بسم الله الرحمن الرحیم

«کتاب آمار و احتمالات کاربردی»

تألیف: دال پل ← ترجمه ذهیبون و آرمان نژاد

در این مبحث از نرم افزار spss استفاده خواهد شد.

داده‌ها
- گسسته
  - کمی (تعداد افراد یک کلاس)
  - کیفی (گروه خون)
- پیوسته ← وزن

جدول فراوانی داده‌های پیوسته

$K = 1 + 3.3322 \log n$

واحد گردشده	داده‌ها
1	عدد صحیح
0/1	تا یک رقم اعشار
0/01	تا دو رقم اعشار
0/003	تا سه رقم اعشار

$S = \frac{\text{واحد گردشده داده‌ها}}{2}$

$\begin{cases} min = \text{کوچکترین داده‌ها} - S \\ max = \text{بزرگترین داده‌ها} + S \end{cases}$

$R = max - min \rightarrow W = \frac{R}{K}$

طول رده‌ها

فراوانی f_i

$Q_1 = P_{0/25}$ $Q_2 = P_{0/50}$ $Q_3 = P_{0/75}$

فراوانی تجمعی g_i

فراوانی نسبی $= r_i$

مرکزی، میانه، مد، میانگین

پراکندگی، دامنه، واریانس، انحراف معیار، ضریب تغییرات

مشخصه‌های آماری

فراوانی نسبی تجمعی $= s_i$

نمایندگی $= m_i$

$D_1 = P_{0/10}$ $D_2 = P_{0/2}$ $D_3 = P_{0/3}$ ---

داده‌های گسسته:

$P_p = (1-z)\, m_{(r)} + z\, m_{(r+1)}$

$r = [(n+1)p]$ جزء صحیح

$z = (n+1)p - r$

پیوسته:

$P_p = L_p + \frac{np - g_F}{f_F}$

فراوانی همان رده $= f_F$

ابتدای رده‌ای که مدرک مرتبه p ام در آن قرار دارد $= L_p$

فراوانی تجمعی داده‌ها قبل از رده مورد نظر g_p

شکست حمله نظامی آمریکا به ایران در طبس(۱۳۵۹ هجری شمسی)

$$\text{حد پیوسته} = M = L_M + \frac{D_2}{D_1+D_2} W$$

در داده های کمی گسسته اگر دو یا چند داده به یک اندازه تکرار شده باشند و بین آنها فاصله نباشد که حد از فرمول (نصف مجموع داده های حد) بدست می آید

هر چند در داده ها نباشد.

$$\begin{cases} \text{میانگین گسسته} \Rightarrow \bar{x}_1 = \frac{1}{n}\sum_{K} x_i \\ \text{میانگین پیوسته} \Rightarrow \bar{x} = \frac{1}{n}\sum_{i=1} x_i F_i \end{cases}$$

واریانس

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

واریانس (میزان پراکندگی):

$$\text{واریانس } \sigma^2 = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x})^2 \geq 0$$

در جامعه

$$\text{انحراف معیار} = \sigma = \sqrt{\sigma^2}$$

در نمونه ←

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

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

$$\bar{x} = \frac{1}{n}\sum F_i x_i$$

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

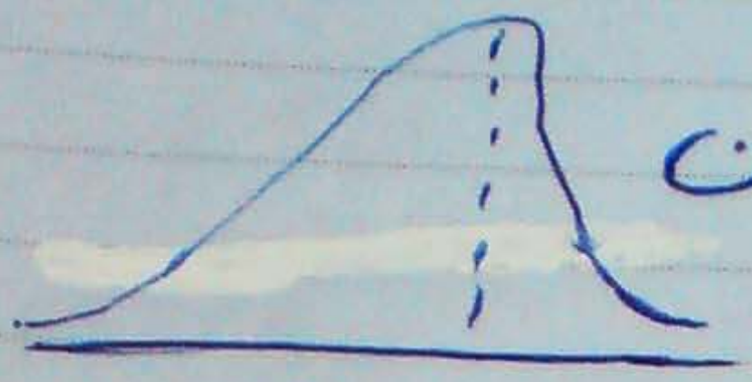
$$\frac{1}{n-1}\left[\sum F_i x_i^2 - \frac{1}{n}\left(\sum F_i x_i\right)^2\right]$$

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

$$\text{ضریب تغییرات} \quad C.V = \frac{s}{\bar{x}}$$

شهادت حضرت فاطمه زهرا (س) ۱۱۲ هجری قمری (تعطیل)

بلبلی برگ گلی خوش رنگ در منقار داشت
واندر آن برگ و نوا خوش ناله های زار داشت
گفتمش در عین وصلی این ناله و فریاد چیست
گفت ما را جلوه معشوق در این کار داشت



داده متقارن

$\bar{x} = M = m$

میانگین میانه مد

$$\begin{cases} x_i \to x_i \\ F_i \xrightarrow{?} F_i \end{cases} \to Z_i = \frac{x_i - \bar{x}}{\sqrt{S^2}} \to \begin{cases} Z_i \to Z_i \\ F_i \to F_n \end{cases} \begin{cases} \bar{Z} = 0 \\ S^2 = 1 \end{cases}$$

ضریب اول چولگی $= \dfrac{2(\bar{x} - m)}{S}$

ضریب دوم چولگی $= \dfrac{\bar{x} - m}{S}$

$\begin{cases} > 0 & \text{چوله به راست} \\ = 0 & \text{چوله متقارن} \\ < 0 & \text{چوله به چپ} \end{cases}$

