۵: آزمون z برای تساوی دو نسبت (توزیع دو جمله‌ای)

موضوع : تحقیق در مورد تساوی دو نسبت $\pi_۱$ و $\pi_۲$ از افراد جامعه، برمبنای نمونه‌ای از هر جامعه

شرایط :

این آزمون تقریبی است و چنین فرض می‌کند که تعداد مشاهدات هر دو نمونه به اندازه کافی بزرگ می‌باشند (یعنی $n_۱$ و $n_۲ \geq ۳۰$) تا بتواند تقریب نرمال را بجای توزیع دو جمله‌ای بکار برد.

روش :

فرض می‌شود که $\pi_۱$ و $\pi_۲$ نسبتهای یک صفت مشخص در دو جامعه باشند. نمونه‌های تصادفی به حجم $n_۱$ و $n_۲$ انتخاب و نسبت‌های متناظر $p_۱$ و $p_۲$ را محاسبه می‌کنیم. آماره آزمون عبارتست از:

$$Z = \frac{(P_۱ - P_۲)}{\sqrt{P(۱-P)\left[\frac{۱}{n_۱} + \frac{۱}{n_۲}\right]}}$$

که در آن :

$$P = \frac{P_۱ n_۱ + P_۲ n_۲}{n_۱ + n_۲}$$

تحت فرضیه صفر که $\pi_۱ = \pi_۲$ است، z تقریباً بصورت نرمال استاندارد توزیع شده است و آزمون حاصله می‌تواند به هر یک از دو صورت آزمون یک دامنه یا دو دامنه باشد.

مثال عددی:

$n_۱ = ۹۵۲$ و $n_۲ = ۱۱۶۸$ و $p_۱ = ۰/۰۰۳۲۵$ و $p_۲ = ۰/۰۵۷۳$

$z = -۲/۷۳$

با توجه به جدول شماره ۱ مقدار بحرانی برابر $z_{\alpha\,(۰/۰۵)} = \pm ۱/۹۶$ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۶: آزمون Z برای مقایسه دو مقدار شمارشی (توزیع پواسن)

موضوع: تعیین معنی‌دار بودن اختلاف بین دو مقدار شمارشی

شرایط:

این آزمون تقریبی است و چنین فرض می‌کند که تعداد نمونه به اندازه کافی بزرگ می‌باشد تا بتواند تقریب نرمال را بجای پواسن به کار برد.

روش:

فرض می‌شود که $N_۱$ و $N_۲$ بترتیب تعداد موارد مشاده شده در فاصله‌های زمانی $t_۱$ و $t_۲$ می‌باشند. بنابراین فراوانی‌های متوسط برابر $R_۱ = \frac{N_۱}{t_۱}$ و $R_۲ = \frac{N_۲}{t_۲}$ هستند. بمنظور آزمودن فرضیه مساوی بودن فراوانیهای متوسط، آماره آزمون زیر مورد استفاده قرار می‌گیرد.

$$Z = \frac{R_۱ - R_۲}{\sqrt{\frac{R_۱}{t_۱} + \frac{R_۲}{t_۲}}}$$

که می‌تواند با توزیع نرمال استاندارد (به هر یک از دو صورت آزمون یک دامنه یا دو دامنه) مقایسه گردد.

مثال عددی:

$N_۱ = ۹۵۲$ و $N_۲ = ۱۱۶۸$ و $t_۱ = ۲۲$ و $t_۲ = ۳۰$

$$R_۱ = \frac{N_۱}{t_۱} = ۴۳/۲۷ \quad و \quad R_۲ = \frac{N_۲}{t_۲} = ۳۸/۹۳$$

$$Z = \frac{۴/۴۸۰۶}{\sqrt{۳/۲۶۵}} = \frac{۴/۴۸۰۶}{۱/۸۰۷} = ۲/۴۷۹$$

با توجه به جدول شماره ۱ مقدار بحرانی برابر $z_{\alpha(۰/۰۵)} = ۱/۹۶$ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۷ : آزمون t برای میانگین یک جامعه (واریانس نا معلوم)

موضوع: تعیین معنی‌دار بودن اختلاف بین میانگین مفروض یک جامعه $\mu_{\circ}$ و میانگین یک نمونه $\bar{x}$.

شرایط:

۱ـ اگر واریانس جامعه σ^2 معلوم باشد، آزمون قوی‌تری وجود دارد. در این صورت به آزمون z برای میانگین یک جامعه مراجعه شود.

۲ـ چنانچه جامعه نرمال توزیع شده باشد آزمون دقت لازم را دارد. ولی اگر جامعه نرمال نباشد، باز هم آزمون راهنمای خوبی خواهد بود.

روش :

از جامعه‌ای با میانگین مفروض $\mu_{\circ}$ و واریانس نامعلوم، نمونه‌ای تصادفی به حجم n انتخاب و سپس میانگین نمونه $\bar{x}$ را محاسبه و نیز انحراف معیار نمونه را با استفاده از فرمول زیر بدست می‌آوریم:

$$S = \left\{ \frac{\sum (x_i - \bar{x})^۲}{n - ۱} \right\}^{\frac{۱}{۲}}$$

آماره آزمون عبارت است از :

$$t = \frac{(\bar{x} - \mu_{\circ})}{s / \sqrt{n}}$$

که می‌تواند با توزیع t استودنت با $(n - ۱)$ درجه آزادی (DF) مقایسه گردد. آزمون می‌تواند به هر یک از دو صورت یک دامنه یا دو دامنه انجام گیرد.

مثال عددی :

$t=-۲/۷$ و $\nu=n-۱$ و $s=۱$ و $\bar{x}=۳/۱$ و $n=۹$ و $m_{\circ}=۴$

با توجه به جدول شماره ۲ مقدار بحرانی برابر $t_{۸\,(۰/۰۲۵)}=\pm ۲/۳$ است. بنابراین با استفاده از آزمون دو دامنه ($H_{\circ}:\mu=\mu_{\circ}$ و $H_{۱}:\mu\neq\mu_{\circ}$) فرضیه صفر رد می‌شود.

آزمون ۸: آزمون t برای میانگین دو جامعه (واریانس‌های نامعلوم ولی مساوی)

موضوع: تعیین معنی‌دار بودن اختلاف میانگین دو جامعه

شرایط:

۱ـ اگر واریانس جامعه‌ها معلوم باشد، آزمون قوی‌تری وجود دارد. در این صورت به آزمون z برای میانگین دو جامعه مراجعه شود.

۲ـ چنانچه جامعه‌ها نرمال توزیع شده باشند آزمون دقت لازم را دارد. ولی اگر جامعه‌ها نرمال نباشند باز هم آزمون راهنمای خوبی خواهد بود.

روش:

دو جامعه با میانگین‌های $\mu_۱$ و $\mu_۲$ در نظر گرفته و نمونه‌های تصادفی مستقل به حجم $n_۱$ و $n_۲$ را انتخاب و میانگین‌های نمونه‌ای $\bar{x}_۱$ و $\bar{x}_۲$ را، همراه با مجموع مربّعات

$$ss_۱ = \sum_{i=۱}^{n_۱} (x_i - \bar{x}_۱)^۲$$

$$ss_۲ = \sum_{i=۱}^{n_۲} (x_i - \bar{x}_۲)^۲$$

محاسبه می‌کنیم. بهترین برآورد واریانس جامعه با $S = (ss_۱ + ss_۲)/(n_۱ + n_۲ - ۲)$ بدست می‌آید.

