Subject:

date:

توزیع مربع کای :

می‌دانیم $Z \sim N(0,1)$ ، بدین صورت اگر متغیر نرمال استاندارد شده یا Z را به توان ۲ برسانیم ، می‌توان گفت که

متغیر Y از توزیع مربع کای با توان دوم کای با درجه آزادی df $\leftarrow$ $(df = n-1)$

$$Y = Z^2 = \sum_{i=1}^{k} \left(\frac{x_i - \mu}{\sigma}\right)^2$$

بنابراین می‌نویسیم $Y \sim X^2(df)$. همچنین اگر از یک جامعه نرمال استاندارد شده ، نمونه‌ای مختلفی استخراج و برای هر نمونه مقدار Z محاسبه گردد ، می‌توان گفت ، سری Z ها ی محاسباتی یعنی $Z_1, Z_2, Z_3, \dots, Z_n$

براساس توزیع ذیل : $Y = Z_1^2 + Z_2^2 + \dots + Z_n^2$ از توزیع مربع کای با درجه آزادی $df = n-1$ تبعیت می‌کند .

تابع چگالی احتمال متغیر تصادفی X^2 که از توزیع کای - اسکوئر با درجه آزادی df تبعیت می‌کند ، عبارت است از :

$$f(x^2) = \begin{cases} \dfrac{1}{2^{df/2}\,\Gamma(df/2)} (x^2)^{\frac{df}{2}-1} e^{-x^2/2} & 0 \le x^2 < \infty \\ 0 & x^2 < 0 \end{cases}$$

در این رابطه Γ تابع گاما می‌باشد که به صورت زیر تعریف می‌شود :

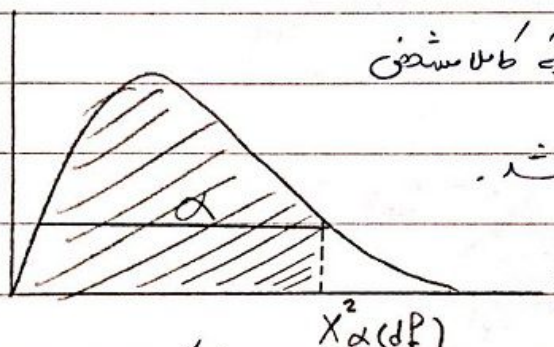
$$\Gamma(\beta) = \int_0^{+\infty} x^{\beta-1} e^{-x}\, dx \qquad \beta > 0$$

اما توضیح مناسب : می‌دانیم در آزمایش تصادفی پرتاب سکه ، زمان لازم برای پیشامد اولین موفقیت دارای توزیع نمایی با پارامتر $\theta = 1/4$ است . در چنین آزمایشی ، زمان لازم برای ۳ امین موفقیت از توزیع گاما تبعیت می‌کند .

منحنی تابع چگالی احتمال توزیع مربع کای به این شکل است ، همانطور کاملا مشخص

است مقادیر X^2 منفی نیست و منحنی توزیع حول $X^2 = 0$ متقارن نمی‌باشد .

جدول شماره ۳ ، سطح زیر منحنی این توزیع و محور x ها را به ازای مقادیر



مختلف X^2 و درجات آزادی مختلف نشان می‌دهد . $X^2_{\alpha(df)}$ مقداری بر روی محور افقی X^2 می‌باشد که سطح ناحیه

سمت راست آن برابر α می‌باشد بنابراین می‌توان گفت :

$$Pr(X^2 \le X^2_{\alpha(df)}) = \alpha$$

Scanned with CamScanner

Subject:

date:

بعنوان مثال با دقت به جدول می‌توان گفت :

$$X^2_{0.9(20)} = 28.4 \Rightarrow Pr(X^2 \leq 28.4_{(20)}) = 28.4$$

$$X^2_{0.95(15)} = 25 \Rightarrow Pr(X^2 \leq 25_{(15)}) = 0.95$$

مثال : اگر متغیر X^2 از توزیع مربع کای با درجه آزادی $df = 18$ تبعیت کند ($X^2 \sim X^2_{(18)}$) مطلوب است :