ضریب کشیدگی $= \dfrac{\frac{Q_3 - Q_1}{2}}{P_{90} - P_1} = 0/263 \begin{cases} > 0 & \text{پهن‌تر} \\ 0 & \text{متقارن} \\ < 0 & \end{cases}$

زیرمجموعه $= A \subset B \Rightarrow$ if $x \in A \Rightarrow x \in B$

دو مجموعه مساوی $= A = B \Leftrightarrow (A \subset B$ و $B \subset A)$

اجتماع ۲ مجموعه $= A \overset{+}{\cup} B = \{x \mid x \in A$ یا $x \in B\}$

اشتراک ۲ مجموعه $= A \overset{\times}{\cap} B = \{x \mid x \in A$ و $x \in B\}$

تفریق $= A - B \{x \mid x \in A$ و $x \notin B\}$

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

روز ملی خلیج فارس

روز شوراها

به مژگان سیه کردی هزاران رخنه در دینم
الا ای همنشین دل که یارانت رفت از یاد

بیا تا گل برافشانیم و می در ساغر اندازیم
فلک را سقف بشکافیم و طرحی نو دراندازیم

بیا تا گل برافشانیم و می در ساغر اندازیم
مروزی مباد آن دم که بی یاد تو بنشینیم

دو پیشامد ناسازگار $A \cap B = \emptyset$

$\forall i \neq j$ if $A_i \cap A_j = \emptyset \rightarrow$ دو به دو ناسازگار A_i

$P : S \rightarrow [0,1]$

۱) $P(S) = 1$

۲) $0 \leq P(A) \leq 1$

۳) $A_1, \dots A_n$ $\quad P(A_1 \cup A_2 \dots) \; P(A_1) + P(A_2) +$

مدلهای احتمال: یک مدل احتمال یعنی نمایش اعضای و فضای نمونه به همراه با احتمال وقوع این اعضاء

متغیر تصادفی: یک متغیر تصادفی تابعی از فضای نمونه به مجموعه اعداد حقیقی است که به همراه هر عضو و فضای نمونه یک عدد حقیقی را نسبت می دهد. متغیرهای تصادفی را با حروف بزرگ لاتین و مقادیر مشاهده شده متغیرهای تصادفی را با حروف کوچک لاتین نمایش می دهیم

«پیوسته»

تابع چگالی احتمال

$F_x(x)$

$0 \leq F_x(x) \leq 1$

$\int F_x(x)\,dx = 1$

«گسسته»

اردیبهشت ۱۳۹۱

تابع احتمال $F_x(x) = P(X = x)$

$\forall x = 0 \leq F_x(x) \leq 1$

$\sum F_x(x) = 1$

تابع برنولی

$$X = \begin{cases} 0 & \text{شکست} \quad 1-p=q \\ 1 & \text{موفقیت} \quad 0<p\leq 1 \end{cases}$$

$X \sim Ber(p)$

$F_X(x) = p^x q^{1-x}$

$E(x) = p$

$V(x) = pq$

۱۲ جمادی الثانی ۱۴۳۳

تعداد موفقیت در n بار آزمایش برنولی مستقل - (دوجمله‌ای)

$X \sim Ber(n, p)$

$F_x(x) = \binom{n}{x} p^x q^{n-x} \longrightarrow E(x) = np$

$Var = npq$

روز بزرگداشت شیخ صدوق

تا که قصر اهل سخت سست بنیاد است

توزیع نرمال مهمترین توزیع یا توزیع پیوسته در صنعت و طبیعت، توزیع نرمال می‌باشد منحنی این توزیع شبیه به یک مد متوسط کاملاً متقارن می‌باشد، گوییم: متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشد هرگاه تابع چگالی آن به فرم زیر باشد:

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

$$-\infty < x < +\infty$$

$$-\infty < \mu < +\infty$$

$$\sigma > 0$$

توزیع نرمال استاندارد توزیع نرمال که دارای میانگین ۰ و واریانس ۱ می‌باشد.

قضیه: اگر σ^2 و μ و N سر X می‌باشد آنگاه به شکل زیر تبدیل بر نرمال استاندارد می‌شود:

$$Z = \frac{X-\mu}{\sqrt{\sigma^2}} \sim N(0,1)$$

$$X \sim N(\mu = 5, \sigma^2 = 16)$$

$$P(8 < x < 12) = P\left(\frac{8-5}{\sqrt{16}} < \frac{X-\mu}{\sqrt{\sigma^2}} < \frac{12-5}{\sqrt{16}}\right)$$

$$= P(0{,}75 < Z < 1{,}75) = P(Z < 1{,}75) - P(Z < 0{,}75)$$

$$= 0{,}9599 - 0{,}7734$$

چرا از دلی غم دیار خود باشم
به شهر خود روم و شهریار خود باشم

چرا به خاک سر کوی یار خود باشم
غم غریبی و غربت چو بر نمی‌تابم

ص ۴۱۳ کتاب

روز بزرگداشت فردوسی

نکته: هر ترکیب خطی از متغیرهای تصادفی نرمال مستقل، دارای توزیع نرمال می‌باشد.