آماره آزمون عبارتست از:

$$t = \frac{(\bar{x}_۱ - \bar{x}_۲) - (\mu_۱ - \mu_۲)}{\sqrt{S\left[\frac{۱}{n_۱} + \frac{۱}{n_۲}\right]}}$$

که می‌تواند با توزیع t استیودنت با درجه آزادی $(n_۱ + n_۲ - ۲)$ مقایسه شود. آزمون می‌تواند به هر یک از دو صورت یک دامنه یا دو دامنه انجام شود.

مثال عددی :

$n_1 = 12$ و $n_2 = 12$ و $\bar{x}_1 = 31/75$ و $\bar{x}_2 = 28/67$

$\nu = n_1 + n_2 - 2$

$ss_1 = 112/25$ و $ss_2 = 66/64$ و $s = 2/85$

$t = 2/65$ و $\nu = 12 + 12 - 2 = 22$

با توجه به جدول شماره ۲ مقدار بحرانی برابر $t_{22(0/025)} = 2/07$ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۹ : آزمون t برای میانگین دو جامعه (واریانس‌های نامعلوم و نامساوی)

موضوع : تعیین معنی‌دار بودن اختلاف میانگین دو جامعه

شرایط :

۱ـ اگر واریانس‌های جامعه معلوم باشند آزمون قوی‌تری نیز وجود دارد. در این صورت به آزمون z برای میانگین دو جامعه مراجعه شود (واریانس‌های معلوم و نامساوی).

۲ـ چنانچه جامعه‌ها نرمال توزیع شده باشند یا حجم نمونه‌ها به اندازه کافی بزرگ باشند آزمون مناسب است.

۳ـ آزمون می‌تواند فقط برای آزمودن فرضیه $\mu_1 = \mu_2$ مورد استفاده قرار گیرد.

روش :

دو جامعه با میانگین‌های μ_1 و μ_2 در نظر گرفته و نمونه‌های تصادفی مستقل به حجم n_1 و n_2 را انتخاب و سپس میانگین‌های نمونه $\bar{x}_1$ و $\bar{x}_2$ و همچنین واریانس‌های $S_1^2 = \frac{\sum(x_i - \bar{x}_1)^2}{n_1 - 1}$ و $S_2^2 = \frac{\sum(x_i - \bar{x}_2)^2}{n_2 - 1}$ را محاسبه می‌کنیم.

آماره آزمون عبارتست از:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

که می‌تواند با توزیع t استیودنت با درجه آزادی حاصله از فرمول زیر مقایسه شود.

$$\nu = \frac{\left\{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right\}}{\frac{s_1^4}{n_1^2(n_1 + 1)} + \frac{s_2^4}{n_2^2(n_2 + 1)}} - 2$$

مثال عددی:

$n_1 = ۴$ و $n_2 = ۹$ و $\bar{x}_1 = ۳۱۶۶/۰$ و $\bar{x}_2 = ۲۲۴۰/۴$

$s_1^2 = ۶۳۲۸/۶۷$ و $s_2^2 = ۲۲۱۶۶۱/۳$ و $t = ۵/۷۲$

$\nu = ۹$ (تقریباً)

با توجه به جدول شماره ۲ مقدار بحرانی برابر $t_{9(0/025)} = ۲/۲۶$ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۱۰: آزمون t برای میانگین دو جامعه (روش مقایسه‌های جفت شده)

موضوع: تعیین معنی‌دار بودن اختلاف میانگین دو جامعه $\mu_۱$ و $\mu_۲$. هیچ پیش فرضی هم در مورد واریانس‌های جامعه وجود ندارد.

شرایط:

۱ـ مشاهدات دو نمونه باید بصورت جفت جفتی بدست آمده باشند. صرف‌نظر از اختلاف دو جامعه، هر جفت مشاهده باید تحت شرایط یکسان یا تقریباً یکسان بدست آمده باشند.

۲ـ چنانچه جامعه‌ها نرمال توزیع شده باشند آزمون دقت لازم را دارد. ولی اگر جامعه‌ها نرمال نباشند، باز هم آزمون راهنمای خوبی خواهد بود.

روش:

تفاضل‌های d_i را برای هر جفت از مشاهدات محاسبه می‌کنیم. اگر n جفت مشاهده وجود داشته باشد، می‌توانیم واریانس تفاضل‌ها را با

$$s^۲ = \sum_{i=۱}^{n} \frac{\left(d_i - \bar{d}\right)^۲}{n-۱}$$

محاسبه می‌کنیم. چنانچه میانگین نمونه‌های دو جامعه را $\bar{x}_۱$ و $\bar{x}_۲$ فرض کنیم. آنگاه آماره آزمون برابر:

$$t = \frac{(\bar{x}_۱ - \bar{x}_۲) - ۰}{s/\sqrt{n}}$$

می‌باشد که از توزیع t استیودنت با درجه آزادی $(n - ۱)$ پیروی میکند. آزمون میتواند به هر یک از دو صورت آزمون یکدامنه یا دو دامنه انجام شود.

مثال عددی:

$\bar{d} = -۰/۱ \quad , \quad n = ۱۰ \quad , \quad \nu = n-۱ \quad , \quad t = ۰/۱۱ \quad , \quad \nu = ۹$

با توجه به جدول شماره ۲ مقدار بحرانی برابر با $t_{۹(۰/۰۲۵)} = ۲/۲۶$ است. بنابراین فرضیه صفر رد نمی‌شود.

آزمون ۱۱: آزمون t برای ضریب رگرسیون

موضوع: تعیین معنی‌دار بودن ضریب رگرسیون y روی x

شرایط:

فرض‌های متداول رگرسیون باید بکار برده شود یعنی:

۱ـ متغیر y بازای هر مقدار x دارای توزیع نرمال باشد.

۲ـ بازای هر مقدار مفروض x، واریانس y ثابت بماند.

روش:

بمنظور برآورد رگرسیون خطی بصورت $y = A + B\ (x_i - \bar{x})$، که در آن B ضریب رگرسیون نامیده می‌شود، نمونه‌ای n جفتی از مقادیر (x_i و y_i) در نظر می‌گیریم. بمنظور آزمودن فرضیه صفر ($H_0 : B = 0$) مقدار برآورد نمونه‌ای ضریب B را با استفاده از فرمول زیر بدست می‌آوریم.

$$b=\frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \left[\frac{1}{n}\sum x_i\right]^2}$$

واریانس xها و واریانس yها حول خط رگرسیون چنین محاسبه می‌شود.

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

$$S_{y.x}^2 = \frac{\sum \{ y_i - y - b\ (x_i - \bar{x})\ \}^2}{n-2}$$

که در آن $\bar{x}$ و $\bar{y}$ به ترتیب میانگین xها و yها هستند.