الف) $Pr(X^2 \leq 7.01)$

$$X^2_{0.01(18)} = 7.01 \Rightarrow Pr(X^2 \leq 7.01) = 0.01$$

ب) اگر $Pr(X^2 > K) = 0.05$ باشد مطلوب محاسبه مقدار K :

$$Pr(X^2 > K) = 0.05 \Rightarrow Pr(X^2 \leq K) = 0.95 \Rightarrow K = X^2_{0.95(18)} = 28.9$$

اگر $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از جامعه نرمال با میانگین μ و واریانس σ^2 باشد، می‌توان گفت :

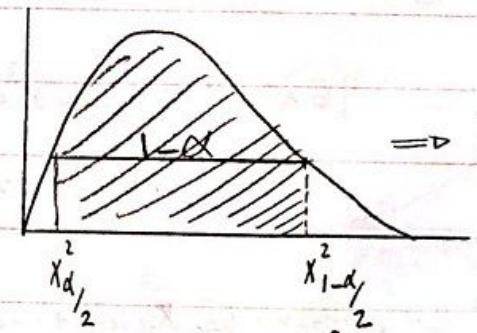
$$\frac{(n-1)S^2}{\sigma^2} \sim X^2_{\alpha(df)}$$

مثال : یک جامعه نرمال دارای واریانس 6 است. اگر نمونه تصادفی 25تایی از این جامعه انتخاب شود احتمال اینکه واریانس نمونه بین 3.45 تا 10.75 باشد را بیابید :

$$P(3.45 < S^2 < 10.75) = Pr\left(\frac{24 \times 3.45}{6} < \frac{(n-1)S^2}{\sigma^2} < \frac{24 \times 10.75}{6}\right)$$

$$Pr(13.8 < X^2_{(24)} < 43) = Pr(X^2_{(24)} < 43) - Pr(X^2_{(24)} < 13.8)$$

$0.99 - 0.05 = 0.94$

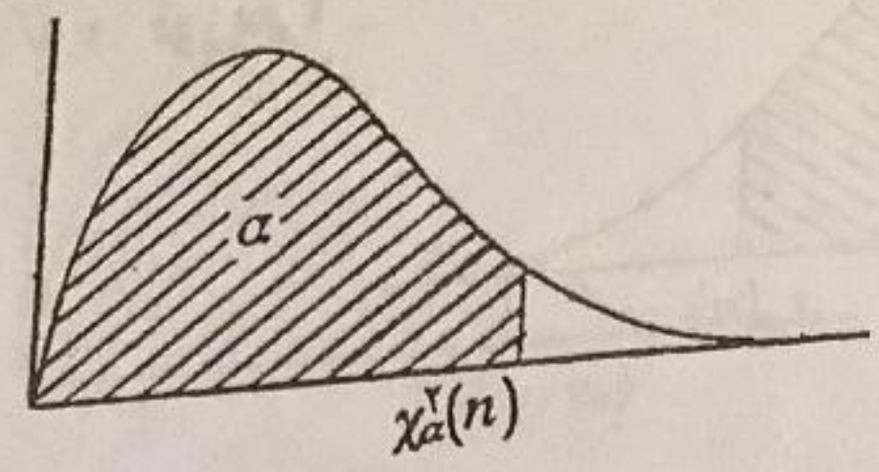


$$\Rightarrow Pr(X^2_{\alpha/2} < X^2 < X^2_{1-\alpha/2}) = 1-\alpha$$

$$X^2_{0.05(24)} = 13.8 \qquad X^2_{0.99(18)} = 43$$

Scanned with CamScanner

جدول IV : توزیع مربع-کای



n	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.115	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Subject:

date:

توزیع t :

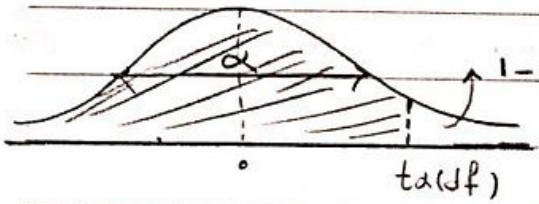
می دانیم Z یک متغیر تصادفی نرمال استاندارد $Z \sim N(0,1)$ و Y یک متغیر تصادفی توان دوم کای با درجه آزادی df می باشد $Y \sim \chi^2_{(df)}$، همچنین می دانیم که این دو متغیر یعنی Z و Y مستقل از هم هستند، در این صورت رابطه $T = \frac{Z}{\sqrt{Y/df}}$، متغیر تصادفی جدیدی به نام t را معرفی می کند که توزیع t - استیودنت با درجه آزادی df نیز نامیده می شود و با این نماد نشان می دهند:

$T \sim t_{\alpha(df)}$.