نمونه تصادفی: اگر جامعه تحت بررسی دارای تابع توزیع و احتمال $F_X(x)$ باشد آنگاه یک نمونه تصادفی به حجم n از این جامعه عبارتست از انتخاب متغیرهای تصادفی $X_1 \dots X_n$ با مقادیر $x_1 \dots x_n$ با مقدارت $x_1, \dots, x_n$

(نمونه تصادفی از لحاظ آماری یعنی نمونه‌ها هم توزیع و مستقل باشند.)

پارامتر و آماره: هر ویژگی مربوط به جامعه را پارامتر می‌نامیم و ویژگی متناظر با آن در نمونه را آماره می‌گوییم

روز بزرگداشت حکیم عمر خیام

پارامتر

میانگین جامعه $= \mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

واریانس جامعه $= \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$

آماره

میانگین نمونه $= \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

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

پارامتر — آماره

جامعه — نمونه

نامعلوم — معلوم

متغیر

لب باز مگیر یک زمان از لب جام
در جام جهان چو تلخ و شیرین به هم است
این از لب یار خواه و آن از لب جام

اردیبهشت ۱۳۹۱

نمونه‌گیری تصادفی ساده

	تعداد نمونه‌های ممکن	احتمال انتخاب
با جاگذاری	N^n	$\frac{1}{N^n}$
بدون جاگذاری	$\binom{N}{n}$	$\frac{1}{\binom{N}{n}}$

N = حجم جامعه

n = حجم نمونه

جامعه = ۱ ۲ ۳ ۴

$N=4$ تصادفی ساده با جاگذاری $N^n = 4^2 = 16$

$n=2$

$\bar{x}$ = (۱و۱) ① (۲و۱) ۱٫۵ (۳و۱) ۲ (۴و۱) ۲٫۵

میانگین (۱و۲) ۱٫۵ (۲و۲) ۲ (۳و۲) ۲٫۵ (۴و۲) ۳

(۱و۳) ② (۲و۳) ۲٫۵ (۳و۳) ۳ (۴و۳) ۳٫۵

(۱و۴) ۲٫۵ (۲و۴) ۳ (۳و۴) ۳٫۵ (۴و۴) ۴

$\bar{x}$	۱	۱٫۵	۲	۲٫۵	۳	۳٫۵	۴
$P(\bar{x}=\bar{x})$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

توزیع نمونه‌ای: همان طور که ملاحظه کردید آماره تابعی از نمونه تصادفی است و مقدار آن از نمونه‌ای به نمونه دیگر تغییر می‌کند یعنی احتمال می‌شود که بتوانیم رفتار آماره را در قالب احتمال بیان کنیم بنابراین خود آماره یک متغیر تصادفی و توزیع احتمال آن را توزیع نمونه‌ای می‌نامیم.

توزیع نمونه‌ای میانگین نمونه $\bar{X}$

* اگر جامعه تحت بررسی دارای توزیع احتمال $F_X(x)$ با میانگین μ و واریانس σ^2 باشد و بخواهیم توزیع میانگین نمونه تصادفی به حجم n یعنی $X_1 - X_n$ را بدست آوریم دو حالت زیر را در نظر می‌گیریم.

۱) اگر جامعه تحت بررسی دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشد چون $\bar{X}$ یک ترکیب خطی از متغیرهای تصادفی مستقل نرمال باشد خود نیز دارای توزیع نرمال با میانگین μ و واریانس $\frac{\sigma^2}{n}$ می‌باشد:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

۲ اگر توزیع جامعه نرمال نباشد، دیگر به راحتی نمی‌توان در مورد توزیع $\bar{X}$ تکرار داد ولی می‌دانیم با افزایش حجم نمونه توزیع $\bar{X}$ نیز به توزیع نرمال می‌گراید، این نکته از میان اهمیتی برخوردار است که آن را در قالب قضیه زیر بیان می‌کنیم:

قضیه حد مرکزی: اگر $X_1, X_2, \dots$ دنباله‌ای از متغیرهای تصادفی نرمال با میانگین μ و واریانس σ^2 (متناهی) باشند طوری که هر تعداد متناهی از این متغیرها از یکدیگر مستقل باشند آنگاه زمانی که $n \to +\infty$

بنابراین توزیع $Z \underset{n \to +\infty}{\overset{\lim}{=}} \dfrac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

نکته: زمانی که $n \geq 30$ باشد قضیه حد مرکزی مناسب می‌باشد.

توزیع مربع χ^2 (کای دو): اگر $Z \sim N(0,1)$ باشد آنگاه Z^2 دارای توزیع χ^2 با یک درجه آزادی است و آن را با نماد زیر نشان می‌دهیم

$Y \sim \chi^2(1)$

اگر $Z_1, Z_2, \dots, Z_n$ n متغیر تصادفی مستقل نرمال استاندارد باشند آنگاه $Y = Z_1^2 + Z_2^2 + Z_n^2$ دارای توزیع χ^2 با n درجه آزادی است.