آماره آزمون برابر:

$$t = \frac{b.\ s_x}{\frac{s_{yx}}{\sqrt{n-1}}}$$

است که از توزیع t استودنت با درجه آزادی ۲ - n پیروی می‌کند. از آنجا که b می‌تواند مقادیر مثبت و یا منفی را اختیار کند، آزمون باید بصورت دو دامنه باشد.

مثال عددی:

$\sum y_i = ۱۷۰۰$ و $\sum x_i^۲ = ۴۹۰۶$ و $\sum x_i = ۷۷۶$

$n = ۱۲$ و $\sum x_i y_i = ۱۰۹۳۸۰$ و $\sum y_i^۲ = ۲۴۶۱۰۰$

$\nu = n - ۲$ و $\bar{y} = ۱۴۱/۶۷$ و $\bar{x} = ۶۳/۸۳$

$s_{y.x}^۲ = ۹۲/۸$ و $S_y^۲ = ۴۷۸/۸$ و $s_x^۲ = ۱۵/۶۱$

$\nu = ۱۰$ و $t = ۶/۸۶$ و $b = ۵/۰۲۹$

با توجه به جدول شماره ۲ مقدار بحرانی برابر $t_{۱۰,\ (۰/۰۲۵)} = ۲/۲۳$ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۱۲ : آزمون t برای ضریب همبستگی

موضوع : تعیین وجود اختلاف معنی‌دار آماری بین ضریب همبستگی نمونه با صفر.

شرایط :

چنین فرض می‌شود که x و y دارای توزیع نرمال دو متغیره بوده و رابطه آنها خطی است.

روش :

نمونه‌ای n جفتی از مقادیر (x_i و y_i) را در نظر گرفته و ضریب همبستگی r که برآوردی از ضریب همبستگی جامعه یعنی ρ می‌باشد، را مطابق فرمول زیر محاسبه می‌کنیم:

$$r = \frac{\sum (x_i - \bar{x})\ (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^۲ \sum (y_i - \bar{y})^۲}}$$

بمنظور آزمودن فرضیه صفر، ($H_۰ ; \rho = ۰$) آماره آزمون را با استفاده از:

$$t = \frac{r}{\sqrt{۱ - r^۲}} \quad \sqrt{n - ۲}$$

محاسبه می‌کنیم که از توزیع t استودنت با درجه آزادی $n - ۲$ پیروی می‌کند. معمولاً این آزمون بصورت دو دامنه می‌باشد، ولی در مواردی خاص می‌تواند به صورت یک دامنه انجام شود.

توضیح :

بمنظور آزمودن ضریب همبستگی جامعه با یک مقدار مشخص غیرصفر، به آزمون z برای ضریب همبستگی مراجعه شود.

مثال عددی :

$$n = ۱۸ \quad , \quad r = ۰/۳۲ \quad , \quad v = n - ۲$$

$$t = \frac{r}{\sqrt{۱ - r^{۲}}} \sqrt{n - ۲} = \frac{۰/۳۲\sqrt{۱۶}}{\sqrt{۱ - (۰/۳۲)^{۲}}} = ۱/۳۵$$

با توجه به جدول شماره ۲ مقدار بحرانی برابر $t_{۱۶(۰/۰۵)} = ۱/۷۵$ است. بنابراین فرضیه صفر رد نمی‌شود.

توجه :

در این مورد متغیرهای xو y مستقل هستند.

آزمون ۱۳ : آزمون Z برای ضریب همبستگی

موضوع : تعیین معنی‌دار بودن اختلاف ضریب همبستگی و مقدار مشخص ρ_0.

شرایط :

۱ـ مقادیر x و y از توزیع نرمال گرفته شده باشد.

۲ـ واریانس مقادیر y مستقل از مقادیر x باشد.

۳ـ رابطه بین x و y خطی باشد.

زمانیکه این شرایط برقرار نباشد، باید به آزمون همبستگی رتبه‌ای کندال مراجعه شود.

روش :

با استفاده از r که بعنوان ضریب همبستگی در آزمون قبلی تعریف شده، مقدار z را محاسبه می‌کنیم.

$$Z_1 = \frac{1}{2} \log_e \left[\frac{1+r}{1-r}\right] = 1/1513 \log_{10} \left[\frac{1+r}{1-r}\right]$$

z_1 دارای توزیع تقریبی نرمال با میانگین μ_{z_1} و انحراف معیار σ_{z_1} می‌باشد که در آن

$$\mu_{z_1} = \frac{1}{2} \log_e \left[\frac{1+\rho_0}{1-\rho_0}\right] = 1/1513 \log_{10} \left[\frac{1+r}{1-r}\right]$$

$$\sigma_{z_1} = \frac{1}{\sqrt{n-3}}$$

آماره آزمون بصورت زیر است.

$$Z = \frac{z_1 - \mu_{z_1}}{\sigma_{z_1}}$$

مثال عددی :

$n = 24$, $\rho_0 = 0/5$, $\mu = 0/75$

$$\mu_{z_۱} = ۱/۱۵۱۳ \quad , \quad log۳ = ۰/۵۴۹۳$$

$$z_۱ = ۱/۱۵۱۳ \, log \left(\frac{۱ + ۰/۷۵}{۱ - ۰/۷۵}\right) = ۰/۹۷۳۰$$

$$\sigma_{z_۱} = ۰/۲۱۸۶ \quad , \quad Z = \frac{z_۱ - \mu_{z_۱}}{\sigma_{z_۱}} = ۱/۹۴$$

با توجه به جدول شماره ۱ مقدار بحرانی در سطح $\alpha = ۰/۱$ برابر ۱/۶۴ است. بنابراین فرضیه صفر رد می‌شود.

آزمون ۱۴ : آزمون Z برای دو ضریب همبستگی

موضوع : تعیین معنی‌دار بودن اختلاف بین ضرایب همبستگی دو جامعه مختلف.

شرایط :

۱ـ مقادیر x و y از توزیع نرمال گرفته شده‌اند.

۲ـ واریانس متغیر y مستقل از مقادیر x باشد.

۳ـ رابطه بین x و y خطی باشد.

روش :

با استفاده از نمادهای آزمون ۱۳، برای جامعه اول با حجم نمونه $n_۱$

$$Z_۱ = \frac{۱}{۲} \log_e\left(\frac{۱ + r_۱}{۱ - r_۱}\right) = ۱/۱۵۱۳ \log_{۱۰}\left(\frac{۱ + r_۱}{۱ - r_۱}\right)$$

را که دارای میانگین $\mu_{z_۱} = \frac{۱}{۲} \log_e\left(\frac{۱ + \rho_۱}{۱ - \rho_۱}\right)$ و انحراف معیار $\sigma_{z_۱} = \frac{۱}{\sqrt{n_۱ - ۳}}$ است را محاسبه نموده، همچنین به طریق مشابه آماره $Z_۲$ را برای جامعه دوم با حجم نمونه $n_۲$ بدست می‌آوریم.