تابع چگالی احتمال متغیر تصادفی t با درجه آزادی df براساس این رابطه بدست می آید:

$$f_{(t)} = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\Gamma\left(\frac{df}{2}\right)\sqrt{\pi df}} \left(1+\frac{t^2}{df}\right)^{-\frac{df+1}{2}} \qquad -\infty < t < +\infty$$

منحنی تابع چگالی احتمال توزیع t به این صورت می باشد ⇐



توزیع t از جهاتی با توزیع نرمال استاندارد که شباهت دارد:

۱- توزیع t همانند توزیع نرمال استاندارد نسبت به خط وسط ($t=0$) متقارن است.

۲- منحنی توزیع t همانند توزیع نرمال استاندارد زنگوله ای شکل می باشد با این تفاوت که ارتفاع منحنی کمتر و کوتاه تر و دو انتهای منحنی بیشتر می باشد. توزیع t استیودنت یک تابع چگالی احتمال برای توضیح رفتار متغیر و محاسبه احتمال وقوع آن در سطح نمونه می باشد لذا با شدت به حجم نمونه یعنی n وابسته می باشد لذا با تغییر درجه آزادی ناشی از تغییر در n (حجم نمونه) شکل منحنی توزیع نیز تغییر می یابد. هرچه df بزرگتر شود توزیع t به سمت توزیع نرمال استاندارد یا Z میل می کند و در صورتی که حجم نمونه بزرگتر از 30 گردد ($n \geq 30$) منحنی توزیع t استیودنت تقریبا منطبق بر منحنی توزیع Z می گردد.



Scanned with CamScanner

Subject:

date:

برای محاسبه احتمال وقوع (سطح زیر منحنی) متغیر t کافی است به جدول شماره 5 وقت شود. در این جدول مقادیر مختلف $t_\alpha(df)$ به ازای مقادیر مختلف α و df نشان داده شده است. در این جدول، مقادیر احتمال براساس این رابطه بدست می آید

$$Pr(T < t_{\alpha(df)}) = \alpha$$

$t_\alpha(df)$ نقطه‌ای بر روی محور افقی تابع چگالی احتمال t می باشد که مساحت سمت چپ آن برابر α می باشد.

مثال: اگر متغیر تصادفی T دارای توزیع t-استیودنت با $df=24$ نباشد، مقدار t را در حالات زیر تعیین کنید:

الف) $Pr(T < t_\alpha) = 0.95 \Rightarrow t_{0.95(24)} = 1.71$

ب) $Pr(T < t_\alpha) = 0.99 \Rightarrow t_{0.99(24)} = 2.49$

ج) $Pr(T > t_\alpha) = 0.025 = 1 - Pr(T < t_\alpha) = 0.025 \Rightarrow Pr(T < t_\alpha) = 0.975$

$\Rightarrow t_{0.975(24)} = 2.06$

مثال: اگر متغیر تصادفی T دارای توزیع t با $df=10$ باشد، مقدار t را در حالات زیر محاسبه کنید:

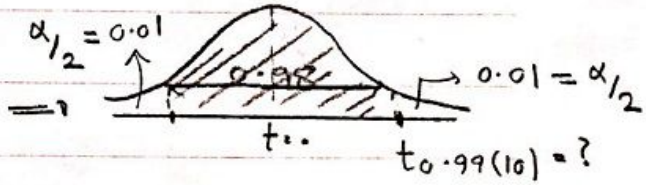
الف) $Pr(T > t) = 0.05$ ب) $Pr(-t < T < t) = 0.98$

ج) $Pr(T < t) = 0.2$ د) $Pr(T > t) = 0.95$

الف) $Pr(T > t) = 0.05 \Rightarrow 1 - Pr(T < t) = 0.05 \Rightarrow Pr(T < t_{(df)}) = 0.95$