حجاب چهره جان می‌شود غبار تنم
خوشا دمی که از آن چهره پرده برفکنم

چنین قفس نه سزای چو من خوش‌الحانیست
گشا ضیاء کز آن حم

$Y \sim \chi^{(2)}(n)$

$y > 0$

$E(Y) = n$

$var(Y) = 2n$

قضیه: اگر از جامعه‌ای نرمال با میانگین μ و واریانس σ^2 نمونه‌ای تصادفی به حجم n با میانگین $\bar{x}$ و واریانس s^2 انتخاب کنیم، آنگاه:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

توزیع t (student): متغیر تصادفی Z دارای توزیع نرمال با میانگین صفر و واریانس ۱ است یعنی $Z \sim N(0,1)$ و $Y \sim \chi^2(n)$ دو متغیر تصادفی مستقل از هم باشند، آنگاه متغیر تصادفی

$$T = \frac{Z}{\sqrt{\frac{Y}{n}}}$$

دارای توزیع T (student) با n درجه آزادی است و آن را با علامت زیر نشان می‌دهیم $T \sim t(n)$ و $-\infty < t < +\infty$

ولادت حضرت امام محمد تقی(ع)(۱۹۵ هجری قمری)

حسب حالی ننوشتی و شد ایامی چند

محرمی کو که فرستم به تو پیغامی چند

قضیه: اگر از جامعه‌ای نرمال با میانگین μ و واریانس σ^2 نمونه‌ای تصادفی به حجم n با میانگین $\bar{X}$ و واریانس S^2 انتخاب کنیم آنگاه

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

$n-1$ درجه آزادی ←

توزیع نمونه‌ای اختلاف میانگین‌ها:

اگر از دو جامعه نرمال مستقل با واریانس‌های σ_1^2 و σ_2^2 و میانگین‌های μ_1 و μ_2 نمونه‌های تصادفی مستقل به حجم‌های n_1 و n_2 با میانگین‌های $\bar{X}_1$ و $\bar{X}_2$ و واریانس‌های S_1^2 و S_2^2 انتخاب می‌کنیم و بخواهیم توزیع نمونه‌ای اختلاف میانگین‌های نمونه‌ای ($\bar{X}_1$ و $\bar{X}_2$) را بدست آوریم به صورت زیر عمل می‌کنیم:

۱ حالت اول: اگر واریانس‌های این دو جامعه نرمال معلوم باشند آنگاه با توجه به مطالب قبل می‌دانیم $\bar{X}_1 \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$ و $\bar{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$

چون $\bar{X}_1$ و $\bar{X}_2$ ترکیب خطی از متغیرهای نرمال مستقل می‌باشند

رحلت حضرت امام خمینی(ره)(۱۳۶۸ هجری شمسی) (تعطیل)-
انتخاب حضرت آیت الله امام خامنه ای به رهبری(۱۳۶۸ هجری شمسی)

و با توجه به مستقل بودن دو جامعه از همدیگر چون $\bar{X}_1 - \bar{X}_2$ نیز یک ترکیب خطی از متغیرهای تصادفی مستقل نرمال می‌باشد، دارای توزیع نرمال به شکل زیر است

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2 , \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

پس

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

۲) اگر واریانس‌های این دو جامعه نرمال معلوم نباشند ولی برابر باشند، آنگاه طبق مطالب خوانده شده خواهیم داشت.

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t(n_1 + n_2 - 2)$$

واریانس آمیخته

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

توزیع فیشر، اگر $X \sim \chi^2(n)$ و $Y \sim \chi^2(m)$ دو متغیر مستقل باشند، آنگاه متغیر تصادفی $F = \dfrac{\frac{X}{n}}{\frac{Y}{m}}$ دارای توزیع فیشر

با درجه آزادی مفروض

$F \sim F(n, m)$ است

دو جامعه آزادی صورت

قضیه اگر σ_1^2 و σ_2^2 به ترتیب واریانس های دو جامعه نرمال مستقل باشند و از این دو جامعه ما نمونه های تصادفی و مستقل به حجم های n_1 و n_2 با واریانس های S_1^2 و S_2^2 انتخاب کنیم آنگاه

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F(n_1 - 1, n_2 - 1)$$

استنباط آماری: روشی است که بوسیله آن اطلاعات حاصل از نمونه تصادفی و تجزیه و تحلیل آنها در مورد پارامتر نامعلوم جامعه یا توزیع یک متغیر تصادفی یا چند متغیر

وفات حضرت زینب (س)(۶۲ هجری قمری) - تغییر قبله مسلمین از بیت المقدس به مکه معظمه(۲ هجری قمری)

تصادفی تکیه شیری های انجام می دهد

استنباط آماری دارای ۲ جنبه زیر است:

۱. برآورد پارامتر مجهول جامعه که به ۲ روش نقطه ای و فاصله ای انجام می گیرد

۲. آزمون فرض در مورد پارامترهای نامعلوم جامعه

برآورد گر نقطه ای:

برآورد نقطه ای یک پارامتر تابعی از نمونه تصادفی است. بعنوان مثال برآورد نقطه ای میانگین جامعه همان میانگین نمونه است و برآورد نقطه ای واریانس جامعه همان واریانس نمونه است یعنی

$$\begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = S^2 \end{cases}$$

برآورد گر نااریب: گوییم برآورد نقطه ای T برای پارامتر نامعلوم θ نااریب است اگر: $E(T) = \theta$ اگر T_1 و T_2 دو برآورد گر نااریب پارامتر مجهول θ باشند، آنگاه برآورد گر T_1 را کارآتر از T_2 می نامیم هرگاه $var(T_2) > var(T_1)$

خوشا دلی که مدام از پی نظر نرود
به هر درش که بخوانندی خبر نرود

نکته: در بین تمام برآوردگرهای نااریب و با کمترین واریانس جامعه به ترتیب ماکسیمم و مینیمم نمونه دارای کمترین واریانس می باشند.