آماره آزمون عبارتست از:

$$Z = \frac{(z_۱ - z_۲) - (\mu_{z_۱} - \mu_{z_۲})}{\sigma}$$

که در آن:

$$\sigma = \sqrt{\sigma_{z_۱}^۲ + \sigma_{z_۲}^۲}$$

است. بنابراین Z از توزیع نرمال با میانگین ۰ و واریانس ۱ پیروی می‌کند.

مثال عددی :

$$n_۱ = ۲۸ ,\ n_۲ = ۳۵ ,\ r_۱ = ۰/۵ ,\ r_۲ = ۰/۳۰ ,\ \alpha = ۰/۰۵$$

با استفاده از جدول ۴ داریم:

$$z_{۱} = ۱/۱۵۱۳\ log\left(\frac{۱ + r_{۱}}{۱ - r_{۱}}\right) = ۰/۵۴۹۳$$

$$z_{۲} = ۱/۱۵۱۳\ log\left(\frac{۱ + r_{۲}}{۱ - r_{۲}}\right) = ۰/۳۰۹۵$$

و

$$\sigma = \sqrt{\left[\frac{۱}{n_{۱} - ۳} + \frac{۱}{n_{۲} - ۳}\right]} = ۰/۲۶۶۹$$

$$z = \frac{۰/۵۴۹۳ - ۰/۳۰۹۵}{۰/۲۶۶۹} = ۰/۸۹۸۵$$

با توجه به جدول شماره ۱ مقدار بحرانی در سطح $\alpha = ۰/۰۵$ برابر ۱/۹۶ است. بنابراین فرضیه صفر رد نمی‌شود.

آزمون ۱۵ : آزمون χ^2 برای واریانس یک جامعه

موضوع : تعیین اختلاف بین واریانس یک نمونه s^2 و واریانس مفروض جامعه σ_0^2

شرایط :

چنین فرض می‌شود جامعه‌ای که نمونه از آن استخراج شده از توزیع نرمال پیروی می‌کند.

روش :

یک نمونه n تایی x_1 و x_2 و ... و x_n در نظر گرفته و مقادیر

$$\bar{x} = \frac{\Sigma x_i}{n} \qquad s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n-1}$$

را محاسبه می‌کنیم. بمنظور آزمون فرضیه صفر، $H_0 : s^2 = \sigma_0^2$

آماره آزمون $\frac{(n-1)\, s^2}{\sigma^2}$ از توزیع کی دو χ^2 با درجه آزادی ۱ - n = ν پیروی می‌کند.
آزمون می‌تواند به هر یک از دو صورت یک دامنه یا دو دامنه انجام شود.

مثال عددی :

$$\bar{x} = ۷۰ \;; \quad \sigma_0^2 = ۹ \;, \quad s^2 = ۱۲ \;, \quad \nu = ۲۴$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{۲۴ \times ۱۲}{۹} = ۳۲$$

با توجه به جدول شماره ۵ مقدار بحرانی برابر $\chi^2_{۲۴(۰/۰۵)} = ۳۶/۴۲$ است.
بنابراین فرضیه صفر رد نمی‌شود. یعنی اختلاف بین واریانس‌ها معنی‌دار نیست.

آزمون ۱۶ : آزمون F برای واریانس دو جامعه (آزمون نسبت واریانسها)

موضوع : تعیین معنی‌دار بودن اختلاف بین واریانس‌های دو جامعه

شرایط :

دو جامعه باید دارای توزیع نرمال باشند (یکسان بودن میانگین دو جامعه ضروری نیست.)

روش :

نمونه‌هایی به حجم $n_۱$ با مقادیر $x_۱$ و $x_۲$ و ... $x_{n_۱}$ و به حجم $n_۲$ با مقادیر $y_۱$ و $y_۲$ و ... و $y_{n_۲}$ از دو جامعه در نظر گرفته و مقادیر

$$\bar{x} = \frac{\Sigma\, \bar{x_i}}{n_۱} \qquad \bar{y} = \frac{\Sigma\, \bar{y_i}}{n_۲}$$

$$s_۱^۲ = \frac{\Sigma\,(x_i - \bar{x})^۲}{n_۱ - ۱} \qquad s_۲^۲ = \frac{\Sigma\,(y_i - \bar{y})^۲}{n_۲ - ۱}$$

محاسبه می‌شوند.

تحت فرضیه صفر که واریانس دو جامعه یکسان است، آماره آزمون $F = \frac{s_۱^۲}{s_۲^۲}$ دارای توزیع F با درجه آزادی ($n_۲$-۱ و $n_۱$-۱) است. آزمون می‌تواند به هر یک ازدو صورت یک دامنه و یا دو دامنه انجام شود.

مثال عددی :

$$n_۱ = ۶ \quad , \quad \Sigma x = ۰/۴ \quad , \quad \Sigma x^۲ = ۰/۳ \quad , \quad s_۱^۲ = ۰/۰۸۷ \quad , \quad n_۲ = ۴$$

$$\Sigma y = ۰/۰۶ \quad , \quad \Sigma y^۲ = ۱/۷۸ \quad , \quad s_۲^۲ = ۰/۳۴ \quad , \quad F(۳و۵) = \frac{۰/۳۴}{۰/۰۸۷} = ۳/۹$$

با توجه به جدول شماره ۳ مقدار بحرانی برابر $F_{(۳و۵)(۰/۰۵)} = ۵/۴۱$ است.

بنابراین فرضیه صفر رد نمی‌شود . یعنی واریانس های دو جامعه با یکدیگر اختلاف

معنی‌داری ندارد.

توضیح:

در تشکیل آماره F توصیه می‌شود که همواره برآورد واریانس بزرگتر در صورت کسر قرار گیرد.

آزمون ۱۷ : آزمون F برای واریانس دو جامعه (با مشاهدات همبسته)

موضوع : تعیین اختلاف بین واریانس دو جامعه وقتیکه بین جفت‌های مشاهدات همبستگی وجود دارد.

شرایط :

چنین فرض می‌شود که مشاهدات بصورت جفت جفتی و بین مشاهدات جفت شده همبستگی وجود داشته و جامعه‌ها نیز به صورت نرمال توزیع شده باشند .

روش :

نمونه‌ای به حجم n از مشاهدات جفت شده (x_1, y_1) و (x_2, y_2) و ... و (x_n, y_n) در نظر گرفته و نسبت واریانس F رامشابه آزمون قبلی (آزمون ۱۶) و همچنین ضریب همبستگی نمونه r

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

را محاسبه می‌کنیم. آماره آزمون

$$F_r = \frac{F - 1}{\sqrt{\left[(F + 1)^2 - 4 r^2 F \right]}}$$

بعنوان ضریب همبستگی با درجه آزادی ۲ - n توزیع شده است.