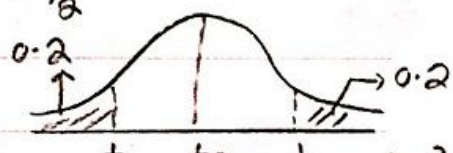
$\Rightarrow t_{0.95(10)} = 1.81$

ب) $Pr(-t < T < t) = 0.98$



$Pr(t_{\alpha/2} < T < t_{1-\alpha/2}) = 0.98 = 1-\alpha \Rightarrow \alpha = 0.02 \quad \alpha/2 = 0.01$

$Pr(T < t) = 0.99 \Rightarrow t_{0.99(10)} = 2.76$



ج) $Pr(T < t) = 0.2 \Rightarrow Pr(T < t_1) = 0.8$

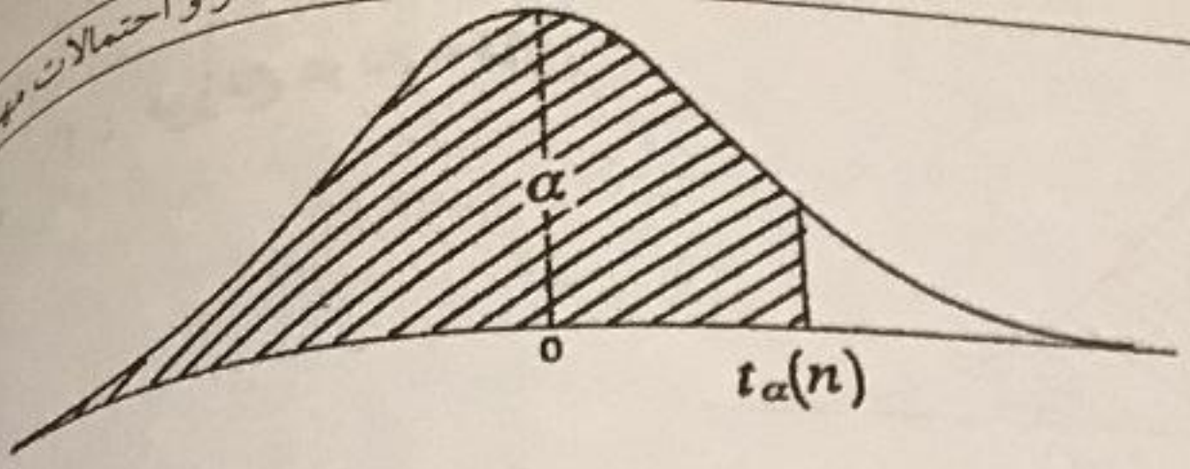
$t_{0.8(10)} = 2.76 \Rightarrow t = -t_1 = -2.76 \Rightarrow t_{0.2(10)} = -1.81$

د) $Pr(T > t) = 0.95 \Rightarrow 1 - Pr(T < t) = 0.95 \Rightarrow Pr(T < t) = 0.05$

$Pr(T < t_1) = 0.95 \Rightarrow t_{0.95(10)} = 1.81 \Rightarrow t = -t_1 \Rightarrow t_{0.05(10)} = -1.81$

Scanned with CamScanner

جدول V : توزیع t



#	$t_{.995}$	$t_{.99}$	$t_{.975}$	$t_{.95}$	$t_{.90}$	$t_{.80}$	$t_{.75}$	$t_{.70}$	$t_{.60}$	$t_{.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
2	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.289	.142
3	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
4	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
5	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
6	3.71	3.14	2.45	1.94	1.44	.906	.718	.553	.265	.131
7	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
8	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
9	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
20	2.84	2.53	2.09	1.72	1.32	.860	.687	.533	.257	.127
21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127
29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
40	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
60	2.66	2.39	2.00	1.67	1.30	.848	.679	.527	.254	.126
120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
∞	2.58	2.33	1.96	1.645	1.28	.842	.674	.524	.253	.126

$$\frac{(n_1+n_2-2)S_P^2}{\sigma^2}=\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2_{(n_1+n_2-2)}$$