برآورد فاصله ای: در روش برآوردی نقطه ای چون با تغییر نمونه، مقدار برآوردگر نیز تغییر می کند بنابراین، برآوردگر نقطه ای دارای خطای زیادی است بجای آن از برآورد فاصله ای که خطای آن کمتر است استفاده می شود در روش برآورد فاصله ای (L و U) از اعداد حقیقی را طوری در نظر می گیریم که با یک احتمال زیادی پارامتر نامعلوم جامعه را در بر داشته باشد. اگر این فاصله با احتمال $(1-\alpha)$ پارامتر نامعلوم جامعه را در بر داشته باشد، آن را یک فاصله اطمینان $(1-\alpha)$ درصدی می نامیم

پارامتر بین فاصله و L و U را حد بالای فاصله و $(1-\alpha)$ را میزان اطمینان $(1-\alpha)$ درصدی می نامیم و بنابراین یک فاصله اطمینان $(1-\alpha)$ درصدی برای پارامتر مجهول جامعه (θ) عبارت است از

$$P(L < \theta < U) = 1-\alpha$$

فاصله‌ی اطمینان برای میانگین جامعه نرمال زمانی که واریانس جامعه معلوم باشد. اگر از جامعه‌ای نرمال با میانگین نامعلوم و واریانس معلوم تصادفی به حجم n میانگین $\bar{x}$ انتخاب کنیم آنگاه با توجه به مشاهدات نمونه‌ی اخذ شده یک فاصله اطمینان $(1-\alpha)$ درصدی برای میانگین این جامعه نرمال به شکل زیر بدست می‌آید:

$$\mu \in \left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

میانگین جامعه

انحراف معیار جامعه

مقدار جدول نرمال استاندارد

حجم نمونه

مثال: ص ۲۳۸

یک تولید کننده لامپ های روشنایی، لامپ‌ها را تولید می‌کند که طول عمر آنها دارای توزیع نرمال با انحراف معیار ۴۰ ساعت است. اگر یک نمونه

۲۶

۲۴ رجب ۱۴۳۳

تصادفی ۳۰ تایی از لامپ‌ها دارای متوسط عمر ۸۷۵ ساعت باشد یک فاصله اطمینان ۹۵٪ درصدی برای متوسط طول عمر تمام لامپ‌های تولیدی این کارخانه بدست آورده و تفسیر کنید.

شهادت سربازان دلیر اسلام: بخارایی، امانی، صفار هرندی و نیکنژاد (۱۳۴۴ مجری شمسی)

جدول ص ۱۳۰

دانی که چنگ و عود چه تقریر می‌کنند — پنهان خورید باده که تعزیر می‌کنند

$\sigma = ۴۰$

$n = ۳۰$ $\qquad \mu \in \left(۸۷۰ - ۱,۹۶\frac{۴۰}{\sqrt{۳۰}} , ۸۷۰ + ۱,۹۶\frac{۴۰}{\sqrt{۳۰}}\right)$

$\bar{x} = ۸۷۰$ $\qquad \mu \in (۸۷۰ - ۱۴,۳۱۳ , ۸۷۰ + ۱۴,۳۱۳)$

$1-\alpha = ۰/۹۵$ $\qquad \mu \in (۸۵۵,۶۸۷ , ۸۸۴,۳۱۳)$

$\alpha = ۰/۰۵$ $\qquad \mu \in (۸۵۵,۶۸۷ , ۸۸۴,۳۱۳)$

$\frac{\alpha}{۲} = ۰/۰۲۵$ $\qquad$ طول عمر لامپ : $۸۷۰ \pm ۱۴,۳۱۳$

$Z_{\alpha/۲} = Z_{۰/۰۲۵} = -۱,۹۶۶$

تفسیر: با ۹۵٪ اطمینان با توجه به نمونهٔ انتخابی متوسط طول عمر لامپ‌های تولیدی کارخانه در این فاصله قرار می‌گیرد.

طول عمر لامپ $= ۸۷۰ \pm ۱۴,۳۱۳$

مثل ص ۲۳۸: طول قد دانشجویان یک دانشکده دارای توزیع نرمال با انحراف معیار ۵cm است. نمونه‌ای تصادفی به حجم ۵۰ از دانشجویان انتخاب کرده و متوسط طول قدآن‌ها ۱۶۸ سانتی‌متر بدست می‌آید. یک فاصله اطمینان ۹۶٪ برای متوسط طول قد همهٔ دانشجویان این دانشکده بدست آورید.

شهادت حضرت امام موسی کاظم(ع)(۱۸۳ هجری قمری) - روز جهاد کشاورزی

در ازل پرتو حسنت ز تجلی دم زد
عشق پیدا شد و آتش به همه عالم زد

$\sigma = 5$

$n = 50$

$\bar{x} = 178$

$1-\alpha = 0/94$

$\alpha = 0/06$

$\frac{\alpha}{2} = 0/02$

در کتاب نیست $z = \frac{(-2/05)+(-2/06)}{2} = -2/055$

فاصلهٔ اطمینان برای میانگین جامعه نرمال زمانی که واریانس جامعه معلوم نباشد؟

اگر از جامعه‌ای نرمال با میانگین μ و واریانس σ^2 نمونه‌ای تصادفی به حجم n با میانگین $\bar{x}$ و واریانس s^2 انتخاب کنیم، یک فاصله اطمینان $(1-\alpha)$٪ برای میانگین این جامعهٔ نرمال به شکل زیر بدست می‌آید:

$$\mu \in \left(\bar{x} - \underbrace{t_{\frac{\alpha}{2}}(n-1)}_{\text{مقدار جدول } t \text{ - استیودنت}} \frac{s}{\sqrt{n}} , \bar{x} + t_{\frac{\alpha}{2}}(n-1) \frac{s}{\sqrt{n}}\right)$$