مثال عددی :

$n_1 = n_2 = ۶$

$\Sigma x = ۰/۴$, $\Sigma x^2 = ۰/۳$, $s_1^2 = ۰/۰۸۷$

$\Sigma y = ۰/۰۶$, $\Sigma y^2 = ۱/۷۸$, $s_2^2 = ۰/۳۴$

$$r = ۰/۸۱۱ \quad , \quad F = \frac{s_۲^۲}{s_۱^۲} = ۳/۹$$

$$F_۲ = \frac{F - ۱}{\sqrt{\left[(F+۱)^۲ - ۴r^۲F\right]}} = \frac{۳/۹ - ۱}{\sqrt{(۴/۹)^۲ - ۴\times(۰/۸۱۱)^۲\times ۳/۹}} = ۰/۷۸۱$$

با توجه به جدول شماره ۶، $r = ۰/۸۱۱$ ، $\nu = n - ۲ = ۴$ ، $\alpha = ۰/۰۵$ است و چون F محاسبه شده ۰/۷۸۱ از $r = ۰/۸۱۱$ کوچکتر است. بنابراین فرضیه صفر را نمی‌توان رد کرد.

آزمون ۱۸ : آزمون T^{2} هتلینگ برای دو سری از میانگین‌های جامعه

موضوع : مقایسه نتایج دو آزمایش ،که هر کدام دارای نتایج مختلفی است. به بیان دیگر آیا میانگین بدست آمده از آزمایش اول با میانگین دومی سازگاری و توافق دارد.

شرایط :

متغیرها را می‌توان مستقل از یکدیگر فرض کرده که از توزیع نرمال چند متغیره پیروی می‌کنند.

روش :

نتایج دو آزمایش را تحت عنوان A و B نامیده و برای سهولت بیشتر، تعداد متغیرها را به سه مورد x و y و z محدود کرده و تعداد مشاهدات را برای دو آزمایش با n_B و n_A نشان میدهیم. برای بدست آوردن پارامترهای a و b و c باید سه معادله زیر را حل نمود .

$$a\,[(xx)_A + (xx)_B] + b\,[(xy)_A + (xy)_B] + c\,[(xz)_A + (xz)_B] = (n_A + n_B - ۲)\,(\bar{x}_A - \bar{x}_B)$$

$$a\,[(xy)_A + (xy)_B] + b\,[(yy)_A + (yy)_B] + c\,[(yz)_A + (yz)_B] = (n_A + n_B - ۲)\,(\bar{y}_A - \bar{y}_B)$$

$$a\,[(xz)_A + (xz)_B] + b\,[(yz)_A + (yz)_B] + c\,[(zz)_A + (zz)_B] = (n_A + n_B - ۲)\,(\bar{z}_A - \bar{z}_B)$$

که در آن:

$$(xx)_A = \Sigma(x_A - \bar{x}_A)^{۲}$$

$$(xy)_A = \Sigma(x_A - \bar{x}_A)\,(y_A - \bar{y}_A)$$

است و تعاریف مشابهی برای سایر جملات وجود دارد.

مقدار $T^۲$ هتلینگ بصورت زیر تعریف می‌شود.

$$T^۲ = \frac{n_A \times n_B}{n_A + n_B} \{a\ (\bar{x}_A - \bar{x}_B) + b\ (\bar{y}_A - \bar{y}_B) + c\ (\bar{z}_A - \bar{z}_B)\}$$

آماره آزمون نیز برابر است با :

$$F = \frac{n_A + n_B - p - ۱}{p\,(n_A + n_B - ۲)} \times T^۲$$

که دارای توزیع F با درجه آزادی $[(n_A+n_B-p-۱)$ و $P]$ است، که در آن P تعداد متغیرها است .

مثال عددی :

$n_A = ۶$, $n_B = ۴$, $DF = \nu = (۶ + ۴ - ۴) = ۶$

$\alpha = ۰/۰۵$, $(xx)_A = ۱۹$, $(yy) = ۳۰$

$(zz) = ۱۸$, $(xy) = -۶$, $\nu = \rho = ۳$

$(xz) = ۱$, $(yz) = -۷$, $\bar{x}_A = ۷$

$\bar{x}_B = ۴/۵$, $\bar{y}_A = ۸$, $\bar{y}_B = ۶$

$\bar{z}_A = ۶$, $\bar{z}_B = ۵$

با حل معادلات معمولی زیر:

$$\begin{cases} ۱۹a - ۶b + c = ۲۰ \\ -۶a + ۳۰b - ۷c = ۱۶ \\ a - ۷b + ۱۸c = ۸ \end{cases}$$

برآورد پارامترهای

$a = ۱/۳۲$

$b = ۰/۹۷۲$

$c = ۰/۷۴۹$

بدست می‌آید.

$$T^{۲} = \frac{۶ \times ۴}{۱۰} (۱/۳۲ \times ۲/۵ + ۰/۹۷۲ \times ۲ + ۰/۷۴۹ \times ۱) = ۱۴/۳۸$$

$$F = \frac{۶}{۳ \times ۸} \times ۱۴/۳۸ = ۳/۶$$

با توجه به جدول شماره ۳ مقدار بحرانی برابر $F_{۳،۶(۰/۰۵)} = ۴/۷۶$ است. لذا فرضیه صفر رد نمی‌شود.

آزمون ۱۹: آزمون تشخیصی برای مرجع یک نمونه P تایی

موضوع: تعیین متعلق بودن مجموعه داده‌ها با P متغیر تصادفی به یک جامعه، وقتیکه مجموعه فوق یقیناً از یکی از دو جامعه گرفته شده باشد.

روش:
با استفاده از نمادهای آزمون T^2 هتلینگ، نمونه‌ای از دو جامعه انتخاب و مقادیر D_A و D_B را برای آن دو جامعه بدست می‌آوریم:

$$D_A = a\bar{x}_A + b\bar{y}_A + c\bar{z}_A$$

$$D_B = a\bar{x}_B + b\bar{y}_B + c\bar{z}_B$$

برای مجموعه داده‌هایی که می‌خواهیم جامعه مرجع آنرا مشخص کنیم ، ملاک D_s باید محاسبه شود.

$$D_s = a\bar{x}_s + b\bar{y}_s + c\bar{z}_s$$

اگر $|D_B - D_S| > |D_A - D_S|$ باشد می‌توان گفت مجموعه داده‌ها متعلق به جامعه A است، در غیراینصورت نتیجه می‌گیریم که مجموعه مورد نظر از جامعه B گرفته شده است .

مثال عددی:

$a = ۱/۳۲$, $b = ۰/۹۷۲$, $c = ۰/۷۴۹$

$\bar{x}_A = ۷$, $\bar{y}_A = ۸$, $\bar{z}_A = ۶$, $\bar{x}_B = ۴/۵$, $\bar{y}_B = ۶$, $\bar{z}_B = ۵$

$D_A = ۱/۳۲ \times ۷ + ۰/۹۷ \times ۸ + ۰/۷۴۹ \times ۶ = ۲۱/۴۹۴$

$D_B = ۱/۳۲ \times ۴/۵ + ۰/۹۷ \times ۶ + ۰/۷۴۹ \times ۵ = ۱۵/۵۰۵$

اگر $\bar{z}_S = ۷$ و $\bar{y}_S = ۶$ و $\bar{x}_S = ۶$ باشد، سپس:

$$D_S = ۱/۳۲ \times ۶ + ۰/۹۷ \times ۶ + ۰/۷۴۹ \times ۷ = ۱۸/۹۸۳$$

$$D_A - D_S = ۲۱/۴۹۴ - ۱۸/۹۸۳ = ۲/۵۱۱$$

$$D_B - D_S = ۱۵/۵۰۵ - ۱۸/۹۸۳ = -۳/۴۷۸$$

مقدار D_S نزدیکتر به D_A است بنابراین D_S متعلق به جامعه A است.