و بنابراین با استفاده از تعریف توزیع t به راحتی می‌توان نشان داد که:

$$T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_P\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t_{(n_1+n_2-2)} \qquad (7.6)$$

۶.۶ توزیع نمونه‌ای نسبت واریانس‌های نمونه

دو جامعهٔ مفروض در بخش ۵.۶ را در نظر بگیرید. در این بخش می‌خواهیم توزیع نمونه‌ای نسبت واریانس‌های نمونه یعنی $\frac{S_1^2}{S_2^2}$ را به‌دست آوریم. توزیع نمونه‌ای این نسبت به توزیعی به‌نام توزیع F مربوط می‌شود که آن را در زیر معرفی می‌کنیم.

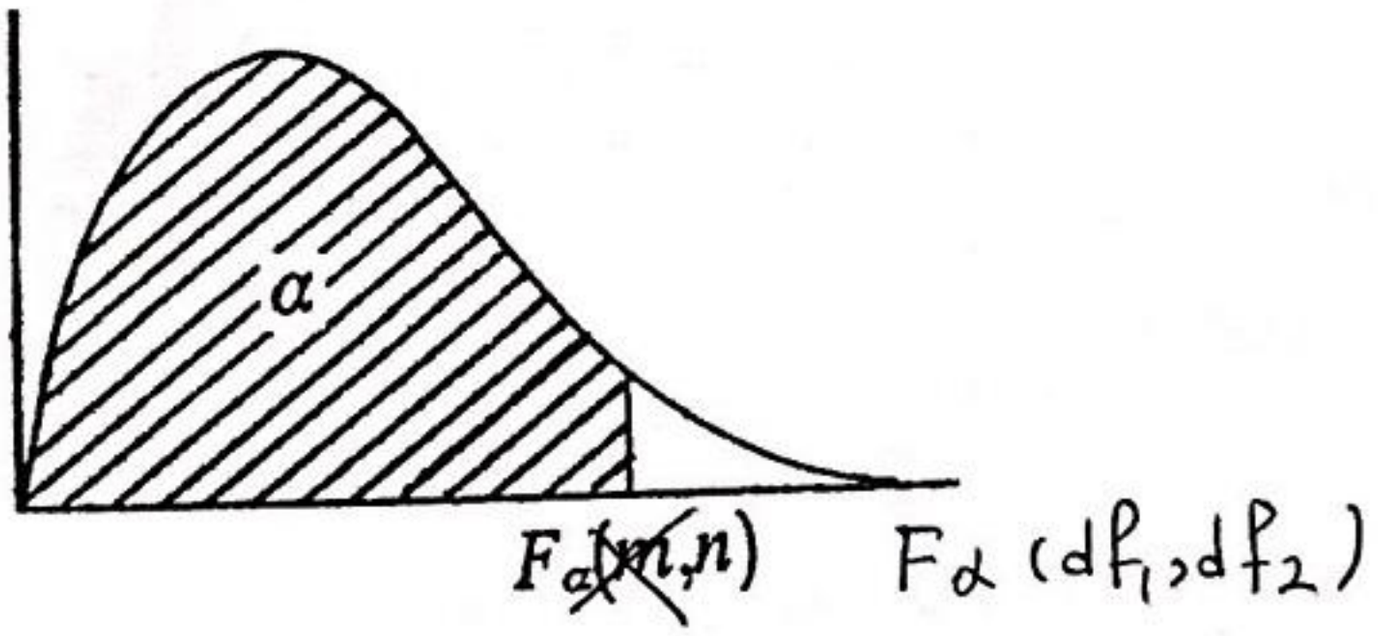
توزیع F[1]

اگر $U\sim\chi^2_{(n)}$ و $V\sim\chi^2_{(m)}$ و متغیرهای تصادفی U و V از یکدیگر مستقل باشند آنگاه توزیع متغیر تصادفی $F=\frac{U/m}{V/n}$ را یک توزیع F با m و n درجه آزادی گویند و آن را با نماد $F\sim F_{(m,n)}$ نمایش می‌دهند.