از افزایش میانگین نمونه

در دیر مغان آمد یارم قدحی در دست
مست از می و میخواران از نرگس مستش مست

برگزیدهٔ سیرت رسول اکرم(ص)(۱۳ سال قبل از هجرت) (تعطیل) -
درگذشت دکتر علی شریعتی (۱۳۵۶ هجری شمسی)

مثال) - وزن ۱۰ جعبه از یک ماده غذایی به شکل زیر بدست آمده با فرض نرمال بودن داده ها یک فاصله اطمینان ۹۹٪ برای متوسط وزن تمام بسته های این نوع ماده غذایی بدست آورید:

۱۰٫۴ و ۹٫۹ و ۹٫۸ و ۱۰٫۱ و ۱۰٫۳ و ۱۰٫۱ و ۹٫۷ و ۱۰٫۲ و ۱۰٫۳ و ۹٫۸

ع غ

$h = 10$

$\bar{x} = 10.06$

$S = 0.245$

$\mu \in \left(10.06 - 3.250 \frac{0.245}{\sqrt{10}} , 10.06 + 3.250 \frac{0.245}{\sqrt{10}}\right)$

$\mu \in (9.890 , 10.311)$

$\mu \in (10.06 - 0.251 , 10.06 + 0.251)$

متوسط وزن بسته ها : 10.06 ± 0.251

$1-\alpha = 0.99 \rightarrow \alpha = 0.01 \rightarrow \frac{\alpha}{2} = 0.005$

$t_{\frac{\alpha}{2}}(n-1) = t_{0.005}(9) = 3.250$

ستون افقی

ستون عمودی

شهادت دکتر مصطفی چمران (۱۳۶۰ هجری شمسی) - روز بسیج اساتید

در نظربازی ما بی خبران حیرانند
من چنینم که نمودم دگر ایشان دانند
عاقلان نقطه پرگار وجودند ولی
عشق داند که در این دایره سرگردانند

$t_{\alpha}(n) = t_{1-\alpha}$

مکمل

اگر α

در چون نور

فاصله اطمینان برای تفاضل میانگین های دو جامعه نرمال:

اگر از دو جامعه نرمال مستقل ما با میانگین های μ_1 و μ_2 و واریانس های σ_1^2 و σ_2^2 نمونه های تصادفی و مستقل به n_1 و n_2 با میانگین های $\bar{x}_1$ و $\bar{x}_2$ انتخاب کنیم و بخواهیم واریانس های (S_1^2 و S_2^2) و بخواهیم برای $\mu_1 - \mu_2$ یک فاصله اطمینان بسازیم دو حالت زیر را در نظر می گیریم:

۱- اگر واریانس های این ۲ جامعه نرمال معلوم باشد، آنگاه با توجه به مشاهدات نمونه ای گردآوری شده یک فاصله اطمینان $(1-\alpha)$ درصدی برای تفاضل میانگین های این دو جامعه نرمال عبارت است از:

$$\mu_1 - \mu_2 \in (\bar{x}_1 - \bar{x}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\bar{x}_1 - \bar{x}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

اگر واریانس های این ۲ جامعه نرمال معلوم نباشند ولی برابر باشند آنگاه یک فاصله اطمینان $(1-\alpha)$ درصدی برای اختلاف میانگین های

ولادت حضرت امام حسین (ع)(۳) هجری قمری) و روز پاسدار

ولادت حضرت ابوالفضل العباس(ع)(۲۶هجری قمری) و روز جانباز

در نازم خم ابروی تو یاد آمد

حالتی رفت که محراب

این دو جامعه فرمان عبارت است از

$$\mu_1 - \mu_2 \in \left((\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right.,$$

$$\left.(\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$$

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

تفسیر فاصله‌ی اطمینان برای اختلاف میانگین های ۲ جامعه نرمال:

در تفسیر فاصله‌ی اطمینان برای اختلاف میانگین های ۲ جامعه نرمال سه حالت زیر را در نظر می‌گیریم

۱) اگر هر طرف فاصله بدست آمده مثبت باشند با $(1-\alpha)$ درصد اطمینان و با توجه به نمونه‌های انتخابی نتیجه می‌گیریم که میانگین جامعه اول از میانگین جامعه دوم بیشتر است.

۲) اگر دو طرف فاصله بدست آمده منفی باشند، با توجه به نمونه‌های انتخابی

ولادت حضرت امام زین العابدین (ع)(۳۸ هجری قمری) - روز مالیات

در همه دیر مغان نیست چو من شیدایی
خرقه جای گرو باده و دفتر جایی

با $(1-\alpha)$ درصد اطمینان نتیجه می‌گیریم که میانگین جامعه دوم از میانگین جامعه اول بیشتر است.

۳) - اگر یک طرف فاصله بدست آمده منفی و طرف دیگر آن مثبت باشد (شامل عدد صفر باشد) با $(1-\alpha)$ درصد اطمینان و با توجه به نمونه‌های انتخابی نتیجه می‌گیریم که میانگین‌های دو جامعه هیچگونه اختلافی با یکدیگر ندارند.