آزمون ۲۰: آزمون کومولان فیشر برای نرمال بودن یک توزیع

موضوع: تعیین معنی‌دار بودن اختلاف بین توزیع فراوانی نمونه با توزیع نرمال

شرایط:

حجم نمونه باید بزرگ مثلاً $n > ۵۰$ باشد.

روش:

گشتاورهای نمونه برای یک توزیع فراوانی را می‌توان با:

$$M_r = \sum_{i=۱} x_i^n f_i \qquad \text{یا} \qquad M_r = \sum_{i=۱} x_i^r$$

بدست آورد و نخستین چهار کومولان نمونه (آمار K فیشر) عبارتند از:

$$K_۱ = \frac{M_۱}{n}$$

$$K_۲ = \frac{nM_۲ - M_۱^۲}{n(n-۱)}$$

$$K_۳ = \frac{n^۲M_۳ - ۳M_۲M_۱ + ۲M_۱^۳}{n(n-۱)(n-۲)}$$

$$K_۴ = \frac{(n^۳ - n^۲)M_۴ - ۴(n^۲+n)M_۳M_۱ - ۳(n^۲-n)M_۲^۲ + ۱۲M_۲M_۱^۲ - ۶M_۱^۴}{n(n-۱)(n-۲)(n-۳)}$$

بمنظور آزمودن چولگی، آماره آزمون عبارتست از:

$$U_۱ = \frac{k_۳}{\sqrt{(k_۲)^۳}} \times \sqrt{\frac{n}{۶}}$$

که از توزیع نرمال استاندارد پیروی می‌کند.

بمنظور آزمودن کشیدگی، آماره آزمون عبارتست از:

$$U_{۲}=\frac{k_{۴}}{(k_{۲})^{۲}}\times\sqrt{\frac{n}{۲۴}}$$

که از توزیع نرمال استاندارد پیروی می‌کند.

آزمون ترکیبی دیگری را می‌توان با استفاده از آماره آزمون زیر بدست آورد:

$$\chi^{۲}=\left[\frac{k_{۳}}{\sqrt{k_{۲}^{۳}}}\times\sqrt{\frac{n}{۶}}\right]^{۲}+\left[\frac{k_{۴}}{(k_{۲})^{۲}}\times\sqrt{\frac{n}{۲۴}}\right]^{۲}$$

که تقریباً از توزیع $\chi^{۲}$ با ۲ درجه آزادی پیروی می‌کند.

مثال عددی :

مثال :(الف)

$$\sum f = n = ۱۹۰ \quad , \quad \sum fx = ۱۵۱ \quad , \quad \sum fx^{۲} = ۸۰۵$$

$$\sum fx^{۳} = ۱۸۳۷ \quad , \quad \sum fx^{۴} = ۱۰۷۵۳$$

یعنی

$$M_{۱} = ۱۵۱ \ , \quad M_{۲} = ۸۰۵ \quad , \quad M_{۳} = ۱۸۳۷ \quad , \quad M_{۴} = ۱۰۷۵۳$$

$$K_{۲} = \frac{(۱۹۰ \times ۸۰۵) - (۱۵۱)^{۲}}{۱۹۰ \times ۱۸۹} = ۳/۶۲۴۳۱$$

$$K_{۳} = \frac{(۱۹۰)^{۲} \times ۱۸۳۷ - ۳ \times ۱۹۰ \times ۸۰۵ \times ۱۵۱ + ۲(۱۵۱)^{۳}}{۱۹۰ \times ۱۸۹ \times ۱۸۸} = ۰/۵۷۹$$

$$K_{۴} = \frac{۲۷۹۵۴۲۱۹۲۴}{۱۹۰ \times ۱۸۹ \times ۱۸۸ \times ۱۸۷} = ۲/۲۱۴۲۷۹$$

آزمون چولگی

$$U_{۱}=\frac{۰/۵۷۹}{۳/۶۲۴\times\sqrt{۳/۶۲}}\times ۵/۶۲۷=۰/۴۷۳$$

با توجه به جدول شماره ۱ مقدار بحرانی در سطح $\alpha = ۰/۰۵$ برابر با ۱/۹۶ است.
بنابراین فرضیه صفر رد نمی‌شود.

آزمون کشیدگی

$$U_۲ = \frac{۲/۲۱۴}{(۳/۶۲۴)^۲} \times \sqrt{\frac{۱۹۰}{۲۴}} = ۰/۴۷۴$$

با توجه به جدول شماره ۱ مقدار بحرانی در سطح $\alpha = ۰/۰۵$ برابر با ۱/۹۶ است.
بنابراین فرضیه صفر رد نمی‌شود.

آزمون ترکیبی

$$\chi^۲ = (۰/۴۷۳)^۲ + (۰/۴۷۴)^۲ = ۰/۴۴۸$$

با توجه به جدول شماره ۵ از مقدار بحرانی ۵/۹۹ کوچکتر است.
مثال (ب): فرض کنید

$$\text{چولگی} = g_۱ = \frac{k_۳}{k_۲\sqrt{k_۲}} = \frac{۰/۵۷۹}{۳/۶۲ \times \sqrt{۳/۶۲}} = ۰/۰۸۴۰۵$$

$$\text{کشیدگی} = g_۲ = \frac{k_۴}{k_۲^۲} = \frac{۲/۲۱۴}{(۳/۶۲)^۲} = ۰/۱۶۸۶$$

$$\sigma(g_۱) = \sqrt{\frac{۶n(n-۱)}{(n-۲)(n+۱)(n+۳)}} = \sqrt{\frac{۶ \times ۱۹۰ \times ۱۸۹}{۱۸۸ \times ۱۹۱ \times ۱۹۳}} = ۰/۱۷۶۳$$

$$\sigma(g_۲)\sqrt{\frac{۲۴n(n-۱)^۲}{(n-۳)(n-۲)(n+۳)(n+۵)}} = \sqrt{\frac{۲۴ \times ۱۹۰ \times ۱۸۹^۲}{۱۸۷ \times ۱۸۸ \times ۱۹۳ \times ۱۹۵}} = ۰/۳۵$$

و در نتیجه :
$$u_۲ = \frac{۰/۱۶۸۶}{۰/۳۵۰۹} = ۰/۴۸۰ \quad و \quad u_۱ = \frac{۰/۰۸۴}{۰/۰۱۷۶} = ۰/۴۷۷$$

با توجه به جدول شماره ۷ مقدار بحرانی برای g_1 بین ۰/۲۸۲ (برای ۲۰۰) و ۰/۳۰۱ (برای ۱۷۵) و مقدار بحرانی برای g_2 بین ۰/۶۲ و ۰/۶۶ قرار می‌گیرد. بنابراین فرضیه صفر را نباید رد کنید.