تابع چگالی احتمال یک توزیع F با درجه‌های آزادی m و n عبارت است از

$$f_F(x)=\frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}(\frac{m}{n})^{\frac{m}{2}}\frac{x^{\frac{m}{2}-1}}{(1+\frac{m}{n}x)^{\frac{m+n}{2}}} \qquad x>0 \qquad (8.6)$$

نمودار تابع چگالی احتمال توزیع F در شکل ۳.۶ رسم شده است.



شکل ۳.۶ نمودار تابع چگالی احتمال توزیع $F_{(m,n)}$

1- F-distribution

Scanned with CamScanner

برای توزیع F جدول (VI) در ضمیمه ارائه گردیده است که در این جدول $F_{\alpha}(m,n)$ به ازای مقادیر مختلف m و n و $\alpha = ۰/۹۰, ۰/۹۵, ۰/۹۹$ محاسبه شده است. که در آن $F_{\alpha}(m,n)$ نقطه‌ای روی محور افقی تابع چگالی احتمال توزیع $F_{(m,n)}$ است که مساحت زیر منحنی قبل از این نقطه برابر α است، برای مثال از جدول (VI) داریم که $F_{۰/۹۰}(۶,۱۴) = ۲/۲۴$ و $F_{۰/۹۵}(۲۰,۱۰) = ۲/۷۷$.

مثال ۱.۶.۶ اگر $F \sim F_{(۱۰,۱۲)}$ مطلوب است:

الف- $P(F < ۲/۷۵)$ را محاسبه کنید.

ب- اگر $P(F > x) = ۰/۰۱$ باشد مقدار x را محاسبه کنید.

حل الف- چون $F_{۰/۹۵}(۱۰,۱۲) = ۲/۷۵$ پس $P(F < ۲/۷۵) = ۰/۹۵$.

ب- $$P(F > x) = ۰/۰۱ \Rightarrow P(F \leq x) = ۰/۹۹ \Rightarrow x = F_{۰/۹۹}(۱۰,۱۲) = ۴/۳۰$$

در قضیه زیر توزیع نمونه‌ای $\frac{S_۱^۲}{S_۲^۲}$ را به دست می‌آوریم.

قضیه ۴.۶ اگر $S_۱^۲$ و $S_۲^۲$ به ترتیب واریانس‌های نمونه‌های تصادفی مستقل به اندازه‌های $n_۱$ و $n_۲$ از جامعه‌های نرمال با واریانس‌های $\sigma_۱^۲$ و $\sigma_۲^۲$ باشند، آنگاه:

$$F = \frac{S_۱^۲/\sigma_۱^۲}{S_۲^۲/\sigma_۲^۲} \sim F_{(n_۱-۱, n_۲-۱)} \qquad (۹.۶)$$

اثبات از قضیه ۲.۶ نتیجه می‌شود که:

$$U = \frac{(n_۱-۱)S_۱^۲}{\sigma_۱^۲} \sim \chi^۲_{(n_۱-۱)} \quad , \quad V = \frac{(n_۲-۱)S_۲^۲}{\sigma_۲^۲} \sim \chi^۲_{(n_۲-۱)}$$

و U و V از یکدیگر مستقل هستند. بنابراین از تعریف توزیع F داریم که:

$$F = \frac{U/(n_۱-۱)}{V/(n_۲-۱)} \sim F_{(n_۱-۱, n_۲-۱)}$$

از طرفی:

$$F = \frac{(n_۱-۱)S_۱^۲/[\sigma_۱^۲(n_۱-۱)]}{(n_۲-۱)S_۲^۲/[\sigma_۲^۲(n_۲-۱)]} = \frac{S_۱^۲/\sigma_۱^۲}{S_۲^۲/\sigma_۲^۲}$$

بنابراین نتیجه قضیه حاصل می‌شود.

مثال ۲.۶.۶ اگر $S_۱^۲$ و $S_۲^۲$ نمایانگر واریانس‌های نمونه‌های تصادفی مستقل به اندازه‌های ۲۵ و

Scanned with CamScanner

۳۱ تایی به ترتیب از جامعه‌های نرمال با واریانس‌های ۱۰ و ۱۵ باشند، احتمال اینکه S_1^2 حداقل ۱/۲۶ برابر S_2^2 باشد چقدر است؟

حل

$$P\left(\frac{S_1^2}{S_2^2} \geq 1/26\right) = P\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \geq 1/26\left(\frac{15}{10}\right)\right) = P(F_{(24,30)} \geq 1/89)$$

$$= 1 - P(F_{(24,30)} < 1/89) = 1 - 0/95 = 0/05$$

۷.۶ تمرینات

۱ جعبه‌ای حاوی ۵ مهره است که از ۱ تا ۵ شماره‌گذاری شده‌اند. از این جعبه نمونه‌ای مرکب از ۲ مهره به تصادف و با جای‌گذاری استخراج می‌شود.