فاصله اطمینان برای واریانس یک جامعه نرمال: اگر از جامعه‌ای نرمال با میانگین μ و واریانس σ^2 نمونه‌ای تصادفی به حجم n با واریانس S^2 انتخاب کنیم، آنگاه یک فاصله اطمینان $(1-\alpha)$ درصدی برای واریانس این جامعه نرمال عبارت است از

روز مبارزه با سلاحهای شیمیایی و میکروبی

$$\sigma^2 \in \left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} , \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right)$$

جدول χ^2 صفحه ۴۱۵ کتاب ستون عمودی سمت آزادی

شهادت مظلومانه آیت الله دکتر بهشتی و ۷۲ تن (۱۳۶۰ هجری شمسی) - روز قوه قضاییه

فاصلهٔ اطمینان برای نسبت واریانس‌های دو جامعه نرمال:

اگر از دو جامعه نرمال با واریانس‌های σ_1^2 و σ_2^2 نمونه‌هایی تصادفی و مستقل و حجم‌های n_1 و n_2 با واریانس‌های S_1^2 و S_2^2 انتخاب کنیم و بخواهیم یک فاصلهٔ اطمینان $(1-\alpha)$ درصدی برای نسبت واریانس‌های این دو جامعهٔ نرمال بدست آوریم از رابطهٔ زیر استفاده می‌شود:

$$\frac{\sigma_1^2}{\sigma_2^2} \in \left(\frac{S_1^2}{S_2^2} \times \frac{1}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)} , \frac{S_1^2}{S_2^2} F_{\frac{\alpha}{2}}(n_2-1, n_1-1) \right)$$

اگر آزمایش در جداول فیشر مقدار مورد نظر وجود نداشت می‌توانیم با استفاده از رابطهٔ زیر آن را بدست آوریم

$$F_\alpha(a,b) = \frac{1}{F_{1-\alpha}(b,a)}$$

در تحلیل فاصلهٔ اطمینان برای نسبت واریانس‌های دو جامعهٔ نرمال ۳ حالت زیر را در نظر می‌گیریم:

۱) اگر هر دو طرف فاصله کمتر از یک باشد، یا کوچک به عنوان نمونه‌های انتخابی و با $(1-\alpha)$ درصد اطمینان نتیجه می‌گیریم که

ولادت حضرت علی اکبر (ع) (۳۳ هجری قمری) و روز جوان - روز صنعت و معدن
شهادت آیت الله صدوقی چهارمین شهید محراب(۱۳۶۱ هجری شمسی)

میزان پراکندگی جامعه اول از جامعه دوم کمتر است.

۲- اگر هر دو طرف فاصله بیشتر از ۱ باشد (با توجه به نمونه های انتخابی) و با (۱-α) درصد اطمینان و با توجه به نمونه های انتخابی نتیجه می گیریم که واریانس جامعه دوم از جامعه اول کمتر است.

۳- اگر یک طرف فاصله کمتر از ۱ و طرف دیگر آن بیشتر از ۱ باشد (فاصله شامل عدد ۱ باشد) با (۱-α) درصد اطمینان و با توجه به نمونه های انتخابی نتیجه می گیریم که میزان پراکندگی دو جامعه هیچ تفاوتی با یکدیگر ندارند.

آزمون فرض آماری:

یک فرض آماری ادعا یا گزاره ای است در مورد پارامتر مجهول جامعه یا توزیع جامعه تحت بررسی که ممکن است درست یا نادرست باشد.

از دو فرض که در یک آزمون فرض آماری مطرح می شود یکی را فرض صفر یا فرض صفر می نامند و آن را با H_0 نمایش می دهیم و دیگری را

سقوط هواپیمای مسافربری ایران توسط ناوگان آمریکای جنایتکار (۱۳۶۷ هجری شمسی)

دست از طلب ندارم تا کام من برآید
یا تن رسد به جانان یا جان ز تن برآید

بگشای تربتم را بعد از وفات و بنگر
کز آتش درونم دود از کفن برآید

فرض مقابل (ادعا) نامیدیم و آن را با H_1 نمایش می‌دهیم.

در تمام مواقع ادعای مطرح شده در فرض H_1 قرار می‌گیرد و ضد ادعا یا فرض موجود در H_0 قرار می‌گیرد بجز موقعی که ادعا حالت مساوی داشته باشد. بنابراین تنها در این مورد ادعا را در H_0 قرار داده و ضد ادعا را در H_1 قرار می‌دهیم.

برای انجام یک آزمون فرض آماری بر اساس n مشاهده یک نمونه تصادفی به شکل زیر عمل می‌کنیم.

۱) اگر مشاهدات نمونه انتخاب شده با فرض H_0 سازگار باشند، بنابراین نتیجه می‌گیریم فرض H_0 درست است یا به عبارتی H_0 را می‌پذیریم

و اگر نتایج حاصله از نمونه انتخابی با H_0 سازگار نباشد فرض H_0 را نمی‌پذیریم یا آن را رد می‌کنیم بنابراین با توجه به مطالب گفته شده برای انجام آزمون آماری نیاز به آماره آزمون و ناحیه رد آزمون نیاز داریم

ولادت حضرت قائم (عج) (۲۵۵ هجری قمری)(تعطیل) و روز جهانی مستضعفان

روز قلم

آماره آزمون و ناحیه بحرانی:

آماره $T=T(X_1, \dots, X_n)$ را آماره آزمون می‌نامیم و براساس آن ادعایی را می‌پذیریم یا رد می‌کنیم ناحیه‌ای که براساس آن فرض را رد می‌کنیم. ناحیه بحرانی نامیده می‌شود و آن را با C نمایش می‌دهیم. مکمل ناحیه بحرانی را ناحیه پذیرش می‌نامیم و آن را با $\bar{C}$ نمایش می‌دهیم. هرگاه براساس مشاهدات یک نمونه تصادفی مقدار مشاهده شده آماره آزمون را بدست آوریم، آنگاه به شکل زیر تصمیم می‌گیریم.