آزمون ۲۱: آزمون دیکسون برای مشاهده دوره افتاده

موضوع : تعیین معنی‌دار بودن اختلاف بین یک مقدار بیش از حد مشکوک و سایر مقادیر نمونه.

شرایط:

۱- حجم نمونه باید بزرگتر از ۳ باشد.

۲- جامعه‌ای که نمونه از آن گرفته شده نرمال فرض شده است.

روش :

نمونه‌ای به حجم n در نظر گرفته و وقتیکه نمونه شامل یک مقدار مشکوک باشد، از نزدیک‌ترین عدد مجاور آن، داده‌ها را به ترتیب صعودی (یا نزولی) مرتب می‌کنیم. بطوریکه مقدار مشکوک کوچکترین یا بزرگترین عدد باشد. مجموعه مرتب شده را با نماد $(x_۱ , x_۲ , ... , x_n)$ نشان می‌دهیم. آماره آزمون عبارتست از:

$$r = (x_۲ - x_۱) / (x_n - x_۱) \qquad \text{اگر} \quad ۳ < n \leq ۷$$

$$r = (x_۲ - x_۱) / (x_{n-۱} - x_۱) \qquad \text{اگر} \quad ۸ \leq n \leq ۱۰$$

$$r = (x_۳ - x_۱) / (x_{n-۱} - x_۱) \qquad \text{اگر} \quad ۱۱ \leq n \leq ۱۳$$

$$r = (x_۳ - x_۱) / (x_{n-۲} - x_۱) \qquad \text{اگر} \quad ۱۴ \leq n \leq ۲۵$$

مقادیر بحرانی برای r را می‌توان از جدول ۸ بدست آورد. اگر مقدار مشاهده شده r بیشتر از مقدار بحرانی باشد فرضیه صفر، یعنی تعلق مقدار دور افتاده به نمونه رد می‌شود.

مثال عددی:

$$x_۱ = ۳۲۶\,,\ x_۲ = ۱۷۷\,,\ x_۳ = ۱۷۶\,,\ x_۴ = ۱۵۷\,,\ n = ۴$$

در اینجا

$$r = \frac{x_۲ - x_۱}{x_n - x_۱} = \frac{۱۷۷ - ۳۲۶}{۱۵۷ - ۳۲۶} = ۰/۸۸۲$$

با توجه به جدول شماره ۸ مقدار بحرانی در سطح $\alpha =$ ۰/۰۵ برابر با ۰/۷۶۵ است. چون مقدار محاسبه شده بیشتر از مقدار بحرانی است. لذا فرضیه صفر، که مقدار $x_۱$ می‌تواند از جامعه فوق باشد، رد می‌شود.

آزمون ۲۲: آزمون F برای میانگین k جامعه (آنالیز واریانس)

موضوع: آزمودن فرضیه صفری که k نمونه از k جامعه با میانگین‌های مساوی گرفته شده‌اند.

شرایط:

چنین فرض می‌شود که جامعه‌ها دارای توزیع نرمال یا واریانس‌های مساوی و نمونه‌ها مستقل از یکدیگر باشند.

روش:

فرض کنید نمونه jام شامل n_j عضو ($j = ۱،۲، ...، k$) باشد. بنابراین تعداد کل اعضا نمونه برابر است با: $N=\sum_{j=۱}^{k} n_j$. که عضو iام از نمونه jام را با نماد x_{ij} ($i = ۱$ و ۲ و ... و n_j) نشان داده و میانگین نمونه jام عبارتست از:

$$N=\sum_{j=۱}^{k}\frac{x_{ij}}{n_{j.}}$$

واریانس مشاهدات با توجه به میانگین نمونه‌ای خودشان عبارتست از:

$$S_۱=\frac{\left(\sum\sum x_{ij}-x_{.j}\right)^۲}{N-K}$$

که درجه آزادی $(N - K)$ را خواهد داشت. بطریق مشابه واریانس میانگین نمونه‌ها با توجه به میانگین کل برابر است با :

$$S_۲=\frac{\sum n_j(x_{.j}-x_{..})^۲}{k-۱}$$

که در آن:

$$x_{..}=\frac{\sum\sum\left[\frac{x_{ij}}{n_j}\right]}{k}$$

و $S_۲$ درجه آزادی $(k - ۱)$ را خواهد داشت.

آماره آزمون $F = \frac{S_۲}{S_۱}$ است که از توزیع F با درجات آزادی $[(K-۱)$ و $(N-K)]$ پیروی میکند چنانچه تشخیص داده شود که $S_۲$ (واریانس بین نمونه‌ها) از $S_۱$ (واریانس داخل نمونه‌ها) بزرگتر است، می‌توان آزمون یک دامنه را بکار برد.

مثال عددی:

$$k = ۳, \quad N = ۱۲ \ , \quad n_۱ = ۳ \ , \quad n_۲ = ۵ \ , \quad n_۳ = ۴ \ , \quad \alpha = ۰/۰۵$$

$$\Sigma x_{i۱} = ۵۳/۵ \ , \quad \Sigma x_{i۲} = ۱۰۲/۵ \ , \quad \Sigma x_{i۳} = ۶۴/۴$$

$$\bar{x}_{.۱} = ۱۷/۸ \ , \quad \bar{x}_{.۲} = ۲۰/۵۰ \ , \quad \bar{x}_{.۳} = ۱۶/۱۰$$

$$x_{..} = \frac{۵۳/۵ + ۱۰۲/۵ + ۶۴/۴}{۳ + ۵ + ۴} = ۱۸/۳۷$$

$$S_۲ = \frac{۱}{۲}\left[(۳ \times ۰/۲۹۱۶) + (۵ \times ۴/۵۳۶۹) + (۴ \times ۵/۱۵۲۹)\right] = \frac{۴۴/۱۷}{۲} = ۲۲/۱$$

$$S_۱ = \frac{۱}{۹}\left[۴۹/۵۳ - ۴۴/۱۷\right] = \frac{۵/۴}{۹} = ۰/۶ \quad , \quad S = ۴۹/۵۳$$

$$F_{۲ \text{ و } ۹} = \frac{۲۲/۱}{۰/۶} = ۳۶/۸$$

با توجه به جدول شماره ۳ مقدار بحرانی برابر $F_{۲ \text{ و } ۹\ (۰/۰۵)} = ۴/۲۶$ است.
چون مقدار محاسبه شده بزرگتر از مقدار بحرانی است. بنابراین واریانس بین نمونه‌ها بطور معنی‌داری بزرگتر از واریانس داخل نمونه‌هاست.

آزمون ۲۳: آزمون Z برای نسبتهای وابسته

موضوع : تعیین معنی‌دار بودن اختلاف بین دو نسبت وابسته در سنجش افکار و نظرخواهی‌ها.

شرایط:

۱ـ افرادی یکسان در دو زمان، براساس بلی ـ خیر مورد سوال واقع شوند.

۲ـ حجم نمونه باید کاملاً بزرگ باشد.

روش:

گروهی از مردم سوالی را بصورت بلی ـ خیر، قبل و بعد از یک برنامه اجتماعی مشخص، پاسخ می‌گویند. آنگاه جدول دو طرف زیر را تشکیل می‌دهیم.

