تحلیل آماری
Analysis Statistical

کلاس استاد

سرکار خانم دکتر نازنین پیله‌وری سلماسی

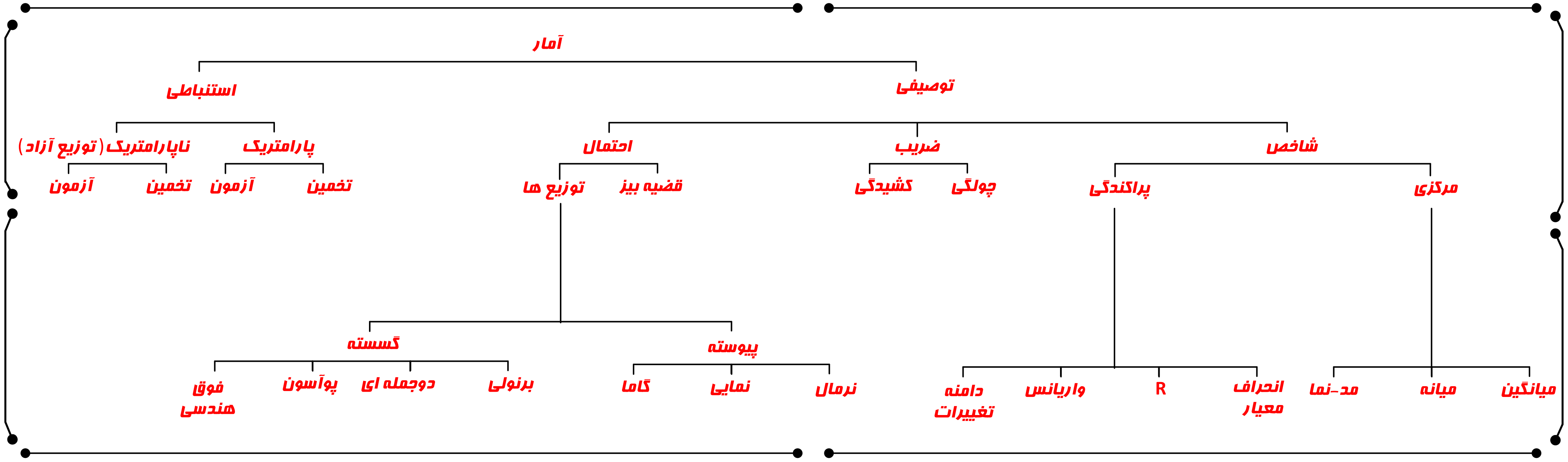
مبتنی بر تدریس استاد ارجمند جناب آقای دکتر پیدایی

تدوین: بابک آشفته یزدی

فروردین ۹۵

www.freebay.ir

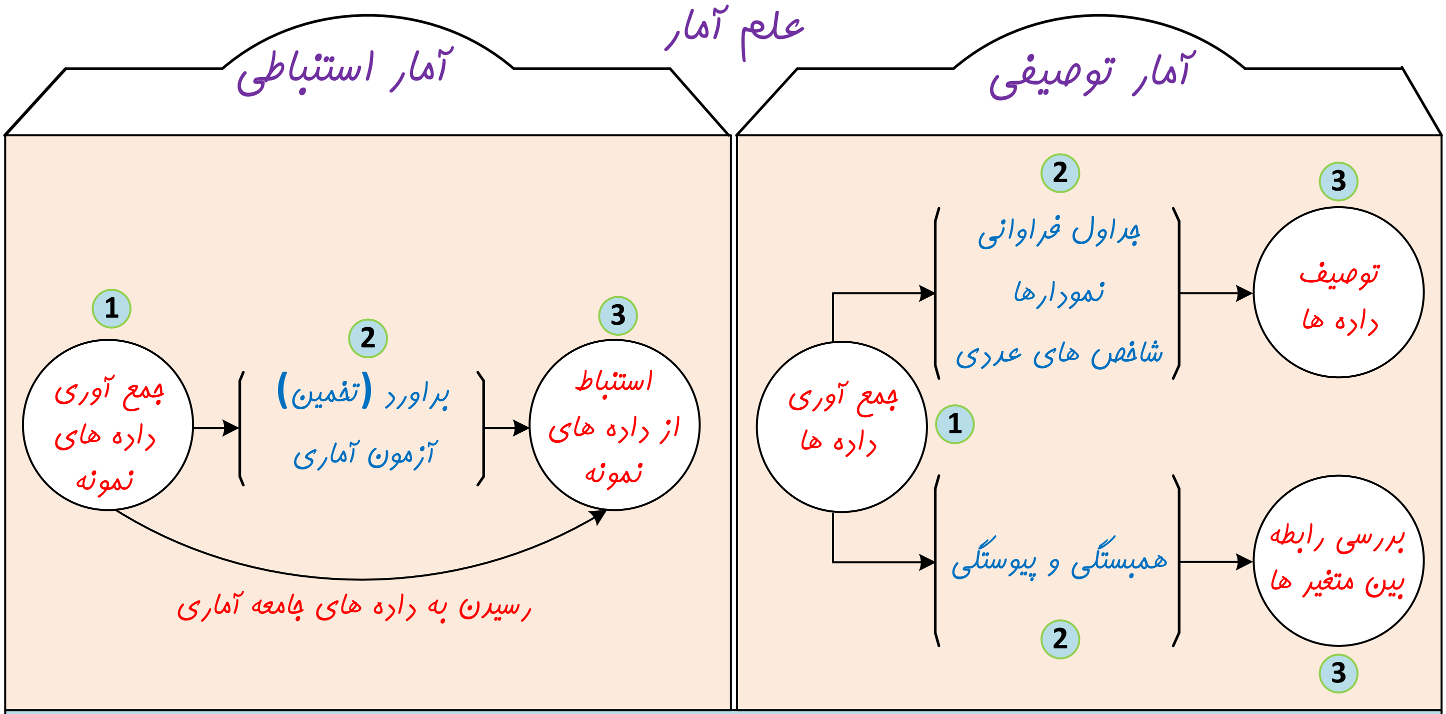
تحلیل آماری
Statistical Analysis



www.freebay.ir

www.freebay.ir

علم آمار



آمار توصیفی Statistics Descriptive

– در آمار توصیفی با داشتن تمام اعضا جامعه به بررسی خصوصیت‌های آماری آن پرداخته می‌شود

– موضوع آمار توصیفی statistics Descriptive تنظیم و طبقه‌بندی داده‌ها، نمایش ترسیمی، و محاسبه مقادیری از قبیل نما، میانگین، میانه و ... می‌باشد که حاکی از مشخصات یکایک اعضای جامعه مورد بحث است. در آمار توصیفی اطلاعات حاصل از یک گروه، همان گروه را توصیف می‌کند و اطلاعات به دست آمده به دسته‌های مشابه تعمیم داده نمی‌شود. به طور کلی از سه روش در آمار توصیفی برای خلاصه‌سازی داده‌ها استفاده می‌شود:

۱– استفاده از جداول ۲– استفاده از نمودار ۳–محاسبه مقادیری خاص که نشان‌دهنده خصوصیات مهمی از داده‌ها باشند.

آمار استنباطی Statistics Inferential

– در آمار استنباطی با بدست آوردن نمونه‌ای از جامعه که خصوصیات اصلی جامعه را بیان می‌کند در مورد جامعه استباط آماری انجام می‌شود.

– چنانچه به جای مطالعه کل اعضای جامعه، بخشی از آن با استفاده از فنون نمونه‌گیری انتخاب شده، و مورد مطالعه قرار گیرد و بخواهیم نتایج حاصل از آن را به کل جامعه تعمیم دهیم از روش‌هایی استفاده می‌شود که موضوع آمار استنباطی statistics Inferential است. آن چه که مهم است این است که در گذر از آمار توصیفی به آمار استنباطی یا به عبارت دیگر از نمونه به جامعه بحث و نقش احتمال شروع می‌شود. در واقع احتمال، پل رابط بین آمار توصیفی و استنباطی به حساب می‌آید.

– در حالی که در نظریه آمار استنباطی اتفاقات تصادفی و عدم قطعیت توسط نظریه احتمالات مدل‌سازی می‌شوند. در این علم، مطالعه و قضاوت معقول در باره موضوع‌های گوناگون، بر مبنای یک نمونه انجام می‌شود و قضاوت در مورد یک فرد خاص، اصلاً مطرح نیست. علم آمار یکی از علوم مرتبط با علم داده‌ها است

لزوم استفاده از آمار در مدیریت

۱ – قابل استفاده کردن داده های جمع آوری شده با توصیف داده ها (بکارگیری عملیاتی نتایج تحقیق)

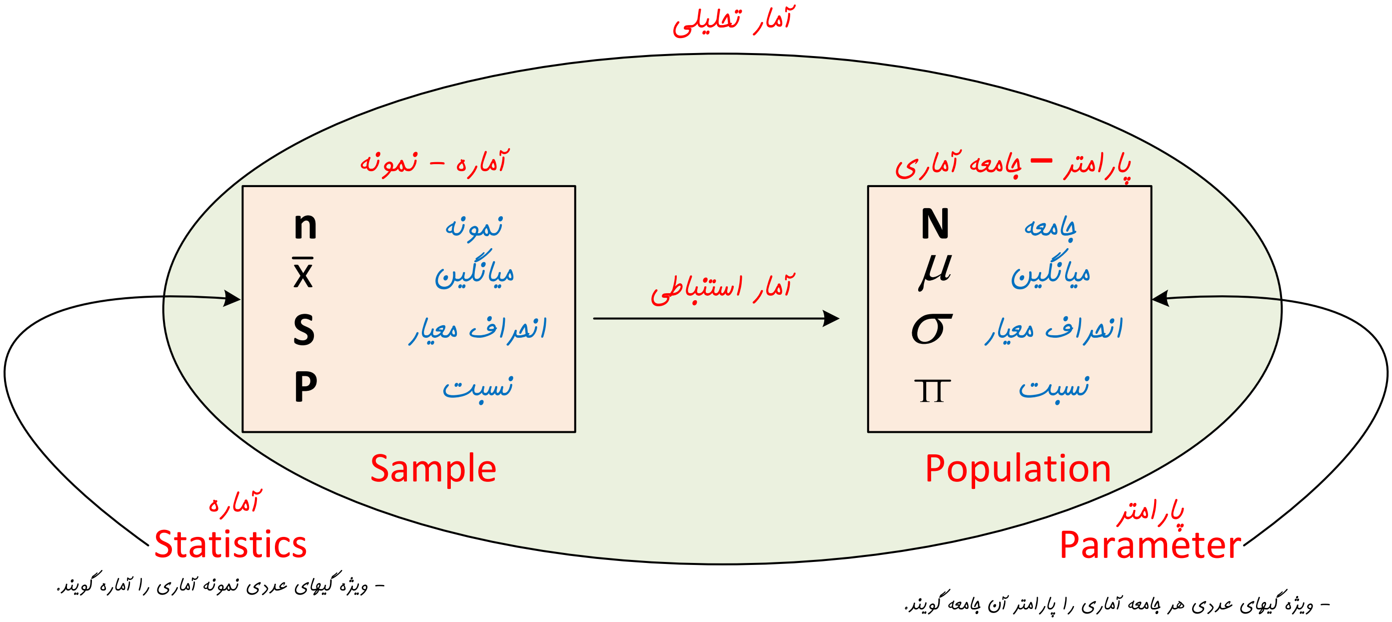
۲ – پاسخ به بررسی سئولات تحقیق

۳ – رد یا تایید فرضیه های پژوهش

مراحل اساسی انجام پژوهش علمی و نیاز به آمار در فصل ۴

	اهمیت مسئله و اهداف ⇐	بیان مسئله ⇐	مقدمه ⇐	فصل ۱ کلیات تحقیق
تدوین فرضیه ها ⇐	تدوین مدل مفهومی ⇐	تحقیقات پیشین ⇐	اصول و مبانی علمی ⇐	فصل ۲ بیان پیشینه
	روایی و اعتبار ⇐	آزمون های مورد استفاده ⇐	جامعه و نمونه آماری ⇐	فصل ۳ مترولوژی تحقیق
	تحلیل استنباطی داده ها ⇐	تحلیل توصیفی داده ها ⇐	تحلیل داده های جمعیت شناختی ⇐	فصل ۴ تحلیل داده ها
		پیشنهادات ⇐	نتیجه گیری ⇐	فصل ۵ جمع بندی

www.freebay.ir



روش های نمونه گیری

از آغاز پیدایش بشر تا کنون ، همیشه و همواره سوالات فراوانی در زندگی و امور جاری مطرح بوده است . پرسش های فراوانی که نائل شدن به پاسخ های آنها منوط به انجام آزمون و تحقیق بوده است . انجام اینگونه تحقیقات نیز مستلزم صرف انرژی ، هزینه ، زمان و ... می باشد که انسانها سعی در کاهش این پارامترها داشته اند . در بسیاری از موارد که رسیدن به پاسخ سوالات مطروحه مستلزم انجام تحقیق و پژوهش در جوامع آماری گسترده می باشد ، استفاده از تکنیک های نمونه گیری از لحاظ موارد مذکور ، نمود بیشتری پیدا کرده اند . اما اینکه نمونه و نمونه گیری چیست و چرا به دنبال نمونه گیری هستیم؟ تعاریف مختلف و گوناگونی جهت نمونه ، نمونه گیری و دیگر اصطلاحات رایج این مقوله عنوان شده است .

یکی از تعاریف موجود ، نمونه را بدین شرح توصیف نموده است که هر جزء جامعه را میتوان نمونه آن جامعه دانست یا در تعریفی دیگر ، نمونه کوچک ترین واحد مستقلی است که یک پاسخ فراهم می آورد. نمونه باید معرف جامعه آماری باشد یعنی صفات جامعه مخصوصاً صفاتی که دارای اهمیت است به تناسب در نمونه وجود داشته باشد . نمونه گیری به روش های منظم انتخاب اطلاق می شود و در تحقیقات اجتماعی برای انتخاب افراد یا موردهای تحقیق بکار می رود .

غیر احتمالی Non-Probability Sampling ②	احتمالی Probability Sampling ①
① نمونه گیری سهمیه ای	① نمونه گیری تصادفی ساده
② نمونه گیری گوله برفی	② سیستماتیک یا منظم
③ نمونه گیری قصدی	③ خوشه ای
④ نمونه‌های در دسترس	④ طبقه ای

احتمالی Probability Sampling

روش های نمونه گیری احتمالی

در نمونه گیری احتمالی هر یک از واحدهای تشکیل دهنده جمعیت ، برای وارد شدن در نمونه از یک احتمال معین ، برابر یا نابرابر ولی نامساوی با صفر برخوردار است

انتخاب نمونه احتمالی با کمک عامل شانس انجام می شود . این عامل شانس است که به جای قضاوت و دانش محقق ، معین می کند کدام واحد باید در نمونه وارد شود بنابراین اشتباهات در نمونه های احتمالی عمدتاً از مقولهٔ اشتباهات تصادفی Error Random است البته اشتباهات تصادفی قابل اندازه گیری می باشند . بر همین اساس با استفاده از آزمونهای آماری معنی دار بودن ، محقق می تواند با اطمینانی معین نسبت به تعمیم پذیری برآوردهای نمونه ای قضاوت کند .

درنمونه گیری احتمالی , هر یک از واحدهای تشکیل دهندهٔ جمعیت برای وارد شدن در نمونه ازیک احتمال معین، برابر یا نابرابر ولی نامساوی با صفر برخوردار است. انتخاب نمونهٔ احتمالی به مدد عامل شانس انجام میشود. این عامل شانس است که به جای قضاوت و دانش محقق، معین میکند کدام واحد باید در نمونه وارد شود. بنابراین، اشتباهات در نمونه های احتمالی عمدتاً از مقولهٔ اشتباهات تصادفی است. هدف از نمونه گیری احتمالی تعمیم، افزایش تعمیم پذیری با انتخاب نمونه های معرف است انتخاب تصادفی مهمترین حالت نمونه گیری احتمالی است.

نمونه گیری احتمالی میتواند یک مرحله ای و یا چند مرحله ای باشد. نمونه گیریهای تصادفی ساده و سیستماتیک (منظم) یک مرحله ای هستند و نمونه گیریهای طبقه بندی شده و خوشه ای چندمرحله ای هستند. در چند مرحله ای ممکن چند واحد هدف داشته باشیم که چند واحد تحلیل متفاوت دارند .

www.freebay.ir
2Bay

ادامه روش های نمونه گیری احتمالی

احتمالی Probability Sampling

۱ نمونه گیری تصادفی ساده

نمونه گیری تصادفی ساده SRS شیوه اصلی انتخاب در نمونه گیری های احتمالی است. تصادفی ساده که عمدتاً با کمک جدول اعداد تصادفی برگزیده میشود، شانس همهٔ واحدهای جمعیت برای ورود به نمونه مساوی است در واقع نمونه گیری تصادفی ساده یکی از مصادیق بارز و با اهمیت نمونه گیری با احتمال برابر است. یک نمونه گیری تصادفی ساده میتواند با جای گذاری یا بدون جای گذاری باشد. نمونه ای را جای گذاری میگویند که افراد باز شانس در انتخاب شدن دارند در نمونه گیری . بدون جای گذاری بطور مثال جامعه ای را در نظر بگیرید که دوتایی میتوان استخراج کنیم. اگر عضو معینی از جامعه یکبار بعنوان عضو اول انتخاب شده خود به خود برای انتخاب عضو دوم کنار گذارده میشود این روش را نمونه برداری بدون جای گذاری یا بدون جانشین میگویند. اگر هر عضو در جامعه بتواند بیش از یکبار در نمونه ظاهر شود را نمونه برداری با جایگذاری میگویند.

اگر حجم جامعه بزرگ باشد نمونه گیری جای گذاری معنای ندارد.

شرایط استفاده ۱ – دسترسی به یک فهرست کامل از واحدهای تشکیل دهندهٔ جمعیت است. ۲- حجم نمونه مشخص شده است ۳- انتخاب نمونه از جامعه بر اساس قرعه کشی یا جدول اعداد تصادفی

۲ سیستماتیک یا منظم

ابتدا حجم جمعیت را بر حجم نمونه تقسیم می کنیم تا فاصله نمونه گیری بدست آید. سپس یک عدد اتفاقی را مبنای شروع قرار می دهیم و به اندازه فاصله نمونه گیری، افراد یا عناصر بعدی را انتخاب می کنیم تا نمونه مورد نظر کامل شود.

۳ خوشه ای

برای مناطق وسیع جغرافیایی که بدست آوردن فهرست کاملی از اجزای آن غیر ممکن است بکار می رود. در این روش نقشه مورد مطالعه را به بخشهایی تقسیم می کنیم سپس به روش اتفاقی ساده یا منظم از آن نمونه گیری می کنیم.

دستور العمل برای انتخاب بهترین نمونه :

به حداکثر رساندن تعداد خوشه ها و کاهش تعداد عناصر هر خوشه.

نمونه گیری خوشه ای با احتمال متناسب با حجم :

در این روش متناسب با حجم هر خوشه از آن نمونه گیری می کنیم.

۴ طبقه ای

برای جامعه آماری که ساخت همگن و متجانس ندارد استفاده می شود. این روش خطای نمونه گیری را کاهش می دهد جامعه را به مورد مطالعه در این روش . زیرمجموعه های همگن تقسیم می کنیم. مثال : اعضای یک دانشگاه.

روش های نمونه گیری غیراحتمالی

غیر احتمالی Non-Probability Sampling

در نمونه گیری غیر احتمالی ،به جای تکیه بر عامل شانس، نمونه به مدد قضاوت انسانی انتخاب میشود. قضاوتی که خود تحت تأثیر معجونی از اصلاحات و علایق شکل میگیرد بنابراین شانس وارد شدن، هر یک از واحدهای جمعیت در نمونه، نامعین و نامعلوم است ،اشتباهات برآورد در نمونه های غیر احتمالی اغلب غیرتصادفی و غیر قابل اندازه گیری است.

۱ نمونه گیری سهمیه ای

باید ساختار جامعه مورد مطالعه مشخص باشد و نیاز به اطلاعات روزآمد دارد. باید همان نسبتهایی که در جمعیت مورد مطالعه وجود دارد در نمونه انتخابی رعایت شود.

مثال : افراد شاغل.

۲ نمونه گیری گوله برفی

نمونه گیری گلوله برفی یک روش نمونه گیری غیر احتمالی برای مواقعی است که واحدهای مورد مطالعه براحتی قابل شناسایی نباشند واحدها این که یی‌اهن‌گام ویژه. بسیار کمیاب یا بخش کوچکی از یک جامعه خیلی بزرگ را تشکیل می دهند. در این روش آمارگیر پس از شناسایی یا انتخاب اولین واحد نمونه گیری از آن برای شناسایی و انتخاب دومین واحد نمونه گیری استفاده یا کمک می گیرد. به همین ترتیب واحدهای دیگر نمونه شناسایی و انتخاب می شوند. این نمونه گیری می تواند به دو صورت خطی و غیر خطی (نمایی) اجرا گردد.