اگر مقدار آماره در ناحیه C قرار بگیرد فرض H را رد می‌کنیم و در غیر این صورت آن را می‌پذیریم از فرض‌هایی که در یک مسأله آزمون فرض مطرح می‌شود یک نوع را فرض ساده و دیگری را فرض مرکب می‌نامیم. فرض ساده فرضی است که براساس آن توزیع جامعه کاملاً مشخص است و فرض مرکب فرضی است که براساس آن

دل می‌رود ز دستم صاحب‌دلان خدا را
کشتی شکستگانیم ای باد شرطه برخیز

درد اگر راز پنهان خواهد شد آشکارا
باشد که باز بینیم دیدار آشنا را

توزیع جامعه بصورت واضح معلوم نیست.

آزمون های آماری به دو صورت یک طرفه و دو طرفه وجود دارد. آزمون یکطرفه، آزمونی است که در آن فرض مقابل یکطرف باشد ($>$ یا $<$) و آزمون مرکب آزمونی است که در آن دو طرف فرض مقابل دو طرف باشد

دو طرفه $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{cases}$ یک طرفه $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta < \theta_0 \text{ یا } \theta > \theta_0 \end{cases}$

انواع خطاهای آماری:

در یک مسئله آزمون فرض آماری، ممکن است مرتکب خطا شویم، زمانی ممکن است فرض H_0 درست باشد و آن را رد کنیم به این نوع خطا خطای نوع ۱ گفته می شود

احتمال وقوع خطای نوع ۱ را با α نمایش داده و آن را سطح معنی داری یا تشخیص آزمون می نامیم و یا ممکن است فرض H_1 درست نباشد و آن را بپذیریم به این خطا، خطای نوع ۲ گفته می شود و احتمال وقوع خطای ۲ را با β نمایش می دهیم

دلا بسوز که سوز تو کارها بکند
نیاز نیم شبی دفع صد بلا بکند

توان آزمون عبارت است از احتمال انجام دادن یک کار درست یعنی فرض H_1 درست نیست و ما نیز به حق فرض H_1 را رد کنیم

$\beta^* = 1 - \beta$

مراحل انجام آزمون فرض یک آزمون آماری:

1. تعیین فرض H_0 و H_1
2. تعیین سطح معنی داری آزمون α
3. تعیین آماره آزمون
4. تعیین ناحیه بحرانی آزمون ما که براساس آماره آزمون و
5. بدست می آید
6. بدست آوردن مقدار مشاهده شده آماره آزمون،

6. نتیجه گیری

اگر مقدار آماره در ناحیه بحرانی قرار گرفت فرض H_0 را رد می کنیم در غیر اینصورت آن را می پذیریم

آزمون برازندگی :

حالات جامعه	۱	۲	----	----	r	جمع
مشاهدات o_i	o_1	o_2	----	----	o_r	n
مقادیر مورد انتظار	e_1	e_2	----	----	e_r	n

H_0 = جامعه دارای توزیع $F_x(\cdot)$ است (وضع جامعه مطلوب است)

H_1 = جامعه دارای توزیع نیست (وضع جامعه مطلوب نیست)

$$\chi^2 = \sum_{i=1}^{r} \frac{(o_i - e_i)^2}{e_i}$$

آماره آزمون

H_0 ناحیه رد = $\chi^2_{\text{آماره}} > \chi^2_{1-\alpha}(r-1)$

آزمون استقلال در جداول توافقی

عامل ۲ \ عامل ۱	۱	۲	...	c	
۱	O_{11}	O_{12}	...	O_{1c}	$O_{1.}$ مقدار مشخص شده
۲	O_{21}	O_{22}	...	O_{2c}	$O_{2.}$
⋮	⋮	⋮	⋮		
r	O_{r1}	O_{r2}	...	O_{rc}	$O_{r.}$
$O_{.j}$	$O_{.1}$	$O_{.2}$	...	$O_{.c}$	$O_{..}$

$$\begin{cases} O_{.i} = \sum_{j=1}^{c} O_{ij} \\ O_{.j} = \sum_{i=1} O_{ij} \\ O_{..} = \sum_{i=1}^{c} \sum_{j=1}^{c} O_{ij} \end{cases}$$

مقدار مورد انتظار $= e_{ij} = \dfrac{O_{i.} \times O_{.j}}{O_{..}}$

$H_0 =$ (بر روی هم تاثیر ندارند) عامل ۱ و ۲ مستقلند

$H_1 =$ (بر روی هم تاثیر دارند) عامل ۱ و ۲ مستقل نمی‌باشند

آماره آزمون $\chi^2 = \sum_{i=1}^{c} \sum_{j=1}^{c} \dfrac{(O_{ij} - e_{ij})^2}{e_{ij}}$

H_0 رد می‌شود $\cdot\ \chi^2_{\text{آماره}} > \chi^2_{1-\alpha}\left[(r-1)(c-1)\right]$