آراء دفعه اول

		بلی	خیر	
آراء دفعه دوم	بلی	a	b	
	خیر	c	d	
				N

بمنظور پاسخگوئی به این سوال که آیا برنامه اجتماعی فوق تغییر معنی‌داری در نسبت پاسخهای «بلی» بوجود آورده است؟

آماره آزمون زیر را محاسبه می‌کنیم.

$$Z = \frac{b - c}{N \, . \, \sigma}$$

که در آن

$$\sigma = \sqrt{\frac{(b+c) - (b-c)^{۲}/N}{N(N-۱)}} \quad (۱)$$

مثال عددی:

$$a = ۳۰ \;,\quad b = ۱۵ \;,\quad c = ۹ \;,\quad d = ۵۱ \;,\quad N = ۱۰۵$$

$$a + c = ۳۹ \;,\quad b + d = ۶۶ \;,\quad a + b = ۴۵ \;,\quad c + d = ۶۰$$

فرضیه صفر عبارتست از اینکه برنامه اجتماعی فوق تغییر چشمگیری را به همراه نداشته است.

۱- به عقیده مترجمین، در یک جدول توافقی که $\overline{d} = \frac{(b-c)}{n}$ تعریف می‌شود. واریانس ارائه شده بوسیله مولف در اصل کتاب بمنظور محاسبه و استفاده برای حدود اعتماد صحیح می‌باشد. امّا برای آزمودن فرضیه، تحت شرایط فرضیه صفر که در آن $H_۰ : \mu = \overline{d} = 0$ است، دقیق و صحیح نمی‌باشد. و باید در آن $\overline{d} = \frac{b-c}{n} = 0$ فرض شود. لذا؛

$$s_d^{۲} = \frac{\Sigma\,(di - \overline{d})^{۲}}{n} = \frac{\Sigma di^{۲}}{n} = \frac{b+c}{n}$$

$$S_{\overline{d}}^{۲} = \frac{s_d^{۲}}{n} = \frac{\frac{b+c}{n}}{n} = \frac{b+c}{n^{۲}}$$

$$Z = \frac{d}{\sqrt{s_{\overline{d}}^{۲}}} = \frac{(b-c)/n}{\sqrt{\frac{b+c}{n^{۲}}}} = \frac{b-c}{\sqrt{b+c}}$$

که این همان واریانس در آزمون مک‌نمار بوده و دقیق‌تر از واریانس فوق است.

اختلاف در نسبت (بلی) برابر است با :

$$\frac{(b-c)}{N} = \frac{۴۵}{۱۰۵} - \frac{۳۹}{۱۰۵} = \frac{۶}{۱۰۵} = ۰/۰۵۷۱$$

$$\sigma = \sqrt{\frac{(۹+۱۵)-(۹-۱۵)^{۲}/۱۰۵}{۱۰۵\times ۱۰۴}} = ۰/۰۴۶۶$$

$$Z = \frac{۰/۰۵۷۱}{۰/۰۴۶۶} = ۱/۲۳$$

با توجه به جدول شماره ۱ مقدار بحرانی در سطح $\alpha =$ ۰/۰۵ برابر با ۱/۹۶ است. چون مقدار محاسبه شده کمتر از مقدار بحرانی است. بنابراین فرضیه صفر رد نمی‌شود.

آزمون ۲۴: آزمون χ^2 برای واریانس یک جامعه مفروض

موضوع: تعیین معنی‌دار بودن اختلاف بین واریانس یک جامعه σ^2 و یک مقدار مفروض σ_0^2

شرایط:

چنین فرض می‌شود که نمونه از یک جامعه نرمال گرفته شده است.

روش:

واریانس نمونه با

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

محاسبه می‌شود. آماره آزمون عبارتست از:

$$\frac{s^2}{\sigma_0^2} \times (n-1)$$

که از توزیع χ^2 با درجه آزادی $(n - 1)$ پیروی می‌کند.

مثال عددی:

$$n = 25 \quad , \quad \bar{x} = 71 \quad , \quad s^2 = 12 \quad , \quad \sigma_0^2 = 9$$

$$H_0 : \ \sigma^2 = \sigma_0^2 \quad , \quad H_1 : \ \sigma^2 \neq \sigma_0^2$$

$$\chi^2 = \frac{s^2}{\sigma_0^2} \times (n-1) = \frac{12}{9} \times 24 = 32$$

با توجه به جدول شماره ۵ مقدار بحرانی برابر $\chi^2_{24(0/05)} = 36/4$ است. بنابراین فرضیه صفر رد نمی‌شود یعنی اختلاف بین واریانس و مقدار مفروض معنی‌دار نیست.

آزمون ۲۵: آزمون F برای دو مقدار شمارشی (توزیع پواسن)

موضوع: تعیین معنی‌دار بودن اختلاف بین دو پیامد شمارشی (مبتنی بر توزیع پواسن).

شرایط:
چنین فرض می‌شود که مقادیر شمارشی دارای توزیع پواسن بوده و دو نمونه تحت شرایط مشابه بدست آمده باشند.

روش:
فرض کنید که μ_1 و μ_2 میانگین‌ها و N_1 و N_2 مقادیر شمارشی دو جامعه را نشان می‌دهند. بمنظور آزمون فرضیه $H_0 : \mu_1 = \mu_2$
آماره آزمون

$$F = \frac{N_1}{N_2 + 1}$$

را محاسبه می‌کنیم که از توزیع F با درجات آزادی $[2N_1$ و $2(N_2 + 1)]$ پیروی می‌کند.
از آنجا که مقادیر شمارشی در فاصله‌های زمانی مختلف t_1 و t_2 بدست آمده‌اند، لازم است که نسبت‌های مقادیر شمارشی $\frac{N_1}{t_1}$ و $\frac{N_2}{t_2}$ با هم مقایسه شوند. لذا آماره آزمون مناسب عبارت است از:

$$F = \frac{\frac{1}{t_1}(N_1 + 0/5)}{\frac{1}{t_2}(N_2 + 0/5)}$$

که از توزیع F با درجات آزادی $[(2N_1 + 1), (2N_2 + 1)]$ پیروی میکند.

مثال عددی:

$N_1 = 13 \quad , \quad N_2 = 3 \quad , \quad t_1 = t_2$

$$f_1 = ۲(N_2 + ۱) = ۲ \times (۳ + ۱) = ۸ \quad , \quad f_2 = ۲N_1 = ۲ \times ۱۳ = ۲۶$$

$$F = \frac{N_1}{N_2 + ۱} = \frac{۱۳}{۳ + ۱} = ۳/۲۵$$

با توجه به جدول۳ مقدار بحرانی برابر $F_{۸,۲۶}(۰/۰۵) = ۲/۳۲$ است. چون مقدار محاسبه شده بیشتر از مقدار جدول است. بنابراین فرضیه صفر رد می‌شود.

توضیح:

بدلیل مساوی بودن زمانهای t_1 و t_2 در مثال عددی از $F = \frac{N_1}{N_2 + ۱}$ استفاده شده است.