هنگامی از نمونه گیری گلوله ی برفی استفاده می شود که چارچوبی برای نمونه گیری وجود ندارد یا‌ه‌گروه ی‌ا‌رد‌ب ش‌ه‌وژپ ی‌ا‌رب ی‌ر‌ی‌گ ه‌ن‌و‌م‌ن ی هویش ن‌ی‌ا ن‌ی‌نچ‌م‌ه. از جامعه کاربرد دارد که به فعالیت هایی اشتغال دارند که بر حسب شرایط جامعه غیر قانونی باشد و بنابراین به دلیل شرایط خاص اجتماعی و امکان دسترسی مستقیم به آنها وجود ندارد . بر اساس این شیوه ی نمونه گیری، پژوهش گر از طریق ایجاد ارتباطات شخصی با کسانی که علاقه مند به انجام مصاحبه یا پر کردن پرسشنامه هستند، از آنها در خواست می کند اگر اشخاص دیگری را می شناسند که دارای ویژگی های مورد نظر پژوهش و مایل به انجام مصاحبه هستند، معرفی کنند. انتخاب جمعیت نمونه از این طریق و به صورت زنجیر وار ادامه می یابند تا زمانی ک دیگر نمونه ای پیدا نشود.

مثال : افراد بی خانمان، مهاجران غیر قانونی

۳ نمونه گیری قصدی

نمونه بر اساس قضاوت شخصی یا اهداف مطالعه انتخاب میشود.

۴ نمونه‌های در دسترس

www.freebay.ir

مقیاس های اندازه گیری

انواع مقیاس

- کمی Quantitative
  - نسبی Ratio
  - فاصله ای Interval
- کیفی Qualitative
  - ترتیبی Ordinal
  - اسمی Nominal

کیفی Qualitative

اسمی Nominal

- مقیاس هایی که فقط بواسطه اسمشان از هم جدا میشوند.

ترتیبی Ordinal

- مقیاس هایی هستند که هم توسط اسمشان از هم جدا میشوند و هم ترتیبی نیز در آن ها وجود دارد.

کمی Quantitative

فاصله ای Interval

- دارای خصوصیات اسمی - ترتیبی میباشد و با عدد نشان داده میشود و دارای صفر قراردادی میباشد.

نسبی Ratio

- دارای خصوصیات اسمی – ترتیبی – فاصله ای است و نیز دارارای صفر مطلق میباشد.

آمار پارامتریک

برای سنجش فرضیه‌هایی که متغیر آن کمی‌اند، از آمار پارامتریک استفاده می‌شود.

متغیرهای کمی به علت کمی بودن و واحد پذیر بودن از این ویژگی برخوردارند که آنها را میانگین‌پذیر و انحراف معیارپذیر می‌کنند و به دلیل همین ویژگی معمولا برای استفاده از آزمون های پارامتریک پیش فرض هایی لازم است که از جمله، نرمال بودن توزیع جامعه است زیرا در حالتی که توزیع جامعه نرمال نباشد، میانگین و انحراف معیار، نمایی واقعی از داده ها را به تصویر نمی‌کشانند.

آمار ناپارامتریک

برای آزمون متغیرهای کیفی و رتبه از آمار ناپارامتریک استفاده می‌شود. این آزمونها که از آنها با عنوان آزمونهای بدون پیش فرض نیز یاد می‌شود به هیچ پیش فرض خاصی نیاز ندارد.

می‌توان متغیرهای کمی را به متغیرهای کیفی تبدیل کرد و آنها را با آزمون‌های ناپارامتریک مورد ارزیابی قرار داد ولی عکس این عمل امکانپذیر نیست.

مقیاس ها	مثال	شاخص مرکزی	آزمون های آماری	نوع
اسمی	مرد - زن	مد	ناپارامتریک	کیفی
رتبه ای	طیف لیکرت	مد و میانه	ناپارامتریک	کیفی
فاصله ای	درجه سانتیگراد فارنهایت	مد - میانه - میانگین	پارامتریک	کمی
نسبی	متر	مد - میانه - میانگین	پارامتریک	کمی

4Bay
www.freebay.ir

ریاضیات مطالعه توصیفی داده های طبقه بندی نشده – کل جامعه آماری

برای مطالعات توصیفی با اندازه های کوچکی از مقادیر $N \le 20$ استفاده میشود و مرتبط با کل جامعه آماری میباشد.

اعدادی را که به منظور بیان کمی توزیع اندازه ها از آن استفاده میشود **شاخص های عددی** گویند.

جمع مشاهدات

$$\sum_{i=1}^{N} X_i = X_1 + X_2 + \cdots + X_N$$

فرمولهای مرتبط

$$\rightarrow \sum_{i=1}^{N} bX_i = b\sum_{i=1}^{N} X_i \rightarrow \sum_{i=1}^{N} a = N.a$$

میانگین حسابی

معدل مجموعه ای از مشاهدات را **میانگین حسابی** مینامند. که نقطه تعادل و مقدار مرکزی مشاهدات میباشد.

میانگین حسابی وزنی

$$\mu_W = \frac{\sum_{i=1}^{N} W_i X_i}{\sum_{i=1}^{N} W_i} = \frac{\sum_{i=1}^{N} W_i X_i}{N}$$

اگر هریک از مشاهدات به تعداد **W** بار تکرارشوند (دارای وزن گردند) در اینصورت **میانگین حسابی وزنی** گویند.

میانگین حسابی ساده

$$\mu_x = \frac{\sum_{i=1}^{N} X_i}{N}$$

اگر مشاهدات یکبار تکرار شده باشند **میانگین حسابی ساده** گفته میشود.

خاصیت های میانگین حسابی

$$y_i = bx_i \rightarrow \mu_i = b\mu_x$$

هر گاه هر یک از مشاهدات آماری با در عدد **b** ضرب شوند میانگین اعداد حاصل شده برابر میانگین اعداد قبلی ضرب در عدد **b** خواهد بود

$$y_i = x_i + a \rightarrow \mu_i = \mu_x + a$$

هر گاه هر یک از مشاهدات با عدد ثابت **a** جمع شود ، میانگین اعداد حاصل شده برابر میانگین مجموعه اعداد قبلی به اضافه **a** خواهد بود

$$\sum_{i=1}^{N} (X_i - \mu_x) = 0$$

جمع جبری اختلاف مجموعه ای از اعداد از میانگینشان برابر صفر است

$$y_i = x_i + z_i \rightarrow \mu_y = \mu_x + \mu_z$$

اگر **Z,X** دو مجموعه از مشاهدات باشند و مجموعه **Y** از جمع دو به دو اعداد **Y,X** حاصل شده باشد ، میانگین مجموعه **Y** برابر است با جمع میانگین مشاهدات **Z,X**

واریانس و انحراف معیار

یکی از شاخص های اندازه گیری پراکندگی داده ها نسبت به میانگین واریانس و انحراف معیار است.

واریانس

$$\sigma_x^2 = \frac{\sum (X_i - \mu_x)^2}{N} \rightarrow \sigma_x^2 = \frac{\sum X_i^2}{N} + \mu_x^2$$

انحراف معیار

$$\sigma_x = \sqrt{\sigma_x^2}$$

هر گاه هر یک از مشاهدات آماری در عدد **b** ضرب شوند واریانس **bx** معادل **b^2** برابر واریانس قبلی خواهد بود.

$$y_i = bx_i \rightarrow \sigma_y^2 = b^2 \sigma_x^2$$

هر گاه هر یک از مشاهدات با عدد ثابت **a** جمع شود ، واریانس مشاهدات جدید فرق نمیکند

$$y_i = x_i + a \rightarrow \sigma_y^2 = \sigma_x^2$$

www.freebay.ir

میانگین حسابی

در دانش آمار میانگین حسابی یا متوسط حسابی به انگلیسی: mean Arithmetic نوعی سنجش گرایش به مرکز است، و عبارت است از مجموع مقادیر موجود در یک مجموعه داده‌ها تقسیم بر تعداد آن‌ها.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + \cdots\cdots + x_n)$$

امید ریاضی

در نظریه احتمالات امید ریاضی، میانگین، مقدار مورد انتظار یا ارزش مورد انتظار یک متغیر تصادفی گسسته برابر است با مجموع حاصل‌ضرب احتمال وقوع هر یک از حالات ممکن در مقدار آن حالت. در نتیجه میانگین برابر است با مقداری که بطور متوسط از یک فرایند تصادفی با بی‌نهایت تکرار انتظار می‌رود. بطور مثال برای تاس داریم:

$$E[X] = \frac{1}{6}\times 1 + \frac{1}{6}\times 2 + \frac{1}{6}\times 3 + \frac{1}{6}\times 4 + \frac{1}{6}\times 5 + \frac{1}{6}\times 6 = 3.5$$

$$E[X] = \sum_{i=1}^{n} x_i P(x_i)$$

در صورتیکه همه احتمالات متغیرها یکسان باشد همانند : توزیع نرمال امید ریاضی برابر میانگین خواهد بود .

واریانس

در نظریه احتمالات و آمار وردایی یا واریانس نوعی سنجش پراکندگی است.

مقدار وردایی با میانگین‌گیری از مربع فاصله مقدار محتمل و یا مشاهده شده با مقدار مورد انتظار محاسبه می‌شود. در مقایسه با میانگین می‌توان گفت که میانگین مکان توزیع را نشان می‌دهد، در حالی که وردایی مقیاسی است که نشان می‌دهد که داده‌ها حول میانگین چگونه پخش شده‌اند. وردایی کمتر بدین معنا است که انتظار می‌رود که اگر نمونه‌ای از توزیع مزبور انتخاب شود مقدار آن به میانگین نزدیک باشد. یکای وردایی مربع یکای کمیت اولیه می‌باشد. ریشه دوم وردایی که انحراف معیار نامیده می‌شود دارای واحدی یکسان با متغیر اولیه است.

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

با توجه به این که تخمین فوق یک تخمین دقیق و بدون خطا برای وردایی نیست لذا برای از بین بردن این خطا در تخمین از وردایی تصحیح شده استفاده می‌کنیم که بصورت زیر تعریف می‌گردد

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

انحراف معیار

در آمار انحراف معیار به انگلیسی: **deviation standard** که با نماد **σ** نشان داده می‌شود یکی از شاخص‌های پراکندگی است که نشان می‌دهد به طور میانگین داده‌ها چه مقدار از مقدار متوسط فاصله دارند. اگر انحراف معیار مجموعه‌ای از داده‌ها نزدیک به صفر باشد، نشانه آن است که داده‌ها نزدیک به میانگین هستند و پراکندگی اندکی دارند؛ در حالی که انحراف معیار بزرگ بیانگر پراکندگی قابل توجه داده‌ها می‌باشد. انحراف معیار برابر با ریشه دوم واریانس است. خوبی آن نسبت به واریانس، این است که هم‌بعد با داده‌ها می‌باشد.

انحراف معیار برای تعیین ضریب اطمینان در تحلیل‌های آماری نیز به کار می‌رود. در مطالعات علمی، معمولاً داده‌های با انحراف معیار بیشتر از دو به عنوان داده‌های پرت در نظر گرفته و از تحلیل، خارج می‌شوند.

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

www.freebay.ir

آمار تحلیلی

آماره – فضای نمونه			پارامتر – جامعه آماری	
نمونه	n	آمار استنباطی →	جامعه	N
میانگین	$\bar{x}$		میانگین	μ
انحراف معیار	S		انحراف معیار	σ
نسبت	P		نسبت	π
Sample			Population	

آماره
Statistics

- ویژه گیهای عددی نمونه آماری را آماره گویند.

پارامتر
Parameter

- ویژه گیهای عددی هر جامعه آماری را پارامتر آن جامعه گویند.

میانگین حسابی ساده

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + \cdots\cdots + x_n)$$

$$\mu_x = \frac{\sum_{i=1}^{N} X_i}{N}$$

انحراف معیار

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

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \qquad S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

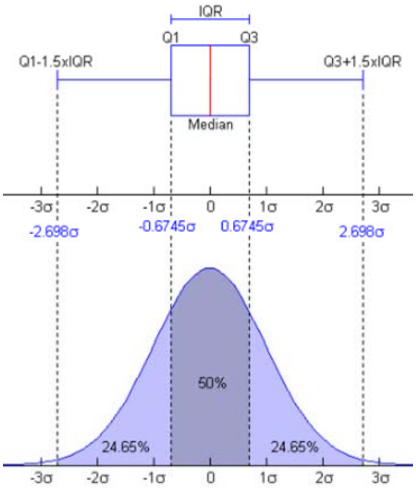
$$\sigma_x = \sqrt{\frac{\sum (X_i - \mu_x)^2}{N}}$$

$$\sigma_x = \sqrt{\frac{\sum X_i^2}{N} - \mu_x^2}$$

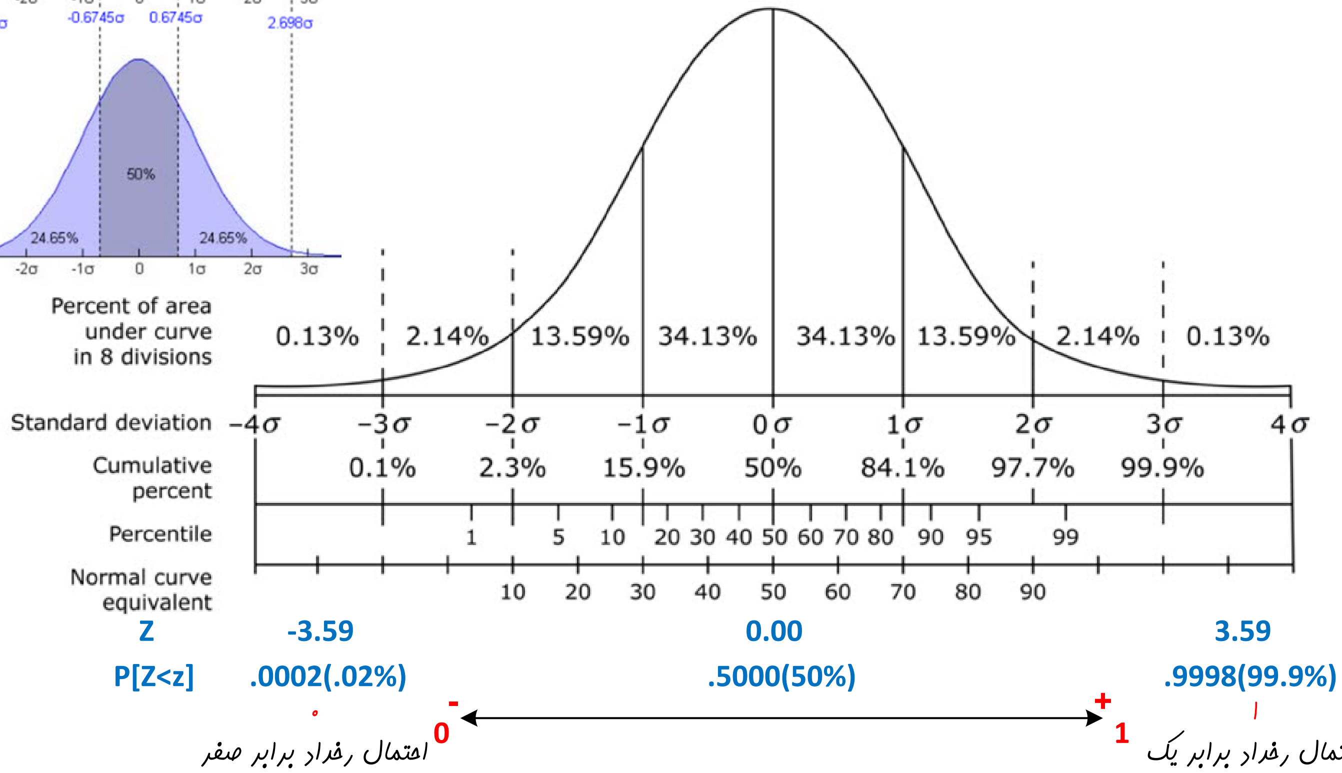
تناظر ریاضی فضای نمونه و تعمیم آن به جامعه آماری

www.freebay.ir

توزیع نرمال



موارد بسیاری وجود دارد که داده ها میل به جمع شدن در اطراف مقدار وسطی دارند بدون این که به چپ یا راست تمایل داشته باشند، و به "توزیع نرمال" نزدیک می شوند.



خمیدگی روی سطح زنگوله یک توزیع نرمال است. و هیستوگرام (بافت نگار) Histogram برخی از داده ها که به این منحنی نزدیک است اما دقیقا منطبق نیست را نشان می دهد.

برخی موارد توزیع نرمال قد مردم-اندازه اجسام شکل یافته توسط ماشین ها-خطاهای اندازه گیری-فشار خون-نمرات یک امتحان

خواص اصلی ریاضی توزیع نرمال میانگین - میانه - مد - تقارن در وسط - ۵۰٪ مقادیر کوچکتر از میانگین ۵۰٪ مقادیر بزرگتر از میانگین است.

قضیه حد مرکزی

قضیه حد مرکزی Central Limit Theorem در نظریه احتمالات بیان می‌کند که با فرض شرایطی خاص، میانگین تعدادی متغیر تصادفی مستقل، که هر یک میانگین و واریانس به خوبی تعریف شده دارند، بطور تقریبی دارای توزیع نرمال خواهد بود.

نکته:

فرض شده است که ما متغیر تصادفی داریم که همگی دارای توزیع احتمال یکنواخت Uniform Probability Distribution هستند. بر اساس قضیه حد مرکزی می توان گفت که اگر ما متغیر تصادفی جدیدی تعریف کنیم به طوری که ، سپس می توان اثبات کرد که فارغ از نوع توزیع احتمالی اولیه ی متغیر های تصادفی در این مثال توزیع یکنواخت توزیع احتمال متغیر جدید، توزیع نرمال خواهد بود.

دنباله از متغیرهای تصادفی مستقل با توزیع یکنواخت پیوسته را که بر یک فضای احتمال تعریف شده‌اند در نظر بگیرید. فرض کنید میانگین برابر و انحراف از معیار آن است. حالا سری را در نظر بگیرید. می‌دانیم که میانگین برابر و انحراف از معیار آن است. بر اساس قضیه حد مرکزی در بی نهایت به سمت توزیع نرمال میل می‌کند

توزیع طبیعی ، یکی از مهمترین توزیع‌های احتمالی پیوسته در نظریه احتمالات است. علت نام‌گذاری و همچنین اهمیت این توزیع، هم‌خوانی بسیاری از مقادیر حاصل شده، هنگام نوسان‌های طبیعی و فیزیکی پیرامون یک مقدار ثابت با مقادیر حاصل از این توزیع است. دلیل اصلی این پدیده، نقش توزیع طبیعی در قضیه حد مرکزی است. به زبان ساده، در قضیه حد مرکزی نشان داده میشود که تحت شرایطی، مجموع مقادیر حاصل از متغیرهای مختلف که هرکدام میانگین و پراکندگی متناهی دارند، با افزایش تعداد متغیرها، دارای توزیعی بسیار نزدیک به توزیع طبیعی است. این قانون که تحت شرایط و مفروضات طبیعی نیز برقرار است، سبب شده که برایند نوسان‌های مختلف تعداد زیادی از متغیرهای ناشناخته، در طبیعت به صورت توزیع طبیعی آشکار شود. بعنوان مثال، با اینکه متغیرهای زیادی بر میزان خطای اندازه‌گیریِ یک کمیت اثر میگذارند،)مانند خطای دید، خطای وسیله اندازه‌گیری، شرایط محیط و ...(اما با اندازه‌گیری های متعدد، برایند این خطاها همواره دارایِ توزیع طبیعی است که حول مقدار ثابتی پراکنده شده است.مثال‌های دیگری از این نوسان‌های طبیعی، طول قد، وزن یا بهره هوشی افراد است.

این توزیع گاهی به دلیل استفاده کارل فردریک گاوس از آن در کارهای خود با نام توزیع یا تابع گوسی (گاوسی) نامیده می‌شود؛ همچنین به دلیل شکل تابع احتمال این توزیع، با نام انحنای زنگوله‌ای نیز معروف است.

www.freebay.ir

اگر تعداد اعداد کمتر از میانگین برابر تعداد اعداد بالاتر از میانگین در سری اعداد مرتب شده صعودی باشد . توزیع مورد نظر نرمال است.

تابع چگالی احتمال توزیع نرمال

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu^2)/2\sigma^2} \qquad e=2.7$$

تابع چگالی احتمال توزیع طبیعی با پارامترهای μ و σ^2 به صورت زیر است

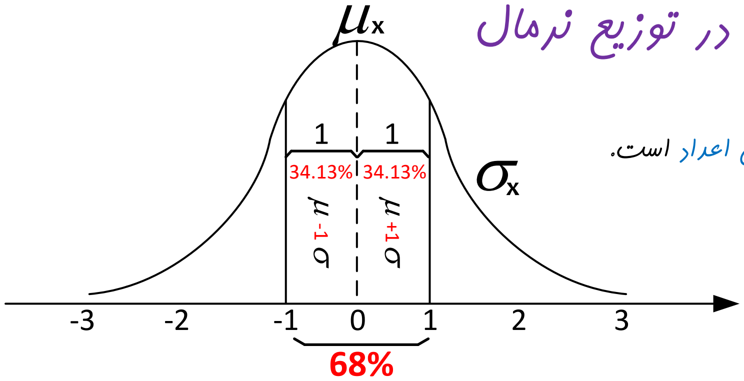
تابع چگالی احتمال ، متقارن حول $X=\mu$ است.

نقاط عطف این منحنی $X=\mu+\sigma \quad X=\mu-\sigma$

تابع چگالی احتمال ، بینهایت بار مشتق پزیر است

همچنین این نقطه میانگین، مد و میانهء توزیع است

تحلیل انحراف معیار در توزیع نرمال



انحراف معیار یا خطای استاندارد Deviation Standard، معیار گستردگی اعداد است.

۶۸٪ از مقدارها در محدوده یک انحراف معیار از میانگین هستند.

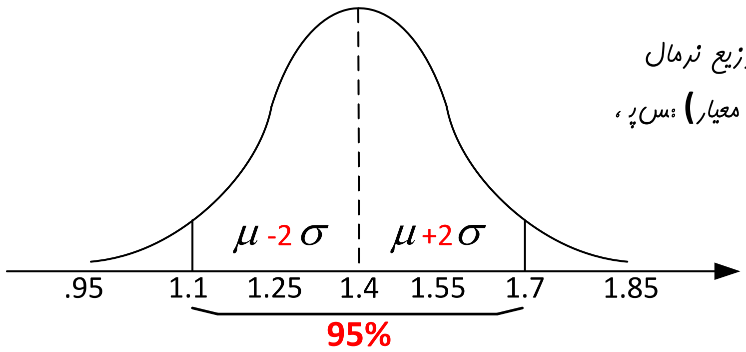
۹۵٪ از مقدارها در محدوده دو انحراف معیار از میانگین هستند.

۹۹,۷٪ از مقدارها در محدوده سه انحراف معیار از میانگین هستند.

مثال :

اگر قد ۹۵٪ از دانش آمزان در مدرسه بین ۱,۱ الی ۱,۷ با در نظر گرفتن توزیع نرمال

۹۵٪ ،برابر دو انحراف از معیار در طرفین میانگین است (مجموع ۴ انحراف معیار) پس: ،



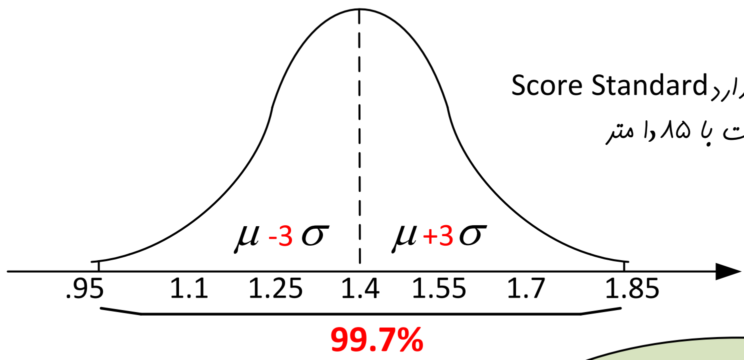
μ= (1.7+1.1)/2=1.4m

σ= (1.7-1.1)/4=0.15m

نمره معیار

تعداد انحراف ها از میانگین همچنین با نام های – نمره معیار یا نمره استاندارد Score Standard

"سیگما" و یا "نمره Z" ذکر می شوند.اگر قد یکی از دانش آموزان برابر است با ۱٫۸۵ متر



در روی نمودار زنگوله ای ۱٫۸۵ متر، ۳ انحراف از میانگین ۱٫۴ متر دارد.

$$Z=\mu+3\sigma=1.85 \rightarrow \sigma=\frac{1.85-\mu}{3}=\frac{1.85-1.4}{3}=0.15$$

با این انحراف معیار توزیع نرمال استاندارد خواهیم داشت.

$X \rightarrow Z$

تبدیل توزیع نرمال به توزیع نرمال استاندارد

$$\boxed{Z_i=\frac{X_i-\mu_x}{\sigma}}$$

یک نظرسنجی از مدت زمان مسافرت، این

مقادیر را به دقیقه در بر دارد.

۲۶, ۳۳, ۶۵, ۲۸, ۳۴, ۵۵, ۲۵, ۴۴, ۵۰, ۳۶, ۲۶, ۳۷, ۴۳, ۶۲, ۳۵, ۳۸, ۴۵, ۳۲, ۲۸, ۳۴

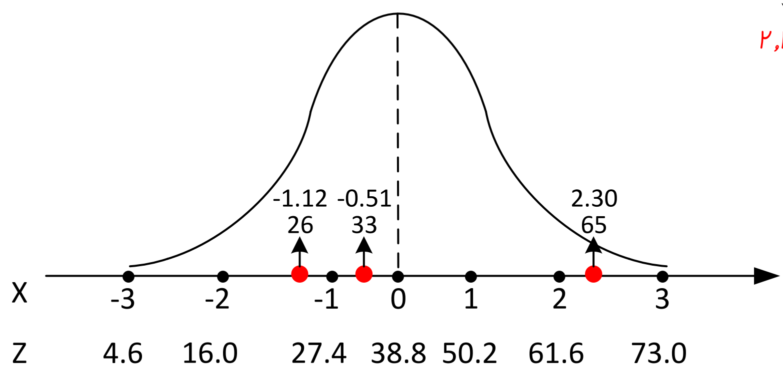
(26-38.5)/11.4=-1.12

-۱,۱۲   ۲,۳۰

$\sum X_i$ = 776

μ_X = 38.8

σ= 11.4



www.freebay.ir

دیاگرام رسیدن از آماره به پارامتر با اطلاعات حاصل از نمونه

برای رسیدن از آماره به پارامتر باید اطلاعات حاصل از نمونه را به جامعه تعمیم دهیم که به یکی از دو روش تخمین آماری و آزمون آماری انجام میگردد.

در فرضیه نویسی پایان نامه ها ، آزمون آماری بیشترین کاربرد را دارد.

در آمار پارامتریک توزیع نمونه گیری نرمال است.

Population
پارامتر
Parameter

Sample
آماره
Statistics

پارامتر – جامعه آماری

جامعه	N
میانگین	μ
انحراف معیار	σ
نسبت	π

تحلیل آماری با استفاده از آمار استنباطی

آماره – فضای نمونه

نمونه	n
میانگین	$\bar{x}$
انحراف معیار	S
نسبت	P

میانگین حسابی ساده

$$\mu_X = \frac{\sum_{i=1}^{N} X_i}{N}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + \cdots\cdots + x_n)$$

انحراف معیار

$$\sigma_x = \sqrt{\frac{\sum (X_i - \mu_x)^2}{N}}$$

$$\sigma_x = \sqrt{\frac{\sum X_i^2}{N} - \mu_x^2}$$

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

حد مرکزی

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

$X \rightarrow Z$

تبدیل توزیع نرمال به توزیع نرمال استاندارد

$$Z_i = \frac{X_i - \mu_x}{\sigma}$$

تابع چگالی احتمال توزیع نرمال

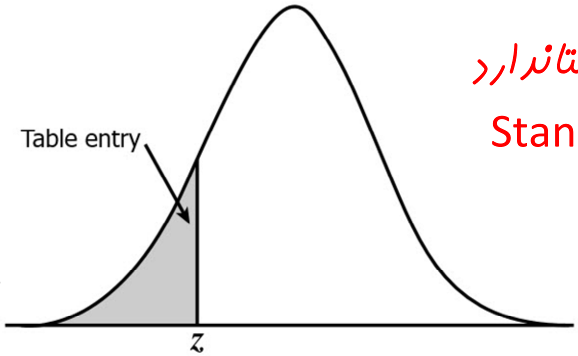
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}$$

$e = 2.7$

www.freebay.ir

جدول احتمال توزیع نرمال استاندارد

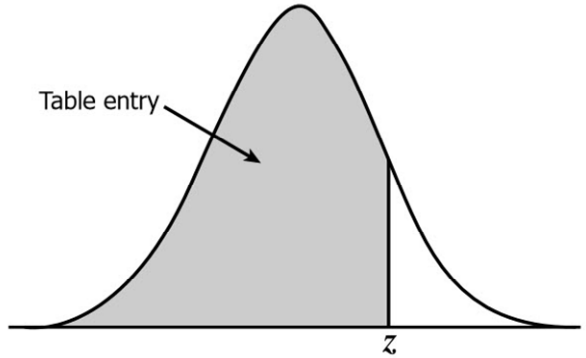
Standard Normal Probalilities



ρ 0.0003
Z -3.40

ρ=0.0002
Z=-3.49

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



ρ=0.0000
Z =0

ρ=0.5359
Z= .09

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

www.freebay.ir

تخمین فاصله ای در جامعه آماری
Estimation Interval

- اگر تحقیق از نوع سوالی و صرفا حاوی پرسش درباره پارامترها باشد برای پاسخ به سوالات از تخمین آماری استفاده میشود
- اگر حاوی فرضیه ها بوده و از مرحله سوال گذر کرده باشد از آزمون فرضیه ها و فنون آماری با نام آزمون آماری استفاده میشود.
- تخمین ها بر دو نوع کلی فاصله ای و نقطه ای می باشند که در چندین حالات بر اساس ویژگی های جامعه . مورد مطالعه قرار میگیرند ،
- در آمار بیشتر از تخمین فاصله ای استفاده میشود. و دارای یک بازه برای تعیین حد بزرگ UCL و حد کوچک LCL است.
- وقتی به سراغ تخمین میرویم که پارامتر میانگین جامعه برای ما مشخص نیست.

توزیع نرمال در جامعه آماری با انحراف معیار معلوم

$$\rho\left(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \leq \mu_x \leq \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}\right) = 1-\alpha$$

این فرمول برای نمونه بالاتر از ۳۰ استفاده میشود

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

انحراف معیار جامعه

تعداد نمونه

α : خطای نوع اول که در تحقیق و پایان نامه عموما برابر 5% است.

خطا	۱۰٪	۵٪	۱٪
اطمینان	۹۰٪	۹۵٪	۹۹٪

$1-\alpha$: فاصله یا سطح اطمینان $\alpha\%$ - ۱۰۰٪

ρ : با احتمال خطا تخمین زده میشود.
ممکن است جهت حصول اطمینان بیشتردر سه بازه 99%,95%,90% تخمینهای برست آمده مقایسه شوند.

داشته ها / مفروضات: σ_x، n، $\bar{x}$، $1-\alpha$، α $\xrightarrow{\text{محاسبه}}$ $\sigma_{\bar{x}}$، ε $\xrightarrow{\text{محاسبه}}$ UCL & LCL

$\bar{x}$: میانگین نمونه

σ_x : انحراف معیار جامعه

$\sigma_{\bar{x}}$: انحراف معیار نمونه

$Z_{\frac{\alpha}{2}}$: مراجعه به جدول نرمال Z برای پیدا کردن احتمال
چون خطا شامل دو انتهای بازه میگردد آلفا تقسیم بر دو شده است.

μ_x : میانگین جامعه؟

$\varepsilon = Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}$: میزان اپسیلونی که به میانگین اضافه و کم میگردد.

$\bar{x}+\varepsilon = UCL$
$\bar{x}-\varepsilon = LCL$
: حد بالا و پایین که باید محاسبه شود.

www.freebay.ir

توزیع نرمال در جامعه آماری با انحراف معیار معلوم

1

سوال

انحراف معیار جامعه معلوم

- بررسی ها نشان میدهد که توزیع وزن شیشه های تولید شده در یک کارخانه بزرگ نرمال و انحراف معیار آن ۲۱ تن است .از آنجا که سنجش وزن شیشه ها بطور روزانه ممکن نیست یک نمونه ۵۰ قطعه ای از تولیدات انتقاب شده است که میانگین وزن آن ۸۷۱ تن است در سطح اطمینان ۹۰ درصد ، میانگین واقعی شیشه های تولید شده را در طی یک روز محاسبه کنید؟

$$\rho(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_x \le \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}) = 1-\alpha$$

$\sigma_x = ۲۱$

$n = ۵۰$

$\bar{x} = ۸۷۱$

$1-\alpha = ۹۰\%$

$\alpha = ۱۰\% = ۰/۱$

$$\sigma_{\bar{x}} = \frac{۲۱}{\sqrt{۵۰}} = ۲/۹۶$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{۰/۱}{2}} = Z_{۰/۰۵} = ۱/۶۴۵$$

$$\varepsilon = 1.645*2.96$$

$$\rho(871-\varepsilon \le \mu_x \le 871+\varepsilon) = 0.9$$

$$\rho(866.11 \le \mu_x \le 875.89) = 0.9$$

تحلیل

با اطمینان ۹۰ درصد میتوان ادعا نمود که میانگین واقعی وزن شیشه های تولید شده بین ۸۶۶ و ۸۷۵ تن باشد و فقط ۱۰ درصد احتمال خطا دارد که میانگین مورد نظر خارج از این محدوده باشد .

۵ درصد احتمال دارد این میانگین بالاتر از ۸۷۵ یا پایین تر از ۸۶۶ تن باشد.

2

سوال

انحراف معیار جامعه معلوم

- مطالعات نشان میدهد توزیع معدل دانشجویان در واحد علوم و تحقیقات نرمال بوده و از انحراف معیار ۱,۵ برخوردار است . یک نمونه ۲۵ نفری از دانشجویان رشته های مختلف انتخاب شده که میانگین نمرات آنها ۱۳,۷۵ میباشد . در سطح اطمینان ۹۵ درصد میانگین واقعی نمرات دانشجویان را تخمین بزنید؟

$$\rho(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_x \le \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}) = 1-\alpha$$

$\sigma_x = ۱/۵$

$n = ۲۵$

$\bar{x} = ۱۳/۷۵$

$1-\alpha = ۹۵\%$

$\alpha = ۵\% = ۰/۰۵$

$$\sigma_{\bar{x}} = \frac{۱/۵}{\sqrt{۲۵}} = ۰/۳$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{۰/۰۵}{2}} = Z_{۰/۰۲۵} = ۱/۹۶$$

$$\varepsilon = ۱/۹۶*۰/۳$$

$$\rho(13.75-\varepsilon \le \mu_x \le 13.75+\varepsilon) = 0.95$$

$$\rho(13.162 \le \mu_x \le 14.338) = 0.95$$

تحلیل

با اطمینان ۹۵ درصد میتوان ادعا نمود که میانگین واقعی نمرات دانشجویان بین ۱۳,۱۶ و ۱۴,۳۳ تن باشد و فقط ۵ درصد احتمال خطا دارد که میانگین مورد نظر خارج از این محدوده باشد .

۲,۵ درصد احتمال دارد این میانگین بالاتر از ۱۴,۳۳ یا پایین تر از ۱۳,۱۶ تن باشد.

3

سوال

انحراف معیار جامعه معلوم

یک ترمینال مسافربری در ایام نوروز طی یک تحقیق ۶۴ اتوبوس را بصورت تصادفی مورد مطالعه قرار داده که میانگین سرعت آنها ۱۰۰ کیلومتر در ساعت بوده است .با فرض انحراف معیار ۱۰ کیلومتر در ساعت ، میانگین سرعت حرکت اتوموبیل ها را تخمین بزنید؟

$\sigma_x = 10$

$n = 64$

$\bar{x} = 100$

$1-\alpha = 95\%$

$\alpha = 5\% = 0.05$

$$\rho(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_x \le \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}) = 1-\alpha$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96 \quad \sigma_{\bar{x}} = \frac{10}{\sqrt{64}} = 1.25 \quad \varepsilon = 1.96*1.25$$

$$\rho(100-\varepsilon \le \mu_x \le 100+\varepsilon) = 0.95 \quad \rho(97.55 \le \mu_x \le 102.45) = 0.95$$

تحلیل

با اطمینان ۹۵ درصد میتوان ادعا نمود که سرعت اتوموبیل ها بین ۹۷,۵۵ و ۱۰۲,۴۵ کیلومتر در ساعت میباشد و تنها ۵ درصد احتمال دارد که خارج از این محدوده باشد و بعبارت دیگر ۲,۵ درصد احتمال میانگین سرعت بالاتر از ۱۰۲,۴۵ و یا پایینتر از ۹۷,۵۵ کیلومتر در ساعت باشد.

www.freebay.ir

Bay

توزیع نرمال در جامعه آماری با انحراف معیار معلوم با درصد اطمینان های
90% 95% 99%

هر چقدر میزان خطا کمتر شود بازه تخمین بیشتر میگردد.

1

انحراف معیار جامعه معلوم

سوال

– تحقیقات انجام شده حاکی از آن است که توزیع قد دانشجویان نرمال بوده و دارای انحراف معیار ۱۰ سانتی متر باشد . یک نمونه ۳۶ نفری از دانشجویان انتخاب شده است که میانگین قد آنها ۱۶۹ سانتی متر میباشد . در سطوح اطمینان ۹۰ درصد ، ۹۵ درصد ، ۹۰ درصد طول واقعی دانشجویان را تخمین بزنید.

$$P(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_x \le \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}) = 1-\alpha$$

فاصله ۵٫۴۶ UCL و LCL

بازه ۹۰ درصد اطمینان

$\sigma_x = ۱۰$

$n = ۳۶$

$\bar{x} = ۱۶۹$

$1-\alpha = ۹۰\%$

$= ۱۰\% = ۰/۱$

$$\sigma_{\bar{x}} = \frac{۱۰}{\sqrt{۳۶}} = 1.66$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = Z_{0.05} = ۱/۶۴۵$$

$$\varepsilon = 1.66*1.645=2.73$$

$$P(169-\varepsilon \le \mu_x \le 169+\varepsilon) = 0.9$$

$$P(166.27 \le \mu_x \le 171.73) = 0.9$$

تحلیل ۹۰ درصد

با اطمینان ۹۰ درصد میتوان ادعا نمود که میانگین واقعی قد دانشجویان بین ۱۶۶٫۲۷ و ۱۷۱٫۷۳ سانتی متر باشد و فقط ۱۰ درصد احتمال خطا دارد که میانگین مورد نظر خارج از این محدوده باشد .
۵ درصد احتمال دارد این میانگین بالاتر از ۱۷۱٫۷۳ یا پایین تر از ۱۶۶٫۲۷ سانتی متر باشد.

فاصله ۶٫۵۰ UCL و LCL

بازه ۹۵ درصد اطمینان

$\sigma_x = ۱۰$

$n = ۳۶$

$\bar{x} = ۱۶۹$

$1-\alpha = ۹۵\%$

$\alpha = ۵\% = ۰/۰۵$

$$\sigma_{\bar{x}} = \frac{۱۰}{\sqrt{۳۶}} = 1.66$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{۰/۰۵}{2}} = Z_{۰/۰۲۵} = ۱/۹۶$$

$$\varepsilon = 1.66*1.96=3.25$$

$$P(169-\varepsilon \le \mu_x \le 169+\varepsilon) = 0.95$$

$$P(165.75 \le \mu_x \le 172.25) = 0.95$$

تحلیل ۹۵ درصد

با اطمینان ۹۵ درصد میتوان ادعا نمود که میانگین واقعی قد دانشجویان بین ۱۶۵٫۷۵ و ۱۷۲٫۲۵ سانتی متر باشد و فقط ۵ درصد احتمال خطا دارد که میانگین مورد نظر خارج از این محدوده باشد .
۲٫۵ درصد احتمال دارد این میانگین بالاتر از ۱۷۱٫۲۵ یا پایین تر از ۱۶۵٫۷۵ سانتی متر باشد.

فاصله ۸٫۵۲ UCL و LCL

بازه ۹۹ درصد اطمینان

$\sigma_x = ۱۰$

$n = ۳۶$

$\bar{x} = ۱۶۹$

$1-\alpha = ۹۹\%$

$\alpha = ۱\% = ۰/۰۱$

$$\sigma_{\bar{x}} = \frac{۱۰}{\sqrt{۳۶}} = 1.66$$

$$Z_{\frac{\alpha}{2}} = Z_{\frac{۰/۰۱}{2}} = Z_{۰/۰۰۵} = ۲/۵۷$$

$$\varepsilon = 1.66*2.57=4.26$$

$$P(169-\varepsilon \le \mu_x \le 169+\varepsilon) = 0.99$$

$$P(164.74 \le \mu_x \le 173.26) = 0.99$$

تحلیل ۹۹ درصد

با اطمینان ۹۹ درصد میتوان ادعا نمود که میانگین واقعی قد دانشجویان بین ۱۶۴٫۷۴ و ۱۷۳٫۲۶ سانتی متر باشد و فقط ۱ درصد احتمال خطا دارد که میانگین مورد نظر خارج از این محدوده باشد .
۲٫۵ درصد احتمال دارد این میانگین بالاتر از ۱۱۷۳٫۲۶ یا پایین تر از ۱۶۴٫۷۴ سانتی متر باشد.

www.freebay.ir

توزیع تی-استیودنت

distribution-t s'Student

degrees of freedom =df درجه آزادی

برای ارزیابی میزان هم‌قوارگی یا یکسان بودن و نبودن میانگین نمونه‌ای با میانگین جامعه در حالتی به کار می‌رود که انحراف معیار جامعه مجهول باشد چون توزیع t در مورد نمونه‌های کوچک با استفاده از درجات آزادی تعدیل می‌شود، می‌توان از این آزمون برای نمونه‌های بسیار کوچک استفاده نمود.

در هنگام تعیین تقریبی میانگین نمونه‌های برداشته شده از یک متغیر تصادفی، توزیع تی-استودنت مطرح می‌شود. این توزیع اساس آزمونی به نام «تست تی» است که مقدار اطمینان از تفاوت دو متغیر تصادفی را از روی نمونه‌هایشان اعلام می‌کند.

تعداد نمونه وقتی از ۳۰ بیشتر میشود توزیع **T** به توزیع **Z** تبدیل میشود

مثال $t_{\frac{\alpha}{2},df} \left\{ \begin{array}{l} \text{If } n=10 \\ \alpha = 0.05 \end{array} \right. \Rightarrow t_{\frac{0.05}{2},9} = t_{0.025,9} = 2.262$

$\frac{\alpha}{2}=0.025$

t Table

n-1=9

df=n-1

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%

Confidence Level

www.freebay.ir

فرمول تخمین فاصله ای در جامعه آماری

Estimation Interval

توزیع تی t در جامعه آماری با انحراف معیار نامعلوم

$$\rho(\bar{x} - t_{\frac{\alpha}{2},df} S_{\bar{x}} \leq \mu_x \leq \bar{x} + t_{\frac{\alpha}{2},df} S_{\bar{x}}) = 1-\alpha$$

این فرمول برای نمونه کمتر از ۳۰ استفاده میشود

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + \cdots\cdots + x_n)$$

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

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

df : درجه آزادی برابر n-1

α : خطای نوع اول که در تحقیق و پایان نامه عموما برابر 5% است.

خطا	۱۰٪	۵٪	۱٪
اطمینان	۹۰٪	۹۵٪	۹۹٪

$1-\alpha$: فاصله یا سطح اطمینان $\% \alpha -$ ۱۰۰٪

ρ : با احتمال خطا تخمین زده میشود.

ممکن است جهت حصول اطمینان بیشترر سه بازه 99%,95%,90% تخمینهای برست آمده مقایسه شوند.

داشته ها / مفروضات: x_i ، n ، $1-\alpha$ ، α $\xrightarrow{\text{محاسبه}}$ $S_{\bar{x}}$ ، ε $\xrightarrow{\text{محاسبه}}$ UCL & LCL

$\bar{x}$: میانگین نمونه

x_i : مقادیر نمونه

σ_x : انحراف معیار جامعه

$S_{\bar{x}}$: انحراف معیار نمونه

$t_{\frac{\alpha}{2},df}$: مراجعه به جدول **t** برای پیدا کردن احتمال
چون خطا شامل دو انتهای بازه میگردد آلفا تقسیم بر دو شده است.

μ_x : میانگین جامعه؟

$\varepsilon = t_{\frac{\alpha}{2},df} S_{\bar{x}}$: میزان اپسیلونی که به میانگین اضافه و کم میگردد.

$\bar{x} + \varepsilon = UCL$
$\bar{x} - \varepsilon = LCL$
: حد بالا و پایین که باید محاسبه شود.

www.freebay.ir

توزیع تی نرمال در جامعه آماری با انحراف معیار نامعلوم

1

سوال

انحراف معیار جامعه نامعلوم

بازاریابی ، درصدد بررسی و برآورد قدرت خرید ساکنان یک محله در تهران میباشد او ناچار است از یک نمونه تصادفی ۱۰ تایی از بین خریداران انتخاب و قدرت خرید هر یک را اندازه گیری نماید . قدرت خرید نمونه فوق بر حسب ده هزار تومان و بدین شکل میباشد **X=8,7,5,4,12,15,10,13,14,12** قدرت خرید ساکنان محله از توزیع نرمال برخوردار است . در سطح اطمینان ۹۵ درصد میانگین قدرت خرید را برآورد نمائید؟

$$\rho\left(\bar{x} - t_{\frac{\alpha}{2},df} S_{\bar{x}} \le \mu_x \le \bar{x} + t_{\frac{\alpha}{2},df} S_{\bar{x}}\right) = 1-\alpha$$

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \quad S_x = \sqrt{\frac{132}{9}} = 3.83$$

$$t_{\frac{\alpha}{2},df} = t_{\frac{0.05}{2},9} = t_{0.025,9} = 2.262$$

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}} = \frac{3.83}{\sqrt{10}} = 1.211 \qquad \varepsilon = 1.211*2.262$$

$$n = 10$$
$$x = 8,7,5,4,12,15,10,13,14,12 \quad \bar{x} = 10$$
$$1-\alpha = 95\%$$
$$\alpha = 5\% = 0.05$$

$$\rho(10-\varepsilon \le \mu_x \le 10+\varepsilon) = 0.95 \qquad \rho(7.261 \le \mu_x \le 12.739) = 0.95$$

تحلیل

با اطمینان ۹۵ درصد میتوان ادعا نمود که قدرت خرید ساکنان محله بین۷۲۶۱۰ و ۱۲۷۳۹۰ تومان میباشد و فقط ۵ درصد احتمال دارد که قدرت خرید ساکنین خارج از این محدوده باشد. یا بعبارت دیگر ۲,۵ درصد احتمال دارد قدرت خرید بالاتر از ۱۲,۷۳۹ و یا پایینتر از ۷,۲۶۱ باشد.

2

سوال

انحراف معیار جامعه نامعلوم

محققی برنبال برآورد سطح عملکرد کارکنان یک سازمان میباشد او بصورت تصادفی ۹ نمونه از بین کارکنان انتخاب کرده که امتیاز عملکرد آنها بدین شرح میباشد ۱۸، ۳۲، ۲۷، ۲۹، ۲۷، ۲۸، ۳۰، ۳۱، ۳۰ با فرض نرمال بودن سطح عملکرد کارکنان ، امتیاز عملکرد کل کارکنان را در سطح اطمینان ۹۹ درصد تخمین بزنید؟

$$\rho\left(\bar{x} - t_{\frac{\alpha}{2},df} S_{\bar{x}} \le \mu_x \le \bar{x} + t_{\frac{\alpha}{2},df} S_{\bar{x}}\right) = 1-\alpha$$

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \quad S_x = \sqrt{\frac{136}{8}} = 4.12$$

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}} = \frac{4.12}{\sqrt{9}} = 1.37$$

$$n = 9$$
$$x = 18,32,27,29,27,28,30,31,30 \quad \bar{x} = 28$$
$$1-\alpha = 99\%$$
$$\alpha = 1\% = 0.01$$

$$t_{\frac{\alpha}{2},df} = t_{\frac{0.01}{2},8} = t_{0.005,8} = 3.355 \qquad \varepsilon = 1.37*3.35=4.58$$

$$\rho(28-\varepsilon \le \mu_x \le 28+\varepsilon) = 0.99 \qquad \rho(23.42 \le \mu_x \le 32.58) = 0.99$$

تحلیل

با اطمینان ۹۹ درصد میتوان ادعا نمود که قدرت خرید ساکنان محله بین ۲۳,۴۲ و ۳۲,۵۸ واحد میباشد و فقط ۱ درصد احتمال دارد که قدرت خرید ساکنین خارج از این محدوده باشد. یا بعبارت دیگر ۰,۵ درصد احتمال دارد قدرت خرید بالاتر از ۳۲,۵۸ و یا پایینتر از ۲۳,۴۲ باشد.

www.freebay.ir

تخمین فاصله ای تفاضل میانگین دو جامعه
Estimation Interval

فرمول توزیع نرمال در دو جامعه با انحراف معیار معلوم

$$P\left[(\bar{x}_1-\bar{x}_2)-Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}\right]=1-\alpha$$

$$\sigma_{\bar{x}_1\bar{x}_2}=\sqrt{\frac{\sigma^2_{x_1}}{n_1}+\frac{\sigma^2_{x_2}}{n_2}}$$

تحلیل هر سه توزیع

تحلیل هر سه توزیع

		UCL		LCL
اگر هر دو دامنه مثبت باشد	1	دامنه دوم +	$\mu_{x_1} > \mu_{x_2}$	+ دامنه اول
اگر هر دو دامنه منفی باشد	2	دامنه دوم -	$\mu_{x_1} < \mu_{x_2}$	- دامنه اول
اگر یک دامنه منفی و دیگری مثبت باشد	3	دامنه دوم +	اختلاف معنی دار نیست	- دامنه اول

فرمول توزیع تی t در دو جامعه با انحراف معیار نا معلوم و برابر

$$P\left[(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2},df}\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2},df}\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha$$

$$S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

$$n_1+n_2-2=df$$

فرمول توزیع نرمال در دو جامعه با انحراف معیار نا معلوم و نابرابر

$$P\left[(\bar{x}_1-\bar{x}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right]=1-\alpha$$

www.freebay.ir

توزیع نرمال در دو جامعه با انحراف معیار معلوم

1

دو جامعه با انحراف معیار معلوم

سوال ۱

محقق قصد دارد عملکرد مدیران دو موسسه را مورد مقایسه قرار دهد ، این محقق از موسسه آلفا یک نمونه تصادفی ۲۵ نفره انتقاب که میانگین امتیاز عملکرد آن ۶۰ میباشد و از سازمان بتا یک نمونه تصادفی ۲۰ نفره انتقاب کرده که میانگین آن ۵۵ میباشد . با فرض نرمال بودن جامعه انحراف معیار آلفا و بتا برابر ۱۰ و ۱۲ ،نحوه عملکرد این دو موسسه را در سطح اطمینان ۹۹ درصد مقایسه کنید؟

$$\rho\left[(\bar{x}_1-\bar{x}_2)-Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}\le \mu_{x_1}-\mu_{x_2}\le(\bar{x}_1-\bar{x}_2)+Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}\right]=1-\alpha$$

α: $n_1=25$, $\bar{x}_1=60$, $\sigma_{x_1}=10$

β: $n_2=20$, $\bar{x}_2=55$, $\sigma_{x_2}=12$

$$\sigma_{\bar{x}_1\bar{x}_2}=\sqrt{\frac{\sigma_{x_1}^2}{n_1}+\frac{\sigma_{x_2}^2}{n_2}}\qquad \sigma_{\bar{x}_1\bar{x}_2}=\sqrt{\frac{100}{25}+\frac{144}{20}}=3.34$$

$$Z_{\frac{\alpha}{2}}=Z_{0.005}=2.58$$

$$Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}=8.617$$

$$\rho\left[(60-55)-8.617\le \overset{\alpha}{\mu_{x_1}}-\overset{\beta}{\mu_{x_2}}\le(60-55)+8.617\right]=0.99$$

$$\rho\left[-3.617\le \mu_{x_1}-\mu_{x_2}\le 13.617\right]=0.99$$

تحلیل

چون یک دامنه منفی و یک دامنه مثبت است اختلاف عملکرد دو موسسه در سطح اطمینان ۹۹ در صد معنی دار نیست (حالت سوم)

www.freebay.ir

توزیع تی t در دو جامعه با انحراف معیار های نا معلوم و برابر

2

دو جامعه با انحراف معیار نامعلوم و برابر

سوال ۲

هدف محقق مقایسه عملکرد سازمان های آلفا و بتا میباشد ، در این تحقیق از سازمان آلفا یک نمونه ۹ تایی دارای میانگین ۴۵ و انحراف معیار ۱۲ و از موسسه بتا یک نمونه ۱۵ نفره با میانگین ۵۵ و انحراف معیار ۱۴ انتقاب شده است .توزیع نمره های سطح آمادگی کارکنان دو سازمان نرمال و واریانس دو جامعه یکسان است . در سطح اطمینان ۹۰ درصد تخمین لازم را برای مقایسه میانگین دو جامعه به عمل آورید؟

$$\rho\left[(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2},df}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le \mu_{x_1}-\mu_{x_2}\le(\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2},df}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha$$

$$t_{\frac{\alpha}{2},df}=t_{0.05,22}=1.717$$

df=9+15-2=22

$$S_p=\sqrt{\frac{8*144+14*196}{22}}=13.31\qquad \varepsilon=\underbrace{t_{\frac{\alpha}{2},df}S_p}_{22.853}\underbrace{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}_{0.421}\qquad \varepsilon=9.621$$

$1-\alpha=90\%=0.9$

$\alpha=10\%=0.1$

$n_1=9$

$n_2=15$

$\bar{x}_1=45$

$\bar{x}_2=55$

$S_1=12$

$S_2=14$

$$\rho\left[-10-\varepsilon\le \mu_{x_1}-\mu_{x_2}\le -10+\varepsilon\right]=0.90\qquad \rho\left[-19.621\le \overset{\alpha}{\mu_{x_1}}-\overset{\beta}{\mu_{x_2}}\le -0.379\right]=0.90$$

$\mu_x<\mu_x \longrightarrow \alpha$ سطح آمادگی $<$ β سطح آمادگی

تحلیل

چون هر دو دامنه منفی است . با اطمینان ۹۰ درصد میتوان گفت میانگین سازمان آلفا کمتر از سازمان بتا میباشد و به عبارتی سازمان بتا بهتر از سازمان آلفا است.(حالت دوم)

دو جامعه با انحراف معیار نامعلوم و نابرابر

سوال ۳

پژوهشگری در صدد است عملکرد کارکنان دو بانک را مورد مقایسه قرار دهد او از بانک اول ۱۶ نفر و از بانک دوم ۱۹ نفر را بصورت تصادفی انتخاب کرده که میانگین امتیاز عملکردشان به ترتیب ۲۸ و ۲۶ میباشد با فرض انحراف معیار نمونه اول برابر ۲ و انحراف معیار نمونه دوم برابر ۳ میباشد. عملکرد کارکنان این دو بانک را مورد مقایسه قرار دهید؟

چون هیچگونه فرضی برای واریانس های دو جامعه ذکر نشده است پس انحراف معیار ها را نابرابر فرض میکنیم.

3

دو جامعه با انحراف معیار نامعلوم و نابرابر

$$\rho\left[(\bar{x}_1-\bar{x}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right]=1-\alpha$$

$$z_{\frac{\alpha}{2}}=z_{0.025}=1.96$$

$$\sqrt{\frac{2^2}{16}+\frac{3^2}{19}}=\sqrt{\frac{4}{16}+\frac{9}{19}}=0.084$$

$$\varepsilon=0.84*1.96=1.64$$

$1-\alpha = 95\%=0.95$
$\alpha = 5\%=0.05$
$n_1 = 16$
$n_2 = 19$
$\bar{x}_1 = 28$
$\bar{x}_2 = 26$
$S_1 = 2$
$S_2 = 3$

$$\rho\left[(28-26)-\varepsilon \le \mu_{x_1}-\mu_{x_2} \le (28-26)+\varepsilon\right]=0.95$$

$$\rho\left[0.36 \le \mu_{x_1}-\mu_{x_2} \le 3.64\right]=0.95$$

تحلیل

با اطمینان ۹۵ درصد میتوان گفت عملکرد کارکنان دو بانک بین ۰٫۳۶ و ۳٫۶۴ خواهد بود و فقط ۵ درصد ممکن است خارج از این بازه باشد . بعبارتی ۲٫۵ درصد بالای ۳٫۶۴ و ۲٫۵ درصد پایین تر از ۰٫۳۶

www.freebay.ir

حل سوال ۳ با فرمول دو جامعه نامعلوم با انحراف معیار برابر

3

دو جامعه با انحراف معیار نامعلوم و نا برابر

$$\rho\left[(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2},df}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2},df}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha$$

$$t_{\frac{\alpha}{2},df}=t_{0.025,33}=2.04$$

$Df=16+19-2=33 \cong 30$

$$S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

$1-\alpha = 95\%=0.95$
$\alpha = 5\%=0.05$
$n_1 = 16$
$n_2 = 19$
$\bar{x}_1 = 28$
$\bar{x}_2 = 26$
$S_1 = 2$
$S_2 = 3$

$$S_p=\sqrt{\frac{15*4+18*9}{33}}==6.72$$

$$\varepsilon=\underbrace{t_{\frac{\alpha}{2},df}}_{2.04}S_p\underbrace{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}_{0.33} \quad \varepsilon=4.52$$

$$\rho\left[(28-26)-\varepsilon \le \mu_{x_1}-\mu_{x_2} \le (28-26)+\varepsilon\right]$$

$$\rho\left[-2.52 \le \mu_{x_1}-\mu_{x_2} \le 6.52\right]$$

تحلیل

اختلاف معنی دار نمی باشد.

حل سئوال ۳ با فرمول دو جامعه با انحراف معیار معلوم

3

دو جامعه با انحراف معیار مشخص

$$\rho\left[(\bar{x}_1-\bar{x}_2)-Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2} \le \mu_{x_1}-\mu_{x_2} \le (\bar{x}_1-\bar{x}_2)+Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}\right]=1-\alpha$$

$$\sigma_{\bar{x}_1\bar{x}_2}=\sqrt{\frac{\sigma^2_{x_1}}{n_1}+\frac{\sigma^2_{x_2}}{n_2}}$$

α	β
$n_1=16$	$n_2=19$
$\bar{x}_1=26$	$\bar{x}_2=28$
$\sigma_{x_1}=2$	$\sigma_{x_2}=3$

$$\sigma_{\bar{x}_1\bar{x}_2}=\sqrt{\frac{4}{16}+\frac{9}{19}}=0.84$$

$$Z_{\frac{\alpha}{2}}=Z_{0.025}=1.96$$

$$Z_{\frac{\alpha}{2}}\sigma_{\bar{x}_1\bar{x}_2}=2.38$$

$$\varepsilon=1.96*0.84=1.64$$

$$\rho\left[2-\varepsilon \le \overset{\alpha}{\mu}_{x_1}-\overset{\beta}{\mu}_{x_2} \le 2+\varepsilon\right]=0.95$$

$$\rho\left[0.36 \le \mu_{x_1}-\mu_{x_2} \le 3.64\right]=0.95$$

تحلیل

با اطمینان ۹۵ درصد میتوان گفت عملکرد کارکنان دو بانک بین ۰٫۳۶ و ۳٫۶۴ خواهد بود و فقط ۵ درصد ممکن است خارج از این بازه باشد . بعبارتی ۲٫۵ درصد بالای ۳٫۶۴ و ۲٫۵ درصد پایین تر از ۰٫۳۶

اختلاف **LCL ,UCL** در دو حالت تحلیل یکسانی دارند و نتیجه دو فرمول یکسان است.

نتیجه فرمول در این مثال توزیع نرمال در دو جامعه   با انحراف معیار های نا معلوم نابرابر

برابر است با

نتیجه توزیع نرمال در دو جامعه با انحراف معیار معلوم

www.freebay.ir

تخمین فاصله ای نسبت موفقیت جامعه (نسبت ، درصد ، کسر)

فرمول نسبت موفقیت جامعه : پایه و اساس تعیین حجم نمونه در مطالعات کیفی با مقیاس های اسمی و رتبه ای

$$\rho\left[\bar{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \le p \le \bar{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\right] = 1-\alpha$$

ρ تخمین با احتمال خطا p نسبت جامعه $\bar{p}$ نسبت نمونه

بعنوان مثال چند در صد از مخاطبین تلویزیونی یک برنامه خاص را مشاهده مینمایند.

سوال

هدف محققی تعیین نسبت افراد ناراضی در سازمان میباشد از آنجائیکه دسترسی به تمام افراد سازمان میسر نیست محققان ۴۰۰ نفر را بطور تصادفی انتخاب کرده اند که فقط ۳۲ نفر از کار خود ناراضی هستند . نسبت افراد ناراضی را در سطح خطای ۱ درصد برآورد نمائید؟

$\bar{p} = 32/400=0.08$

$Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.58$

$1-\alpha = 99\% = 0.99$

$\alpha = 1\% = 0.01$

$n = 400$

$x = 32$

$$\rho\left[\bar{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \le p \le \bar{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\right] = 1-\alpha$$

$$\sqrt{\frac{0.08*(1-0.08)}{400}} = 0.0135 \qquad \varepsilon = .0135*2.58=0.034$$

$$\rho\left[0.08 - \varepsilon \le p \le 0.08 + \varepsilon\right] = 0.99$$

$$\rho\left[0.045 \le p \le 0.115\right] = 0.99$$

تحلیل

نسبت افراد ناراضی در سازمان بین ۴,۵ درصد تا ۱۱,۵ درصد کل کارکنان می باشد ، بعبارت دیگر نسبت افراد ناراضی با ۹۹ درصد اطمینان حداکثر برابر ۱۱,۵ درصد و حداقل ۴,۵ درصد می باشد.

تعیین حجم نمونه n =?

روش های تعیین حجم نمونه

1. اصل تجربه (بعلت دقیق نبودن توصیه نمیشود)
   - میانگین سه تحقیق مشابه
   - استفاده از نسبت هزینه
2. استفاده از تکنیک های آماری (در پایان نامه و تحقیقها بالاترین استفاده را دارد)
   - برآورد میانگین (جامعه محدود 1، جامعه نامحدود 2)
   - برآورد نسبت (جامعه محدود 3، جامعه نامحدود 4)
3. استفاده از جدول مورگان (برای گزارشات درون سازمانی میشود استفاده کرد ولی به هیچ وجه جامعیت تکنیکهای آماری را ندارد)

www.freebay.ir

فرمول های روش های تعیین حجم نمونه

تعیین تعداد نمونه زمانی که هدف تحقیق برآورد میانگین جامعه باشد.

برای جامعه محدود

1

$$n=\frac{N\,Z_{\frac{\alpha}{2}}^{2}\,\sigma_x^{2}}{(N-1)\varepsilon^{2}+Z_{\frac{\alpha}{2}}^{2}\,\sigma_x^{2}}$$

حجم جامعه مشخص است N

برای جامعه نامحدود

2

$$n=\frac{Z_{\frac{\alpha}{2}}^{2}\,\sigma_x^{2}}{\varepsilon^{2}}$$

حجم جامعه نا مشخص است

ε دقت برآورد

(اگر مقدار اپسیلون در مسئله مشخص نبود برابر ۰,۰۷ در نظر میگیریم)

1

جامعه نامحدود مقدار N نامشخص

سوال

پژوهشگری علاقه مند است میانگین رشد کاری کارکنان یک سازمان را تعیین کنداو دقت برآورد را ۵ در نظر گرفته است. و تصور میکند انحراف معیار نمره های رشد کاری کارکنان ۲۰ نمره باشد اندازه نمونه را در سطح خطای ۵ درصد برآورد نمائید؟

$$Z_{\frac{\alpha}{2}}=\frac{۰/۰۵}{2}=Z_{۰/۰۲۵}=۱/۹۶$$

$$\sigma_x=20$$

$$n=\frac{Z_{\frac{\alpha}{2}}^{2}\,\sigma_x}{\varepsilon^{2}}=\frac{1.96^{2}*20^{2}}{5^{2}}=61.5\approx 62$$

هر عددی برست آمد به بزرگترین عدد بعدی روند (گرد) میکنیم.

تعیین تعداد نمونه زمانی که هدف تحقیق برآورد نسبت موفقیت باشد.

برای جامعه محدود

3

$$n=\frac{N\,Z_{\frac{\alpha}{2}}^{2}\,pq}{(N-1)\varepsilon^{2}+Z_{\frac{\alpha}{2}}^{2}\,pq}$$

برای جامعه نامحدود

4

$$n=\frac{Z_{\frac{\alpha}{2}}^{2}\,pq}{\varepsilon^{2}}$$

P نسبت موفقیت

q نسبت عدم موفقیت

P+q=1

(اگر مقدار P و qمشخص نبود هرکدام را ۰,۵ در نظر میگیریم)

2

جامعه نامحدود مقدار N نامشخص

سوال

مطالعه ای برای تعیین نسبت موفقیت مدیران وظیفه مدار در سطح سازمان های دولتی کشور برنامه ریزی شده است . این تصور وجود دارد که نسبت مزبور بزرگتر از ۰,۴۵ نیست . با حدود اطمینان ۹۵ درصد و سطح خطای مجاز ۰,۰۸ چند مدیر باید مورد مطالعه قرار گیرند؟

$$\varepsilon=0.08$$

$$Z_{\frac{\alpha}{2}}=\frac{۰/۰۵}{2}=Z_{۰/۰۲۵}=۱/۹۶$$

P=0.45 q=1-p=0.55

$$n=\frac{Z_{\frac{\alpha}{2}}^{2}\,pq}{\varepsilon^{2}}=\frac{۱/۹۶^{2}*0.45*0.55}{0.08^{2}}=148.6\approx 149$$

www.freebay.ir

جامعه محدود مقدار N مشخص

سوال

محققی قصد دارد در یک سازمان ۱۵۹۲ نفری اقدام به نمونه گیری کند و هیچ اطلاعات دیگری از تحقیقات پیشین بدست نیاورده است . مطلوبترین تعداد نمونه برای انجام این تحقیق چقدر است؟

$$Z_{\frac{\alpha}{2}}=\frac{۰/۰۵}{2}=Z_{۰/۰۲۵}=۱/۹۶$$

$\varepsilon=0.07$

P=0.5 & q=0.5

N=1592

$$n=\frac{N\,Z_{\frac{\alpha}{2}}^{2}\,pq}{(N-1)\varepsilon^{2}+Z_{\frac{\alpha}{2}}^{2}\,pq}=\frac{1592*1.96^{2}*0.5*0.5}{(1591-1)*0.07^{2}+1.96^{2}*0.5*0.5}=\frac{1528.9568}{8.7514}=174.70\approx175$$

جدول مورگان

یکی از ساده ترین راه های تعیین حجم نمونه استفاده از جدول مورگان است . در جدول مورگان حجم جامعه و نمونه متناظر با آن آمده است.

TABLE 1
Table for Determining Sample Size from a Given Population

N	*S*	*N*	*S*	*N*	*S*
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

Note.—*N* is population size.
S is sample size.

- اگر تعداد جامعه N نامشخص باشد نمیتوان از جدول مورگان استفاده نمود.
- این جدول دقیق نیست و فاصله طبقات در اعداد بزرگ بالا میباشد.

www.freebay.ir

آزمون فرضیه های آماری
Testing Hypothesis

اگر تحقیق از نوع سوالی و صرفا حاوی پرسش درباره پارامترها باشد برای پاسخ به سوالات از تخمین آماری استفاده می شود و اگر حاوی فرضیه ها بوده و از مرحله سوال گذر کرده باشد آزمون فرضیه ها و فنون آماری آن بکار برده می شود.

استفاده از آزمون زمانی است که علاوه بر سوال در تحقیق خود فرضیه نیز داشته باشیم که بطور کلی هدف از آزمون آماری آن است که با توجه به اطلاعات بدست آمده از داده های نمونه، حدس خود را در مورد جامعه به طور قوی رد یا قبول کنیم.

1 فرضیه صفر ، فرضیه مهم — H_0 — Null Hypothesis — همیشه به رد یا تایید فرضیه **H0** میپردازیم

2 فرضیه مقابل ، فرضیه آلترناتیو یا جانشین — H_1 — Alternative Hypothesis

1 در حالت عادی یعنی وقتی که ادعای فرضیه مساوی = ندارد آن را در **H1** قرار میدهیم و نقیض آن را در **H0** قرار میگیرد. — ادعا $= H_1$

سوال

نسبت مشارکت جو در سازمان بیش از ۷۰ در صد میباشد.

ادعا $H_1: \rho > 0.7$

نقیض ادعا $H_0: \rho \leq 0.7$ $\longrightarrow$ برای اثبات فرضیه باید اثبات کنیم که $\rho \leq 0.7$ نیست.

بعبارتی باید نقیض ادعا ثابت شود.

2 در حالت که ادعای فرضیه مساوی = دارد آن را در **H0** قرار میدهیم و نقیض آن را در **H1** قرار میگیرد. — ادعا $= H_0$

آنچه **H0** را آزمون پذیر میکند آنست که حتما باید حاوی علامت تساوی باشد
H0,H1 هر کدام میتوانند ادعا یا نقیض ادعا باشند.

سوال

میانگین معدل دانشجویان کلاس دست کم ۱۳ است.

ادعا $H_0: \mu \geq 13$ $\longrightarrow$ برای اثبات فرضیه باید اثبات کنیم که $\mu \geq 13$ است.

نقیض ادعا $H_1: \mu < 13$

بعبارتی باید ادعا ثابت شود.

مسئله حقوقی

در موارد حقوقی فرض بر برائت یا بیگناهی افراد است مگر اینکه عکس آن اثبات شود ، پس **H0** در اینحالت بیگناهی میباشد. و **H1** شامل الباقی مقادیر میباشد که همان گناهکار بودن است. پس از مشاهده شواهد و غیره به تایید یا رد **H0** پرداخته میشود.

ادعا $H_0:$ بی گناه بودن

نقیض ادعا $H_1:$ گناه کار بودن

ادعای قاضی

متهم = بی گناه بودن

سوال

m f

میانگین بهره وری خانم ها بیشتر از آقایان است.

ادعا $H_1: \mu_f > \mu_m$

نقیض ادعا $H_0: \mu_f \leq \mu_m$

برای اثبات فرضیه باید اثبات کنیم که میانگین بهره وری خانم ها کوچکتر یا مساوی آقایان نیست.

سوال

o y

میانگین بهره وری مدیران جوان در حد مدیران مسن است .

ادعا $H_0: \mu_y = \mu_o$

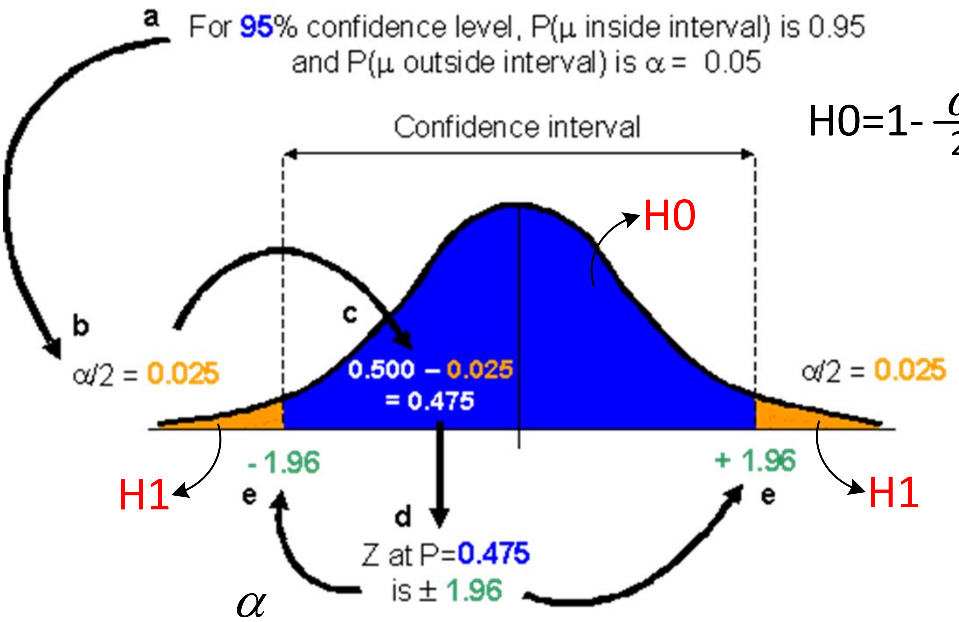
نقیض ادعا $H_1: \mu_y \neq \mu_o$

برای اثبات فرضیه باید اثبات کنیم که میانگین بهره وری مدیران جوان کمتر یا بیشتر از مدیران مسن نیست

www.freebay.ir

سطح معنی داری
Significant Level

روش کار این است که فرض H_0 را بنفع H_1 رد کنیم بشرط اینکه از یک آزمون آماری، مقداری بدست آوریم که احتمال وقوع آن مقدار با توجه به H_0 برابر یا کمتر از یک احتمال بسیار کوچک باشد . این احتمال کوچک را سطح معنی داری گویند



$$H0=1-\frac{\alpha}{2}$$

$$\text{If } \frac{\alpha}{2}=0.025 \quad H0=\begin{cases} 1-(2*0.025)=1-0.05=0.95 & \left(\frac{\alpha}{2}\right) \\ 2*(0.500-0.025)=2*0.475=0.95 \end{cases} \quad \alpha=0.05 \quad \boxed{P=0.95} \longrightarrow -1.96\leq Z \leq +1.96$$

$$\text{If } \frac{\alpha}{2}=0.05 \quad H0=\begin{cases} 1-(2*0.05)=1-0.1=0.9 & \left(\frac{\alpha}{2}\right) \\ 2*(0.500-0.05)=2*0.45=0.9 \end{cases} \quad \alpha=0.10 \quad \boxed{P=0.90} \longrightarrow -1.64\leq Z \leq +1.64$$

$$\text{If } \frac{\alpha}{2}=0.25 \quad H0=\begin{cases} 1-(2*0.5)=1-0.5=0.5 & \left(\frac{\alpha}{2}\right) \\ 2*(0.500-0.25)=2*0.25=0.5 \end{cases} \quad \alpha=0.50 \quad \boxed{P=0.50} \longrightarrow -0.67\leq Z \leq +0.67$$

خطاهای آماری

منفی کاذب
خطای نوع اول : α رد کردن **H0** در حالیکه **H0** درست است. ⟵ فرض صفر درست باشد و آزمون فرض آن را رد کند.

مثبت کاذب
خطای نوع دوم : β پذیرفتن **H0** در حالیکه **H0** غلط است. ⟵ فرض صفر درست نباشد و آزمون فرض آن را قبول کند.

تصمیم گیری

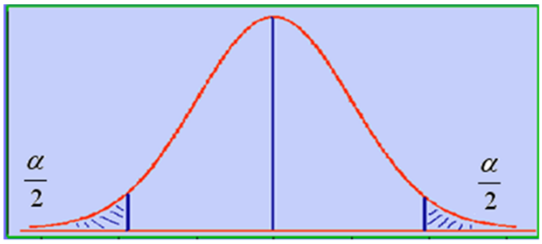
	قبول H_0	رد H_0
واقعیت H_0	تصمیم درست	خطای نوع اول
واقعیت H_1	خطای نوع دوم	تصمیم درست

واقعیت

تصمیم گیری: نتیجه گیری از نمونه	درست است	غلط است
H0 پذیرفته میشود	تصمیم درست است	β
H0 رد میشود	α	تصمیم درست است

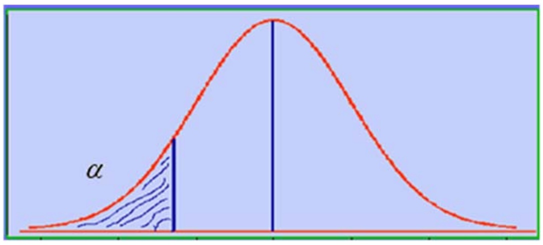
www.freebay.ir

خطاهای آماری با توجه به فرضیه پژوهش



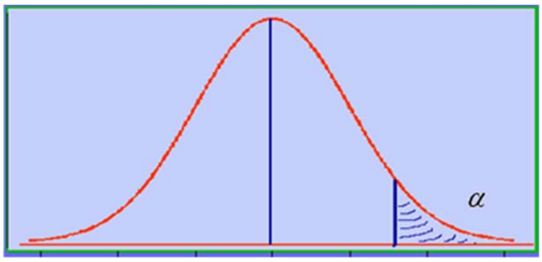
دو دنباله
Two Tailed

اگر علامت عدم تساوی در **H1** بود.



یک دنباله چپ
One Tailed Left

اگر علامت کوچکتر در **H1** بود دنباله راست است.



یک دنباله راست
One Tailed Right

اگر علامت بزرگتر در **H1** بود دنباله راست است.

سوال ۱

میانگین دستمزد کارکنان یک کارخانه کمتر از ۲۰۰۰۰ هزار تومان است.

نقیض ادعا $H_0: \mu \geq 20000$

ادعا $H_1: \mu < 20000$

→ H1 یک دنباله چپ

سوال ۲

مردان و زنان تفاوت سطح درجه هوشی ندارند.

ادعا $H_0: \mu_w = \mu_m$

نقیض ادعا $H_1: \mu_w \neq \mu_m$

→ H1 دو دنباله

سوال ۳

حداکثر ۴۰ درصد از مشتریان شرکت راضی هستند.

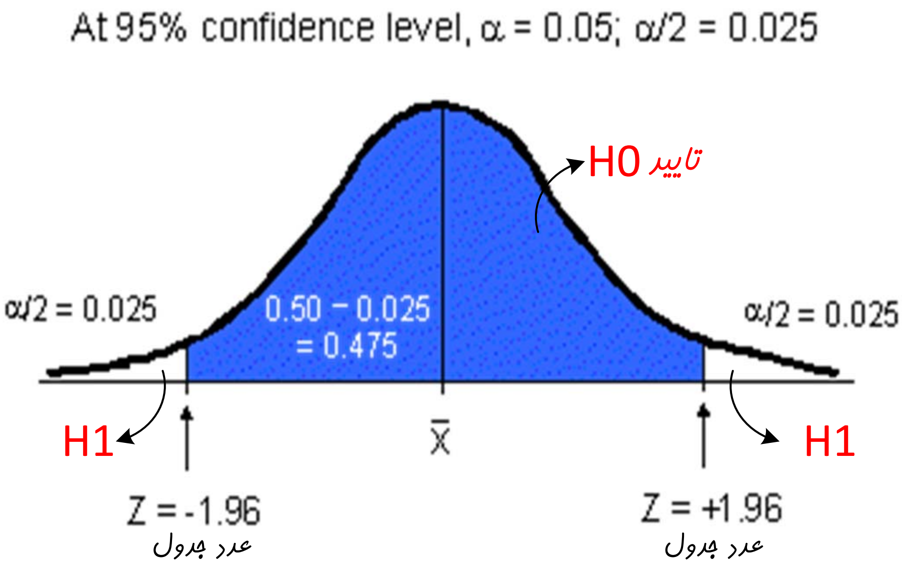
نقیض ادعا $H_0: \rho \leq 0.40$

ادعا $H_1: \rho > 0.40$

→ H1 یک دنباله راست

www.freebay.ir

مراحل عمومی آزمون فرضیه های آماری



گام اول

تعریف **H0** و **H1** (تعیین ادعا و نقیض ادعا)

گام دوم

تعیین مقدار آماره (اعداد محاسبه شده از فرمول ها)

گام سوم

تعیین سطح بحرانی (بر اساس عدد جدول)

گام چهارم

تصمیم گیری (با مقایسه عدد جدول و عدد محاسبه شده)

- **مقياس اسمي :**
  - براي بررسي اختلاف مقادير: آزمون خي دو -Chi square (X^2)
- **مقياس رتبه اي :**
  - بررسي ارتباط :
    - ضريب همبستگي اسپيرمن
  - بررسي اختلاف مقادير رتبه ها (مستقل):
    - دوگروه : يو من - ويتني Mann-Whitney U
    - سه گروه و بيشتر : كراسكال واليس Kruskal-Wallis
  - بررسي اختلاف مقادير رتبه ها (وابسته):
    - دو گروه متغير: ويلكاكسون Wilcoxon
    - چند گروه متغير: فريدمن Freidman test

- **مقياس نسبي فاصله اي:**
  - ارتباط متغيرها:
    - دو گروه: ضريب همبستگي پيرسون Pearson`s
    - چند گروه : رگرسيون چند متغيره Multiple regression
  - اختلاف ميانگين (مستقل):
    - دوگروه: T-studentمستقل
    - چند گروه: ANOVA
  - اختلاف ميانگين (وابسته):
    - T-student وابسته يا زوج

آزمون فرضیه آماری میانگین یک جامعه

استفاده از این آزمون زمانی است که فرضیه محقق در مورد میانگین یک جامعه باشد.

به عنوان مثال قصد داشته باشد میانگین قد دانشجویان را در کل جامعه آماری مورد مطالعه قرار دهد

گام اول

حالت اول	حالت دوم	حالت سوم
$\begin{cases} H_0 : \mu_x = \mu_0 \\ H_1 : \mu_x \neq \mu_0 \end{cases}$	$\begin{cases} H_0 : \mu_x \geq \mu_0 \\ H_1 : \mu_x < \mu_0 \end{cases}$	$\begin{cases} H_0 : \mu_x \leq \mu_0 \\ H_1 : \mu_x > \mu_0 \end{cases}$

گام دوم

جامعه نرمال و انحراف معیار آن نامعلوم

n <= 30

$$t = \frac{\overline{x} - \mu_0}{S_{\overline{x}}}$$

n >30

$$z = \frac{\overline{x} - \mu_0}{S_{\overline{x}}}$$

جامعه نرمال و انحراف معیار آن معلوم

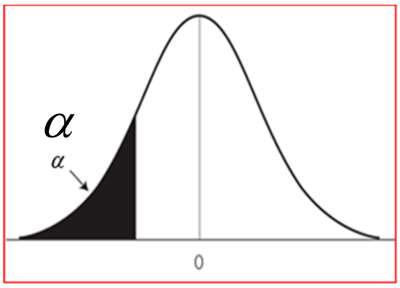
$$z = \frac{\overline{x} - \mu_0}{\sigma_{\overline{x}}}$$

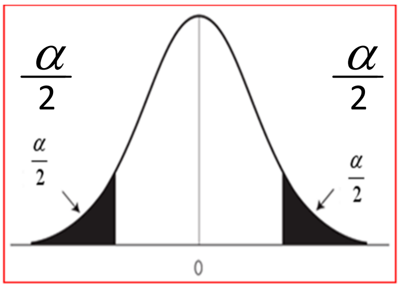
$$S_{\overline{x}} = \frac{S_x}{\sqrt{n}}$$

www.freebay.ir

گام سوم

حالت سوم	حالت دوم	حالت اول
$\begin{cases} H_0 : \mu_x \le \mu_0 \\ H_1 : \mu_x > \mu_0 \end{cases}$	$\begin{cases} H_0 : \mu_x \ge \mu_0 \\ H_1 : \mu_x < \mu_0 \end{cases}$	$\begin{cases} H_0 : \mu_x = \mu_0 \\ H_1 : \mu_x \ne \mu_0 \end{cases}$





گام چهارم

تصمیم گیری : اگر چنانچه ملاک آزمون در ناحیه بحرانی قرار گیرد فرضیه **H0** رد میشود.

سوال ۱

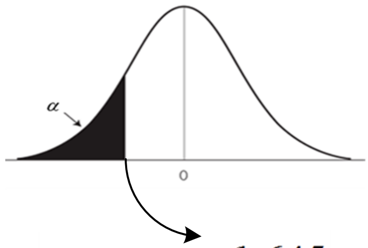
پژوهشگری فرضیه خود را بدین شکل صورتبندی نموده است: "میانگین امتیاز عملکرد کارکنان در سازمان دست کم ۵۰ است" وی برای آزمون فرضیه فوق یک نمونه ۶۴ تایی به شکل تصادفی انتخاب کرده که میانگین و انحراف معیار آن ۴۵ و ۱۶ می باشد. در سطح خطای ٪۵ فرضیه فوق را آزمون نمایید.

ادعا $H_0 : \mu_x \ge 50$

نقیض ادعا $H_1 : \mu_x < 50$

n >30

$$z = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \qquad Z = \frac{45-50}{\frac{16}{\sqrt{64}}} = -2.5$$



$z_\alpha = z_{0.05} = -1.645$

تحلیل

چون ملاک آزمون (-2.5) در ناحیه بحرانی قرار رفته پس فرضیه صفر رد می شود و در نتیجه نقیض ادعا تایید می گردد و فرضیه پژوهشی رد می شود.

سوال ۲

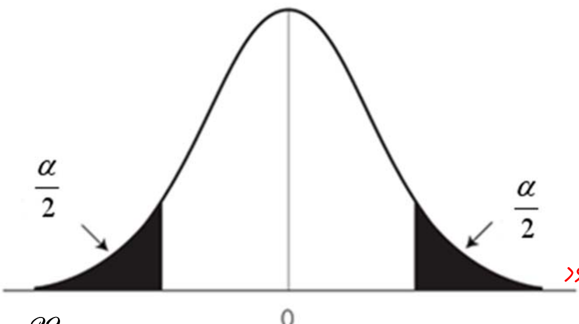
پژوهشگری فرضیه خود را بدین شکل صورتبندی نموده است: "میانگین رضایت شغلی کارکنان ۵۵ می باشد "ایشان یک نمونه تصادفی ۱۲ تایی انتخاب کرده که میانگین و انحراف معیار آن ۶۰ و ۱۵ است با فرض نرمال بودن سطح رضایت شغلی کارکنان فرضیه محقق را در سطح اطمینان ۹۹ درصد آزمون کنید

ادعا $H_0 : \mu_x = 55$

نقیض ادعا $H_1 : \mu_x \ne 55$

n <= 30

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \qquad t = \frac{60-55}{\frac{15}{\sqrt{12}}} = 1.155$$



$t_{\frac{\alpha}{2},df} = t_{0.005,11} = \pm 3.106$

تحلیل

چون ملاک آزمون(1.155) درخارج از ناحیه بحرانی قرار رفته پس پس فرضیه صفر رد نمی شود و در نتیجه ادعای محقق تایید می گردد و فرضیه پژوهشی پزیرفته می شود.

www.freebay.ir

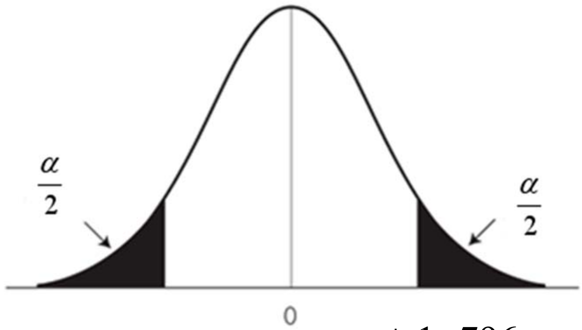
سوال ۲

پژوهشگری فرضیه خود را بدین شکل صورتبندی نموده است: "میانگین رضایت شغلی کارکنان ۵۵ می باشد "ایشان یک نمونه تصادفی ۱۲ تایی انتقاب کرده که میانگین و انحراف معیار آن ۶۰ و ۱۵ است با فرض نرمال بودن سطح رضایت شغلی کارکنان فرضیه محقق را در سطح اطمینان ۹۰ درصد آزمون کنید

$$\begin{cases} H_0 : \mu_x = 55 & \text{ادعا} \\ H_1 : \mu_x \neq 55 & \text{نقیض ادعا} \end{cases}$$

n <= 30

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \qquad t = \frac{60-55}{\frac{15}{\sqrt{12}}} = 1.155$$



$$t_{\frac{\alpha}{2}, df} = t_{0.05, 11} = \pm 1.796$$

تحلیل

چون ملاک آزمون(1.155) درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه ادعای محقق تایید می گردد و فرضیه پژوهشی پزیرفته می شود.

آزمون فرضیه آماری میانگین دو جامعه

استفاده از این آزمون زمانی است که فرضیه محقق در مورد میانگین دو جامعه باشد. یا به عبارت دیگر بفواهد میانگین شاخص مورد بررسی را در دو سازمان مورد مقایسه قرار دهد.

دو جامعه مستقل با انحراف معیارهای جامعه معلوم

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

دو جامعه مستقل با انحراف معیارهای جامعه نامعلوم

df > 30

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

df<= 30

$\sigma_1 = \sigma_2$

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

$\sigma_1 \neq \sigma_2$

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

$$df = n_1 + n_2 - 2$$

$$S_p = \sqrt{\frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}}$$

www.freebay.ir

سوال ۱

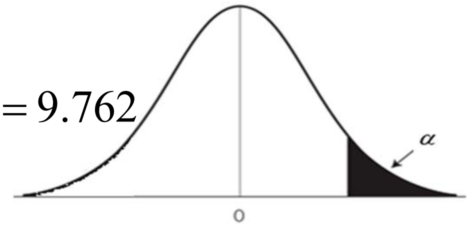
تمرین دو جامعه مستقل با انحراف معیارهای جامعه نامعلوم و مساوی

محققی به منظور مقایسه عملکرد مدیران دو سازمان که به روش های مشارکتی و دستوری اداره می شود فرضیه زیر را صورتبندی نموده است: "بهره وری سبک مدیریت مشارکتی بهتر از سبک مدیریت دستوری است" او بدین منظور از هر سازمان چند نمونه تصادفی انتقاب کرد که اطلاعات آنها بشرح زیر می باشد. با فرض تساوی واریانس ها و نرمال بودن دو جامعه ، فرضیه داده شده را در سطح اطمینان ۹۹ درصد آزمون نمایید.

سبک دستوری	سبک مشارکتی
$n_2=15$	$n_1=10$
$\bar{x}_2=45$	$\bar{x}_1=52$
$s_2=8$	$s_1=12$

df=10+15-2=23<30

$$s_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}=9.762$$



$$t_{\alpha,df}=t_{0.01,23}=2.5$$

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{(52-45)-(0)}{9.762\sqrt{\frac{1}{10}+\frac{1}{15}}}=1.757$$

نقیض ادعا $H_0: \mu_1 \le \mu_2$

ادعا $H_1: \mu_1 > \mu_2$

تحلیل

چون ملاک آزمون(1.157) درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه نقیض ادعای محقق تایید می گردد و فرضیه پژوهشی رد می شود.

سوال ۱

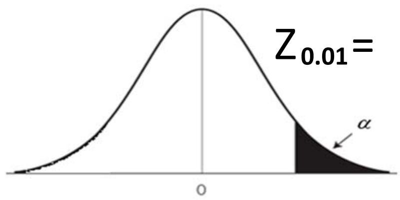
تمرین دو جامعه مستقل با انحراف معیارهای جامعه معلوم

محققی به منظور مقایسه عملکرد مدیران دو سازمان که به روش های مشارکتی و دستوری اداره می شود فرضیه زیر را صورتبندی نموده است: "بهره وری سبک مدیریت مشارکتی بهتر از سبک مدیریت دستوری است" او بدین منظور از هر سازمان چند نمونه تصادفی انتقاب کرد که اطلاعات آنها بشرح زیر می باشد. با فرض انحراف معیار جامعه اول ۸ و دوم ۱۲ ،فرضیه داده شده را در سطح اطمینان ۹۹ درصد آزمون نمایید.

سبک دستوری	سبک مشارکتی
$n_2=15$	$n_1=10$
$\bar{x}_2=45$	$\bar{x}_1=52$
$\sigma_2=12$	$\sigma_1=8$

نقیض ادعا $H_0: \mu_1 \le \mu_2$

ادعا $H_1: \mu_1 > \mu_2$



$Z_{0.01}=$ 2.33

$$z=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(52-45)-(0)}{\sqrt{\frac{64}{10}+\frac{144}{15}}}=\frac{7}{\sqrt{16}}=\frac{7}{4}=1.75$$

تحلیل

چون ملاک آزمون(1.75) درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه نقیض ادعای محقق تایید می گردد و فرضیه پژوهشی رد می شود.

سوال ۱

تمرین دو جامعه مستقل با انحراف معیارهای جامعه نامعلوم و مساوی

محققی به منظور مقایسه عملکرد مدیران دو سازمان که به روش های مشارکتی و دستوری اداره می شود فرضیه زیر را صورتبندی نموده است: "بهره وری سبک مدیریت مشارکتی بهتر از سبک مدیریت دستوری است" او بدین منظور از هر سازمان چند نمونه تصادفی انتقاب کرد که اطلاعات آنها بشرح زیر می باشد. با فرض تساوی واریانس ها و نرمال بودن دو جامعه ، فرضیه داده شده را در سطح اطمینان ۹۹ درصد آزمون نمایید.

سبک دستوری	سبک مشارکتی
$n_2=25$	$n_1=10$
$\bar{x}_2=45$	$\bar{x}_1=52$
$s_2=8$	$s_1=12$

df=10+25-2=33>30

$Z_{0.01}=$ 2.33

$$z=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}=\frac{7}{\sqrt{16}}=\frac{7}{4.11}=1.70$$

تحلیل

چون ملاک آزمون(1.70) درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه نقیض ادعای محقق تایید می گردد و فرضیه پژوهشی رد می شود.

www.freebay.ir

آزمون مقایسه زوجها

Paired Sample Test

نمونه های دو جامعه که به هم وابسته اند
یک جامعه با دو نمونه گیری

μ_2 μ_1

یک جامعه با دو نمونه گیری

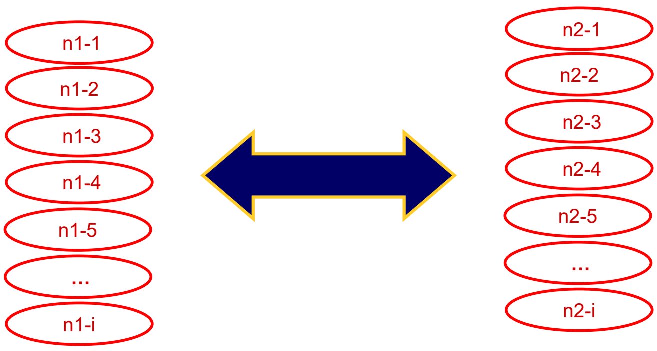
یعنی دو سری نمونه از یک جامعه به عنوان پیش آزمون (Pre Test) و پس آزمون (Post Test)

در این آزمون باید مشخص گردد میانگین دو نمونه ، معنی دار هستند یا خیر. جوامع وابسته هستند یا مستقل.

استفاده از این آزمون زمانی است که محقق بخواهد اعضای دو مجموعه متناظر را که دارای تعداد نمونه یکسان هستند مورد مقایسه قرار دهد.

عنوان مثال قصد مقایسه یک مجموعه ای از شاخص ها را در قبل و بعد از یک اقدام مورد مقایسه قرار دهد.

مراحل عمومی این آزمون همانند آزمونهای قبلی می باشد با این تفاوت که ملاک آزمون از فرمول زیر بدست می آید.



آماره آزمون در جوامع وابسته یا به اصطلاح مقایسه ای زوجی همیشه بر اساس t میباشد.

$$\bar{d} = \frac{\Sigma d_i}{N}$$

1

2

$$s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}}$$

d=difrential H0

$$\mu_d = \mu_2 - \mu_1 \quad \mu_d > 0$$

$$d_i = \mu_2 - \mu_1$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

3

4

$$t = \frac{\bar{d}}{s_{\bar{d}}}$$

www.freebay.ir

سوال ۱

نمرات پیش آزمون و پس آزمون پنج فراگیر در یک دوره آموزشی بدین شکل می باشد:

آیا فرضیه زیر در سطح اطمینان ۹۹درصد مورد تایید است؟

	شماره نمونه	۱	۲	۳	۴	۵
μ_1	پیش آزمون	۱۲	۸	۴	۱۶	۱۰
μ_2	پس آزمون	۱۴	۱۳	۱۰	۱۸	۱۵

نقیض ادعا $H_0: \mu_d \leq 0$

ادعا $H_1: \mu_d > 0$

$$t_{\alpha,df} = t_{0.01,4} = 3.747$$

t=4.81

آماره آزمون

	۱	۲	۳	۴	۵
d_i	۲	۵	۶	۲	۵
$(\bar{d}-d_i)^2$	۴	۱	۴	۴	۱

$$\bar{d} = \frac{\Sigma d_i}{N} \qquad \Sigma d_i = 20 \longrightarrow \bar{d} = \frac{\Sigma d_i}{N} = \frac{20}{5} = 4$$

$$s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{14}{5-1}} = 1.87$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{1.87}{\sqrt{5}} = 0.83$$

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{4}{0.83} = 4.81$$

تحلیل

چون ملاک آزمون در ناحیه بحرانی قرار رفته پس فرضیه صفر رد می شود و در نتیجه ادعای محقق تایید می گردد و فرضیه پژوهشی پذیرفته می شود.

دوره آموزشی مورد نظر اثربخش می باشد

www.freebay.ir

سوال ۲

نمرات وضع موجود و وضع مطلوب شایستگی های مدیریتی پنج مدیر به شکل جدول زیر می باشد:

آیا فرضیه زیر در سطح اطمینان ۹۹درصد مورد تایید است؟

	شماره نمونه	۱	۲	۳	۴	۵
μ_1	وضع موجود	۵۰	۵۹	۵۰	۵۸	۵۰
μ_2	وضع مطلوب	۴۰	۵۷	۴۷	۵۰	۴۸

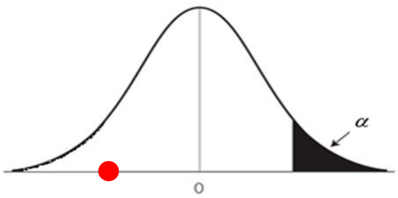
	۱	۲	۳	۴	۵
d_i	-۱۰	-۲	-۳	-۸	-۲
$(\bar{d}-d_i)^2$	۲۵	۹	۴	۹	۹

$$\bar{d} = \frac{\Sigma d_i}{N} \qquad \Sigma d_i = -25 \longrightarrow \bar{d} = \frac{\Sigma d_i}{N} = \frac{-25}{5} = -5$$

$$s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{56}{5-1}} = 3.742$$

نقیض ادعا $H_0: \mu_d \leq 0$

ادعا $H_1: \mu_d > 0$



$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{3.742}{\sqrt{5}} = 1.674$$

$$t_{\alpha,df} = t_{0.01,4} = 3.747$$

t=-2.988

آماره آزمون

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{-5}{1.674} = -2.988$$

تحلیل

چون ملاک آزمون درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه نقیض ادعای محقق تایید می گردد و فرضیه پژوهشی رد می شود.

وضعیت مطلوب شایستگی مدیران مطلوب نمیباشد.

آزمون فرضیه آماری نسبت موفقیت در جامعه

استفاده از این آزمون زمانی است که فرضیه محقق در مورد نسبت موفقیت یک جامعه باشد. و مقیاس های بکارگرفته شده از نوع مقیاس های کیفی باشند.

گام اول تعیین **H0 , H1**

حالت اول

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

حالت دوم

$$\begin{cases} H_0 : p \geq p_0 \\ H_1 : p < p_0 \end{cases}$$

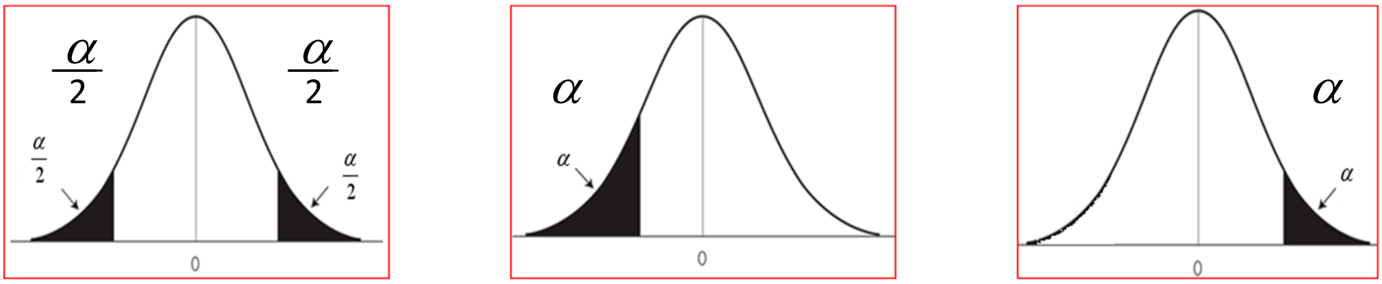
حالت سوم

$$\begin{cases} H_0 : p \leq p_0 \\ H_1 : p > p_0 \end{cases}$$

گام دوم آماره آزمون

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

گام سوم تعیین ناحیه بحرانی



گام چهارم تصمیم گیری: اگر چنانچه ملاک آزمون در ناحیه بحرانی قرار بگیرد فرضیه **H0** رد می شود.

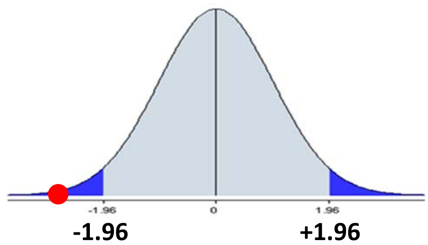
سوال ۱

فرضیه ای به این صورت بیان شده است: "شصت درصد مدیران سازمان دارای سبک مدیریت دستوری هستند "محقق ۲۰۰ نفر مدیر را بصورت تصادفی انتخاب کرده که بطور متوسط نیمی از آنها دارای سبک دستوری هستند. فرضیه فوق را در سطح اطمینان ۹۵درصد آزمون نمایید.

$$\begin{cases} H_0 : p = 0.6 & \text{ادعا} \\ H_1 : p \neq 0.6 & \text{نقیض ادعا} \end{cases}$$

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.5 - 0.6}{\sqrt{\frac{0.6(1-0.6)}{200}}} = -2.89$$

$$z_{\frac{\alpha}{2}} = z_{0.025} = \pm 1.96$$



از آنجاییکه ملاک آزمون در دامنه چپ قرار گرفته لذا می توان گفت که درصد مدیران دستوری کمتر از ۶۰ درصد می باشد.

تحلیل

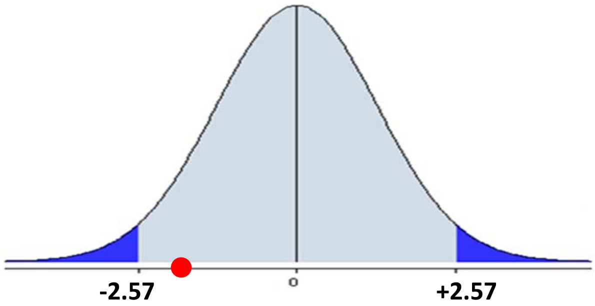
از آنجاییکه ملاک آزمون در ناحیه بحرانی قرار گرفته لذا فرضیه صفر رد می شود و در نتیجه ادعا (فرضیه پژوهشی) رد می شود.

Bay
www.freebay.ir

سوال ۲

فرضیه ای به این صورت بیان شده است: "۶۵ درصد از دانشجویان دانشگاه آزاد اسلامی را خانم ها تشکیل می دهند" محقق برای آزمون این فرضیه ۱۵۰ نفر را بصورت تصادفی انتخاب کرده که بطور متوسط ۶۰ درصد انها خانم هستند. فرضیه مورد نظر را در سطح اطمینان ۹۹ درصد آزمون نمایید.

$$\begin{cases} H_0 : p = 0.65 & \text{ادعا} \\ H_1 : p \neq 0.65 & \text{نقیض ادعا} \end{cases} \qquad \begin{cases} p_0 = 0.65 \\ \bar{p} = 0.6 = \dfrac{60}{100} \end{cases}$$



$$z = \frac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.6 - 0.65}{\sqrt{\dfrac{0.65(1-0.65)}{150}}} = -1.28$$

$$z_{\frac{\alpha}{2}} = \pm 2.57$$
$$Z_{0.005}$$

تحلیل

از آنجاییکه ملاک آزمون در خارج از ناحیه بحرانی قرار گرفته لذا فرضیه صفر رد نمی شود و در نتیجه ادعا (فرضیه پژوهشی) پذیرفته می شود.

سوال ۳

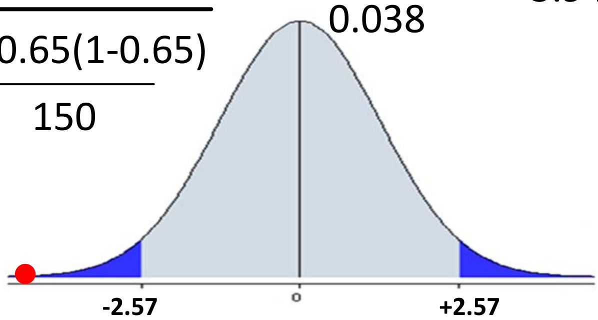
فرضیه ای به این صورت بیان شده است: "۶۵ درصد از دانشجویان دانشگاه آزاد اسلامی را خانم ها تشکیل می دهند" محقق برای آزمون این فرضیه ۱۵۰ نفر را بصورت تصادفی انتخاب کرده که ۷۵ نفر از آنها خانم هستند. فرضیه مورد نظر را در سطح اطمینان ۹۹ درصد آزمون نمایید.

$$\begin{cases} H_0 : p = 0.65 & \text{ادعا} \\ H_1 : p \neq 0.65 & \text{نقیض ادعا} \end{cases}$$

$$z = \frac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.5-0.65}{\sqrt{\dfrac{0.65(1-0.65)}{150}}} = \frac{-0.15}{0.038} = -3.94$$

$$z_{\frac{\alpha}{2}} = \pm 2.57$$
$$Z_{0.005}$$

$$\begin{cases} p_0 = 0.65 \\ \bar{p} = 0.5 = \dfrac{75}{150} \end{cases}$$



تحلیل

از آنجاییکه ملاک آزمون در ناحیه بحرانی قرار گرفته لذا فرضیه صفر رد می شود و در نتیجه ادعا (فرضیه پژوهشی) پذیرفته نمی شود.

سوال ۴

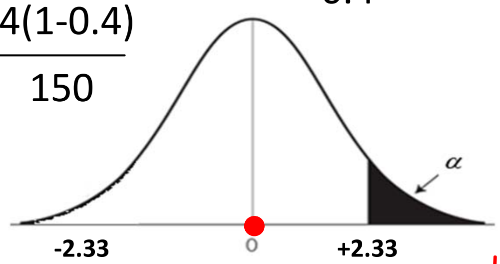
فرضیه ای به این صورت بیان شده است: "حداکثر ۴۰ درصد دانشجویان را پسران تشکلیل می دهند" محقق برای آزمون این فرضیه ۱۵۰ نفر را بصورت تصادفی انتخاب کرده که بطور متوسط ۶۰ درصد انها خانم هستند. فرضیه مورد نظر را در سطح اطمینان ۹۹ درصد آزمون نمایید.

$$\begin{cases} H_0 : p \leqslant 0.4 & \text{ادعا} \\ H_1 : p > 0.4 & \text{نقیض ادعا} \end{cases}$$

$$z = \frac{\bar{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.4-0.4}{\sqrt{\dfrac{0.4(1-0.4)}{150}}} = \frac{0}{0.4} = 0$$

$$Z_{\alpha} = Z_{0.01} = 2.33 \qquad \begin{cases} p_0 = 0.4 = \dfrac{40}{100} \\ \bar{p} = 0.4 \end{cases}$$

60% خانم = 40% آقا



تحلیل

چون ملاک آزمون درخارج از ناحیه بحرانی قرار رفته پس فرضیه صفر رد نمی شود و در نتیجه نقیض ادعای محقق تایید می گردد و فرضیه پژوهشی رد می شود.

www.freebay.ir

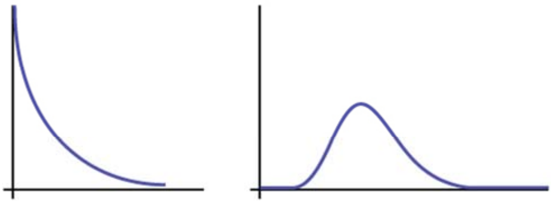
جدول سطح زیر منحنی توزیع کای دو – کای مربع

Table :Chi-Square Probabilities

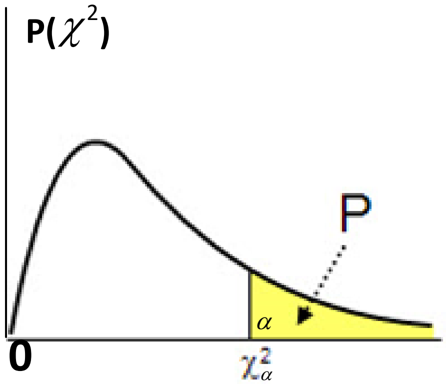
-No negative values

- Mean is equal to the degrees of freedom
- The standard deviation increases as degrees of freedom increase, so the chi-square curve spreads out more as the degrees of freedom increase.

- As the degrees of freedom become very large, the shape becomes more like the normal distribution.



df = 2 (a) df = 24 (b)



df	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

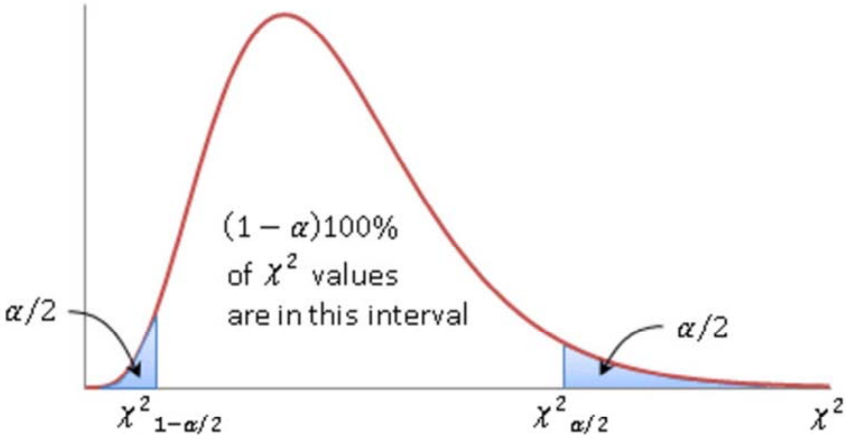
www.freebay.ir

آزمون فرضیه آماری برای واریانس جامعه

هرگاه فرضیه ای در مورد پراکندگی جامعه آماری باشد صحت یا سقم آن را می توان با استفاده از این آزمون مورد بررسی قرار داد.

گام اول تعیین **H0 , H1**

حالت سوم	حالت دوم	حالت اول
$\begin{cases} H_0 : \sigma_x^2 \le \sigma_0^2 \\ H_1 : \sigma_x^2 > \sigma_0^2 \end{cases}$	$\begin{cases} H_0 : \sigma_x^2 \ge \sigma_0^2 \\ H_1 : \sigma_x^2 < \sigma_0^2 \end{cases}$	$\begin{cases} H_0 : \sigma_x^2 = \sigma_0^2 \\ H_1 : \sigma_x^2 \ne \sigma_0^2 \end{cases}$

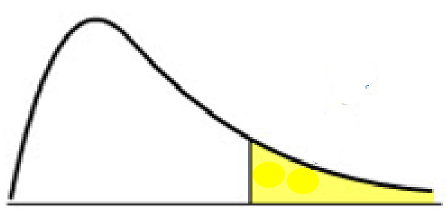
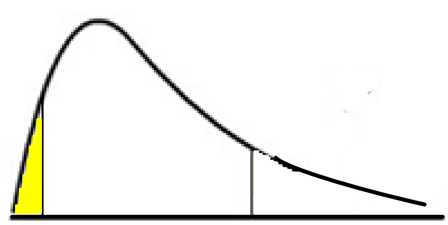
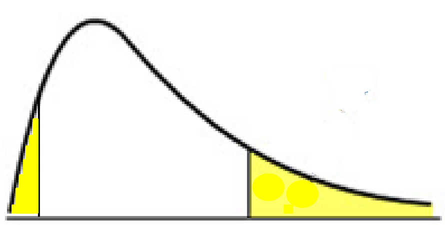


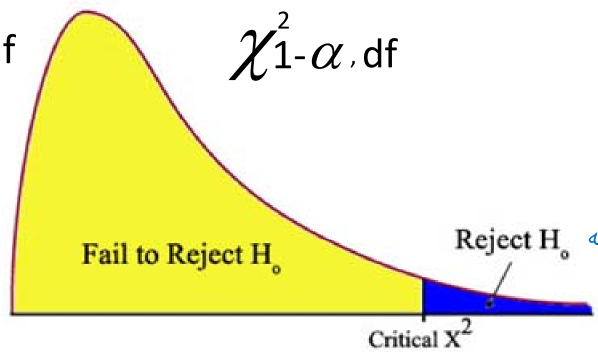
گام دوم آماره آزمون یا ملاک آزمون

$$\chi^2 = \frac{(n-1)*S_x^2}{\sigma_0^2}$$

$$df=(n-1)$$

گام سوم تعیین ناحیه بحرانی

حالت سوم	حالت دوم	حالت اول
		
$\chi^2_{\alpha,df}$	$\chi^2_{1-\alpha,df}$	$\chi^2_{1-\frac{\alpha}{2},df}$, $\chi^2_{\frac{\alpha}{2},df}$



گام چهارم

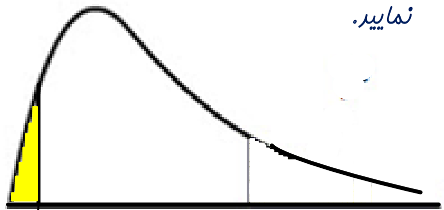
تصمیم گیری:اگر مقدار کای دو در ناحیه بحرانی قرار بگیرد فرضیه صفر رد می شود.

سوال ۱

مدیرعامل بازار بورس ادعا کرده است که ریسک(انحراف معیار) بازده سهام شرکتهای عرضه کننده سهام در بازار بورس کمتر از ۵ تومان است. بدین منظور یکی از کارگزاران ۲۵ شرکت را بطور تصادفی انتخاب از بین شرکتهای عرضه کننده انتخاب کرده که میانگین بازده آنها ۱۴ و انحراف معیارشان ۴ تومان است. اگر بازده شرکتهای بازار بورس از توزیع نرمال برخوردار باشد ادعا را در سطح اطمینان ۹۵ درصد آزمون نمایید.

$$\begin{cases} H_0 : \sigma_x^2 \geqslant 25 & \text{نقیض ادعا} \\ H_1 : \sigma_x^2 < 25 & \text{ادعا} \end{cases}$$

$$\chi^2 = \frac{(n-1)*S_x^2}{\sigma_0^2} = \frac{(25-1)*4^2}{5^2} = 15.36$$



$$\chi^2_{1-\alpha,df} = \chi^2_{0.95,24} = 13.85$$

تحلیل

چون ملاک آزمون در خارج از ناحیه بحرانی قرار گرفته لذا فرضیه صفر رد نمی شود و ادعای مدیرعامل پذیرفته نیست.

www.freebay.ir

آزمون کای مربع χ^2

یکی از مهمترین آزمونهای ناپارامتریک آزمون کای دو است. اساس و پایه این آزمون بررسی فراوانی مشاهده شده که در طرح های تحقیقاتی جمع شده اند با فراوانی های مورد انتظار است. یعنی می خواهیم برانیم آیا بین فراوانی های مشاهده شده و فراوانی های مورد انتظار تفاوتی معنی دار وجود دارد یا آنکه این تفاوت ناچیز و حاصل شانس است. در واقع می خواهیم برانیم که بین دو متغیر ارتباطی وجود دارد یا آن دو متغیر مستقل از هم می باشند.

توزیع ۲ را معمولا وقتی بکار می برند که داده های جمع آوری شده به صورت فراوانی بوده و فرضیه ها بصورت رابطه ای و تفاوتی باشند. داده های جمع آوری شده برای متغیر در یک جدول که شامل **r** سطر و **C** ستون است خلاصه می شود که به چنین جدولی، جدول «توافقی» گویند. در حالت کلی جدول زیر را خواهیم داشت:

توزیع کای‌دو از توزیع نرمال برست می‌آیر. اگر متغیر **X** متعلق به جامعه نرمال با میانگین و انحراف معیار باشد و از این جامعه نمونه تصادفی به حجم **n** انتقاب شود، سپس هریک از اعضای نمونه با استفاده از فرمول $Z=\frac{X-\mu_x}{\mu_x}$ استاندارد شوند، متغیر تصادفی زیراز توزیع کای‌دو با **n** درجه آزادی پیروی می‌کند.

$$\chi^2=Z_1^2+Z_2^2+\dots+Z_n^2$$

متغیر دوم / متغیر اول	۱	۲	۲	...	C	مجموع
۱	n_{11}	n_{12}	n_{13}	...	n_{1c}	$n_{1.}$
۲	n_{21}	n_{22}	n_{23}	...	n_{2c}	$n_{2.}$
۳	n_{31}	n_{32}	n_{33}	...	n_{3c}	$n_{3.}$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
r	n_{r1}	n_{r2}	n_{r3}	...	n_{rc}	$n_{r.}$
مجموع	$n_{.1}$	$n_{.2}$	$n_{.3}$	...	$n_{.c}$	n

مقادیر **nij** فراوانیهای مشاهده شده در هر سلول است که فصل مشترک سطر **i** و ستون **j** می باشر. جمع سطر **i** ، **ni** و جمع ستون **j** ، **n.j** علامتگزاری شده است.

www.freebay.ir

مراحل آزمون استقلال کای دو

گام اول تعریف **H0 , H1**

H_0 بین **X,Y** ارتباط وجود ندارد یعنی مستقل هستند.

H_1 بین **X,Y** ارتباط وجود دارد یعنی وابسته هستند.

گام دوم آماره آزمون یا ملاک آزمون

F_e فراوانی مورد انتظار Expected

F_o فراوانی مشاهده شده Observed

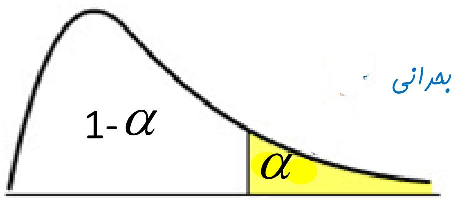
$$\chi^2=\sum\frac{(F_o-F_e)^2}{F_e}$$

$$df=(r-1)(c-1)$$

$$F_e=\frac{(n_{i.}*n_{.j})}{n}$$

$$F_e=\frac{\text{مجموع سطر * مجموع ستون}}{\text{کل فراوانی}}$$

با استفاده از این آزمون می فهمیم که بین گزینه ارتباط وجود دارد سپس با **همبستگی اسپیرمن** شدت آنرا برآورد می کنیم



گام سوم تعیین ناحیه بحرانی

$$\chi^2_{\alpha,df}$$

گام چهارم

تصمیم گیری:اگر مقدار کای دو در ناحیه بحرانی قرار بگیرد فرضیه صفر رد می شود.

سوال ۱

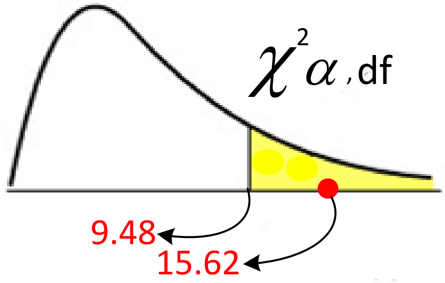
فرضیه ای بصورت زیر تدوین شده است:"بین عملکرد کارکنان و میزان رضایت شغلی آنها ارتباط وجود دارد "برای بررسی فرضیه فوق یک نمونه ۱۸۰ نفره بطور تصادفی انتخاب شده و میزان عملکرد و رضایت شغلی آنها اندازه گیری شده است که اطلاعات آن در جدول زیر ارائه شده است. در سطح خطای ۵ درصد فرضیه فوق را آزمون نمایید.

رضایت / عملکرد	بالا	متوسط	پایین
خوب	18	20	7
متوسط	15	37	38
ضعیف	7	23	15

$F_e = \frac{40*45}{180}$ فراوانی هر عملکرد

رضایت / عملکرد	بالا (F_o F_e)	متوسط	پایین	جمع
خوب	18 10	20 20	7 15	45
متوسط	15 20	37 40	38 30	90
ضعیف	7 10	23 20	15 15	45
جمع	40	80	60	F_e 180 فراوانی کل

$\frac{60*45}{180}$



$$\chi^2_{\alpha,df} = \chi^2_{0.05,4} = 9.48$$

$$\chi^2 = \Sigma\frac{(F_0 - F_e)^2}{F_e} = \frac{(18-10)^2}{10} + \frac{(20-20)^2}{20} + \frac{(7-15)^2}{15} + \dots + \frac{(15-15)^2}{15} = 15.62$$

$$df = (r-1)(c-1) = (3-1)(3-1) = 4$$

H_0 بین **X,Y** ارتباط وجود ندارد یعنی مستقل هستند.

H_1 بین **X,Y** ارتباط وجود دارد یعنی وابسته هستند.

تحلیل (تصمیم گیری)

چون ملاک آزمون در ناحیه بحرانی قرار گرفته لذا فرضیه صفر رد می شود لذا می توان گفت که بین عملکرد و رضایت ارتباط وجود دارد و مستقل از هم نیستند بلکه وابسته هستند . دراین حالت در صورت نیاز به سراغ آزمون همبستگی اسپیرمن می رویم.

رگرسیون

Regression

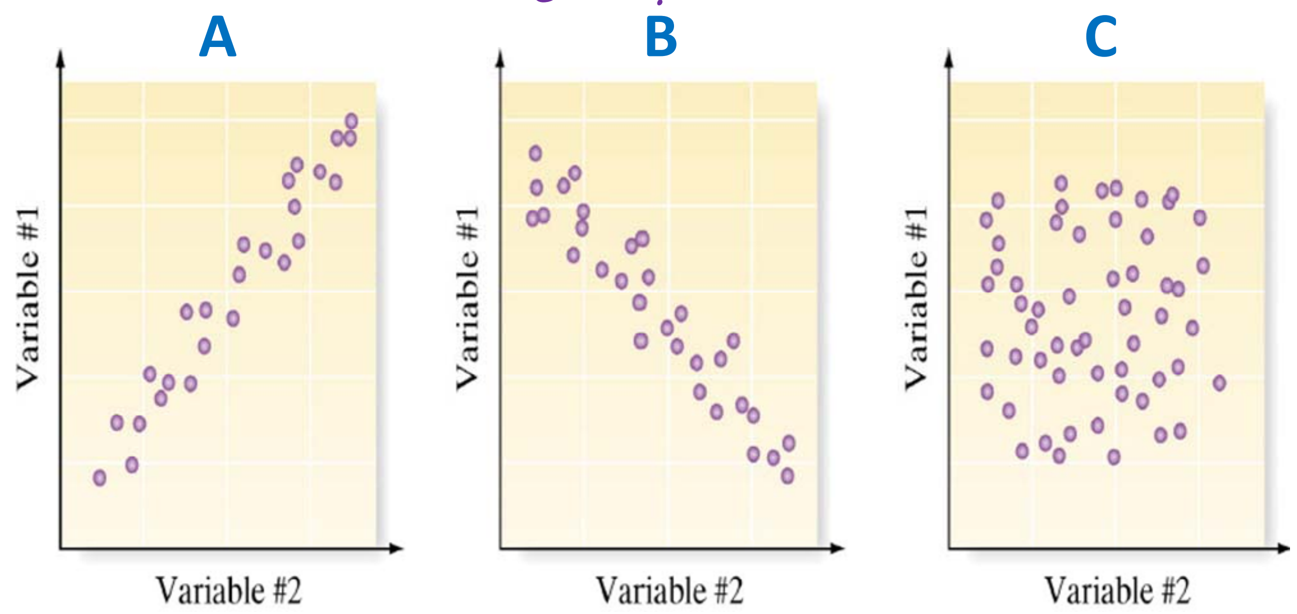
واژه رگرسیون در فرهنگ لغت به معنی بازگشت است و اغلب جهت رساندن مفهوم "بازگشت به یک مقدار متوسط یا میانگین "به کار می رود. بدین معنی که برخی پدیده ها به مرور زمان از نظر کمی به طرف یک مقدار متوسط میل می کنند . بیش از۱۰۰ سال پیش در سال ۱۸۷۷ فرانسیس گالتون(Francis Galton) در مقاله ای که در همین زمینه منتشر کرد اظهار داشت که متوسط قد پسران دارای پدران قد بلند ، کمتر از قد پدرانشان می باشد . به نحو مشابه متوسط قد پسران دارای پدران کوتاه قد نیز بیشتر از قد پدرانشان گزارش شده است. به این ترتیب گالتون پدیده بازگشت به طرف میانگین را در داده هایش مورد تأکید قرار داد . برای گالتون رگرسیون مفهومی زیست شناختی داشت اما کارهای او توسط کارل پیرسون (Karl Pearson) برای مفاهیم آماری توسعه داده شده . گرچه گالتون برای تأکید بر پدیده "بازگشت به سمت مقدار متوسط" از تحلیل رگرسیون استفاده کرد، اما به هر حال امروزه واژه تحلیل رگرسیون جهت اشاره به مطالعات مربوط به روابط بین متغیرها به کار برده می شود .

نمودار پراکندگی

در حقیقت تحلیل رگرسیونی فن و تکنیکی آماری برای بررسی و مدل سازی ارتباط بین متغیرها است. رگرسیون تقریباً در هر زمینه ای از جمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی، بیولوژی و علوم اجتماعی برای برآورد و پیشبینی مورد نیاز است . می توان گفت تحلیل رگرسیونی، پرکاربردترین روش در بین تکنیک های آماری است. شمایی کلی و خلاصه شده از یک تحلیل رگرسیونی ساده به صورت زیر می باشد: در ابتدا تحلیل گر حدس می زند که بین دو متغیر نوعی ارتباط وجود دارد ، در حقیقت حدس می زند که یک رابطه به شکل یک خط بین دو متغیر وجود دارد و سپس به جمع آوری اطلاعات کمی از دو متغیر می پردازد و این داده ها را به صورت نقاطی در یک نمودار دو بعدی رسم می کند.

www.freebay.ir

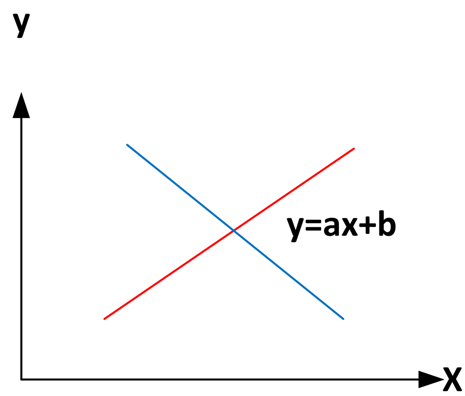
نمودار پراکندگی



A.Positive Relationship **B.Negative Relationship** **C.No Relationship**

این نمودار که به آن نمودار پراکندگی **(Scatter Plot)** گفته می شود نقش بسیار مهمی را در تحلیل های رگرسیونی و نمایش ارتباط بین متغیرها ایفا می کند. در صورتی که نمودار نشان دهنده این باشد که داده ها تقریباً **(** نه لزوماً دقیق **)** در امتداد یک خط مستقیم پراکنده شده اند، حدس تحلیل گر تأیید شده و این ارتباط خطی به صورت زیر نمایش داده می شود:

y=ax+b که در آن **a** عرض از مبدأ و **b** شیب این خط است.



همبستگی پیرسون

$$r=\frac{\overline{xy}-\bar{x}\bar{y}}{\sqrt{S_x^2\,S_y^2}}$$

$$S_x^2=\frac{\sum x^2}{N}-\mu_x^2$$

$$S_y^2=\frac{\sum y^2}{N}-\mu_y^2$$

$$a=\frac{\overline{xy}-\bar{x}\bar{y}}{S_x^2}$$

$$b=\bar{y}-a\bar{x}$$

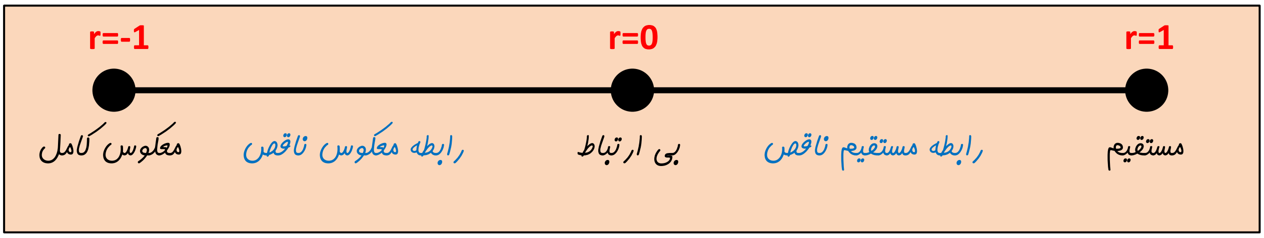
$$\bar{y}=|a\bar{x}|+b$$

معادله خط رگرسیون

- با استفاده از جدول r در سطح 0.05 و 0.01
  - Df = n – 2
  - اگر قدر مطلق r بیشتر از عدد جدول باشد
- ضریب همبستگی بدست آمده معنی دار است

بسیار پایین	0 – 0.19
پایین	0.2 – 0.39
متوسط	0.4 – 0.69
بالا	0.7 – 0.89
بسیار بالا	0.9 - 1

عدد **r** باید حتما بین بازه ۱- تا ۱ باشد . درغیر اینصورت کاربردی نمیباشد . لازم بذکر است که در هر دو ضریب همبستگی پیرسون و اسپیرمن تحلیل **r** به یک شکل میباشد.



www.freebay.ir

سوال ۱

داده های زیر هزینه تبلیغات شرکتی را همراه با تعداد فروش محصولش در طی ۹ سال مختلف نشان می دهد. اگر چنانچه هزینه تبلیغات ۱۰ واحد باشد مقدار فروش را پیش بینی نمایید.و شرت همبستگی بین فروش و تبلیغات را تحلیل نمایید.

تعداد فروش Y	11	20	16	24	26	15	21	18	27
هزینه تبلیغات X	3	5	4	7	9	6	5	4	8

$\bar{y} = \frac{178}{9} = 19.77$

$\bar{x} = \frac{50}{9} = 5.55$

x	y	x^2	y^2	xy
3	11	9	121	33
5	20	25	400	100
4	16	16	256	64
7	24	49	576	168
9	26	81	676	234
6	15	36	225	90
5	21	25	441	105
4	18	16	324	72
8	27	64	729	216
5.55	19.77	35.66	416.77	120.22
$\bar{x}$	$\bar{y}$	$\overline{x^2}$	$\overline{y^2}$	$\overline{xy}$
μ_x	μ_y	$\frac{\sum x^2}{N}$	$\frac{\sum y^2}{N}$	

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{S_x^2 S_y^2}} \qquad r = \frac{120.22 - 5.55*19.77 = 10.49}{\sqrt{4.85*25.91 = 125.66}} = \frac{10.49}{11.20} = 0.93$$

$$S_x^2 = \frac{\sum x^2}{N} - \mu_x^2 = 35.66 - 5.55*5.55 = 4.85$$

$$S_y^2 = \frac{\sum y^2}{N} - \mu_y^2 = 416.77 - 19.77*19.77 = 25.91$$

$$a = \frac{\overline{xy} - \bar{x}\bar{y}}{S_x^2} = \frac{10.49}{4.58} = 2.29$$

$$b = \bar{y} - a\bar{x} = 19.77 - 2.29*5.55 = 7.06$$

$$y = |ax| + b$$

$$Y = 2.29X + 7.06$$

$$Y = 2.29*10 + 7.06 = 29.96$$

هزینه تبلیغات ۱۰ واحد

تحلیل

بین دو متغیر رابطه همبستگی مستقیم و ناقص وجود دارد

تقریبا هیچ موقع به عدد **r=1** نمیرسیم و به سمت آن در رابطه نزدیک میشویم.

نکته

$r^2 = 0.93*0.93 = 0.86$ → یعنی **86%** موضوع مربوط به تبلیغات میباشد و **14%** دیگر مربوط به شاخص هایی است که در اینجا دیده نشده اند.

ضریب همبستگی رتبه ای اسپیرمن

در زمانی که بخواهیم رابطه دو متغیر ترتیبی را بررسی نماییم از ضریب همبستگی اسپیرمن استفاده می کنیم.

برای پیدا کردن ضریب همبستگی اسپیرمن از فرمول زیر استفاده می کنیم؛

ضریب همبستگی اسپیر من معادل نا پارامتریک ضریب همبستگی پیرسون میباشد.

$$r = 1 - \frac{6 * \sum d_i^2}{n(n^2 - 1)}$$

$$d_i = x - y$$

قابل ذکر است که در محاسبه ضریب همبستگی اسپیرمن ابتدا باید به این مسئله دقت کنیم که متغیرهای ما به صورت ترتیبی چیده شده است یا نه؟

اگر بصورت ترتیبی (صعودی یا نزولی) چیده شده باشند براحتی با استفاده از فرمول گفته شده همبستگی اسپیرمن را محاسبه می کنیم اما اگر به ترتیب چیده نشده باشند ابتدا به ترتیب داده ها را می چینیم و بعد با استفاده از فرمول, همبستگی اسپیرمن رامحاسبه می نماییم.

سوال ۱

در متغیرهای زیر ضریب همبستگی اسپیرمن را محاسبه نمایید؟

رتبه نزولی

X	2	4	5	2
Y	2	4	1	3
R_x	3	2	1	3
R_y	3	1	4	2
d_i	0	1	-3	1
d_i^2	0	1	9	1

$$r_s = 1 - \frac{6 * \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6*13}{4(16-1)} = -0.1$$

رتبه صعودی

X	2	4	5	2
Y	2	4	1	3
R_x	1	2	3	1
R_y	2	4	1	3
d_i	-1	-2	2	-2
d_i^2	1	4	4	4

$$r_s = 1 - \frac{6 * \sum d_i^2}{n(n^2 - 1)} = 1 - \frac{6*9}{4(16-1)} = 0.1$$

www.freebay.ir

سوال ۲

در متغییرهای زیر ضریب همبستگی اسپیرمن را محاسبه نمایید؟

	X	15	16	17	18	19
	Y	80	69	41	78	53
Rate	R_x	1	2	3	4	5
Rate	R_y	5	3	1	4	2
	d_i	-4	-1	2	0	3
	d_i^2	16	1	4	0	9

رتبه نزولی

$$r_s = 1 - \frac{6 * \sum d_i^2}{n(n^2-1)} = 1 - \frac{6*30}{5(25-1)} = 1 - \frac{180}{120} = 1 - \frac{3}{2} = -0.5$$

با توجه به اینکه متغیر **X** بصورت صعودی مرتب شده بصورت صعودی از ۱ تا ۵ رتبه دهی میکنیم . از همین رو متناظر آن متغیر **Y** بر اساس مقدار رتبه دهی میگردد.

X	15	16	17	18	19
Y	80	69	41	78	53
R_x	1	2	3	4	5
R_y	5	3	1	4	2
d_i	-4	-1	2	0	3
d_i^2	16	1	4	0	9

رتبه صعودی

$$r_s = 1 - \frac{6 * \sum d_i^2}{n(n^2-1)} = 1 - \frac{6*30}{5(25-1)} = 1 - \frac{180}{120} = 1 - \frac{3}{2} = -0.5$$

$r^2 = 0.25$ یعنی ۲۵ درصد متغیر **Y** به متغیر **X** وابسته است و ۷۵ درصد الباقی به عوامل دیگری وابسته است.

با توجه به دو مثال قبلی و مقایسه جواب ها بهتر است در رتبه بندی از روش نزولی استفاده شود.

تمرین

پاسخ ۳۰ نفر از کارکنان یک سازمان در مورد دو متغیر رضایت شغلی و تعهد سازمانی به شرح جدول زیر می باشد که این سوالات در طیف لیکرت ارزیابی شده و به شرح زیر ارزشگزاری شده است . آیا بین ایندو متغیر ارتباط وجود دارد؟ در صورت وجود ارتباط شدت آنرا بدست آورید.

		ارزش	۱	۲	۳	۴	۵
متغیر اول	رضایت شغلی	گزینه ها	کاملا ناراضی	ناراضی	تاحدی راضی	راضی	کاملا راضی
متغیر دوم	تعهد سازمانی	گزینه ها	خیلی ضعیف	ضعیف	متوسط	خوب	خیلی خوب

	رضایت شغلی	تعهد سازمانی
پاسخگوی ۱	1	1
پاسخگوی ۲	5	3
پاسخگوی ۳	1	1
پاسخگوی ۴	1	4
پاسخگوی ۵	2	1
پاسخگوی ۶	3	2
پاسخگوی ۷	2	3
پاسخگوی ۸	1	2
پاسخگوی ۹	1	3
پاسخگوی ۱۰	5	4
پاسخگوی ۱۱	4	1
پاسخگوی ۱۲	1	2
پاسخگوی ۱۳	2	3
پاسخگوی ۱۴	3	5
پاسخگوی ۱۵	2	4

	رضایت شغلی	تعهد سازمانی
پاسخگوی ۱۶	1	1
پاسخگوی ۱۷	3	1
پاسخگوی ۱۸	1	3
پاسخگوی ۱۹	2	2
پاسخگوی ۲۰	1	5
پاسخگوی ۲۱	3	4
پاسخگوی ۲۲	2	1
پاسخگوی ۲۳	5	5
پاسخگوی ۲۴	4	5
پاسخگوی ۲۵	1	1
پاسخگوی ۲۶	5	5
پاسخگوی ۲۷	5	4
پاسخگوی ۲۸	4	1
پاسخگوی ۲۹	4	2
پاسخگوی ۳۰	2	1

ابتدا با آزمون کای ۲ ،ارتباط بین دو متغیر بررسی میشود، سپس با استفاده از ضریب همبستگی رتبه ای اسپیرمن، شدت ارتباط را بدست میآوریم

H_0 بین رضایت شغلی و تعهد سازمانی ارتباط وجود ندارد یعنی مستقل هستند.

H_1 بین رضایت شغلی و تعهد سازمانی ارتباط وجود دارد یعنی وابسته هستند.

www.freebay.ir

رضایت شغلی \ تعهد سازمانی	1 خیلی ضعیف	2 ضعیف	3 متوسط	4 خوب	5 خیلی خوب	جمع
1 کاملا ناراضی	F_o 4 3.3 F_e	2 1.7	2 1.7	1 1.7	1 1.7	10
2 ناراضی	3 2.3	1 1.2	2 1.2	1 1.2	0 1.2	7
3 تا حدی راضی	1 1.3	1 0.7	0 0.7	1 0.7	1 0.7	4
4 راضی	2 1.3	1 0.7	0 0.7	0 0.7	1 0.7	4
5 کاملا راضی	0 1.7	0 0.8	1 0.8	2 0.8	2 0.8	5
جمع	10	5	5	5	5	30 F_e فراوانی کل

$$F_e=\frac{(n_{i.}\ *n_{.j}\ )}{n}=\frac{\text{مجموع سطر * مجموع ستون}}{\text{کل فراوانی}}=\frac{10*5}{30}=1.7$$

$$\chi^2=\sum\frac{(F_0-F_e)^2}{F_e}=\overset{0.14}{\frac{(4-3.3)^2}{3.3}}+\overset{0.21}{\frac{(3-2.3)^2}{2.3}}+\overset{0.06}{\frac{(1-1.3)^2}{1.3}}+\overset{0.37}{\frac{(2-1.3)^2}{1.3}}+\overset{1.7}{\frac{(0-1.7)^2}{1.7}}+\overset{0.05}{\frac{(2-1.7)^2}{1.7}}+\overset{0.04}{\frac{(1-1.2)^2}{1.2}}+$$

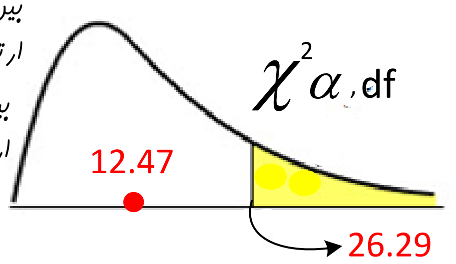
$$\overset{0.12}{\frac{(1-0.7)^2}{0.7}}+\overset{0.12}{\frac{(1-0.7)^2}{0.7}}+\overset{0.8}{\frac{(0-0.8)^2}{0.8}}+\overset{0.05}{\frac{(2-1.7)^2}{1.7}}+\overset{0.91}{\frac{(2-1.2)^2}{1.2}}+\overset{0.7}{\frac{(0-0.7)^2}{0.7}}+\overset{0.7}{\frac{(0-0.7)^2}{0.7}}+\overset{0.05}{\frac{(1-0.8)^2}{0.8}}+\overset{0.28}{\frac{(1-1.7)^2}{1.7}}+$$

$$\overset{0.03}{\frac{(1-1.2)^2}{1.2}}+\overset{0.12}{\frac{(1-0.7)^2}{0.7}}+\overset{0.7}{\frac{(0-0.7)^2}{0.7}}+\overset{1.8}{\frac{(2-0.8)^2}{0.8}}+\overset{0.28}{\frac{(1-1.7)^2}{1.7}}+\overset{1.2}{\frac{(0-1.2)^2}{1.2}}+\overset{0.12}{\frac{(1-0.7)^2}{0.7}}+\overset{0.12}{\frac{(1-0.7)^2}{0.7}}+\overset{1.8}{\frac{(2-0.8)^2}{0.8}}=12.47$$

$$\begin{cases} df=(r-1)(c-1)=4*4=16 \\ \chi^2_{\alpha,df}=\chi^2_{0.05,16}=26.29 \end{cases}$$

H_0 بین رضایت شغلی و تعهد سازمانی ارتباط وجود ندارد یعنی مستقل هستند.

H_1 بین رضایت شغلی و تعهد سازمانی ارتباط وجود دارد یعنی وابسته هستند.



تحلیل

چون ملاک آزمون در خارج از ناحیه بحرانی قرار گرفته، لذا فرضیه صفر رد نمیشود. پس میتوان گفت که بین تعهد سازمانی و رضایت شغلی ارتباط وجود ندارد و مستقل از هم هستند.

با توجه به عدم وجود وابستگی بین متغیرها نیاز به محاسبه شدت رابطه (بطور مثال ضریب همبستگی اسپیرمن نیست.)

با فرض اینکه متغیر ها وابسته باشند بعنوان تمرین فرمول اسپیرمن محاسبه شده است

چون در این مثال از ابتدا رتبه بندی شده است ، صرفا نیاز دارد تا متغیر **X** را بصورت صعودی مرتب نمائیم.

X	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3	3	4	4	4	4
Y	1	1	4	2	3	2	1	3	5	1	1	3	3	4	2	1	1	2	5	1	4	1	5	1	2
d_i	0	0	-3	-1	-2	-1	0	-2	-4	0	1	-1	-1	-2	0	1	1	1	-2	2	-1	3	-1	3	2
d_i^2	0	0	9	1	4	1	0	4	16	0	1	1	1	4	0	1	1	1	4	4	1	9	1	9	4

X	5	5	5	5	5
Y	3	4	5	5	4
d_i	2	1	0	0	1
d_i^2	4	1	0	0	1

$$r_s=1-\frac{6*\sum d_i^2}{n(n^2-1)}=1-\frac{6*83}{30(900-1)}=1-\frac{498}{26970}=0.982$$

استنتاج به این صورت است که در هر شرایطی ضریب همبستگی برست می آید . پس آزمون استقلال برای محاسبه ضرایب همبستگی مهم است.

www.freebay.ir