الف- با تشکیل جدول توزیع احتمال توأم X_1 و X_2، توزیع $\overline{X} = \frac{X_1 + X_2}{2}$ را محاسبه کنید.

ب- میانگین و واریانس $\overline{X}$ را از روی توزیع آن محاسبه کنید.

ج- میانگین و واریانس جامعه را پیدا کنید.

د- نمودار توزیع جامعه (برای یک مشاهده X) و نمودار توزیع نمونه‌ای $\overline{X}$ را رسم کنید.

۲ جعبه‌ای پر از مهره‌های هم‌شکل است که یک‌سوم آن‌ها با شماره ۲ و یک‌سوم با شماره ۴ و یک‌سوم با شماره ۶ مشخص شده‌اند.

الف- یک مهره استخراج می‌گردد اگر شماره آن X باشد، میانگین μ و انحراف معیار σ جامعه را بیابید.

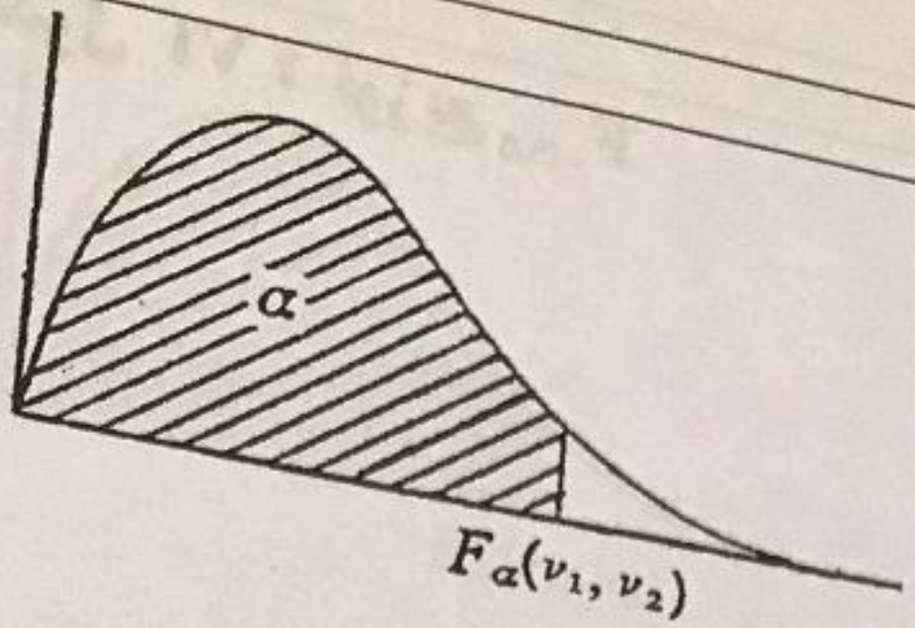
ب- نمونه‌ای مرکب از دو مهره با جای‌گذاری استخراج می‌شود. اگر $\overline{X}$ میانگین نمونه باشد، جدول توزیع احتمال $\overline{X}$ را به‌دست آورید و از روی آن میانگین و واریانس $\overline{X}$ را حساب کنید.

ج- قسمت (ب) را برای نمونه‌ای متشکل از ۳ مهره حل کنید.

۳ اندازه‌گیری قطر گام نوعی پیچ دارای توزیع نرمال با میانگین ۰/۴۰۰۸ اینچ و انحراف معیار ۰/۰۰۰۳ اینچ است. حدود مشخصات طراحی به صورت ۰/۴±۰/۰۰۱ اینچ ارائه شده است.

Scanned with CamScanner

ضمیمه جدول VI : توزیع $F_{0/95}$

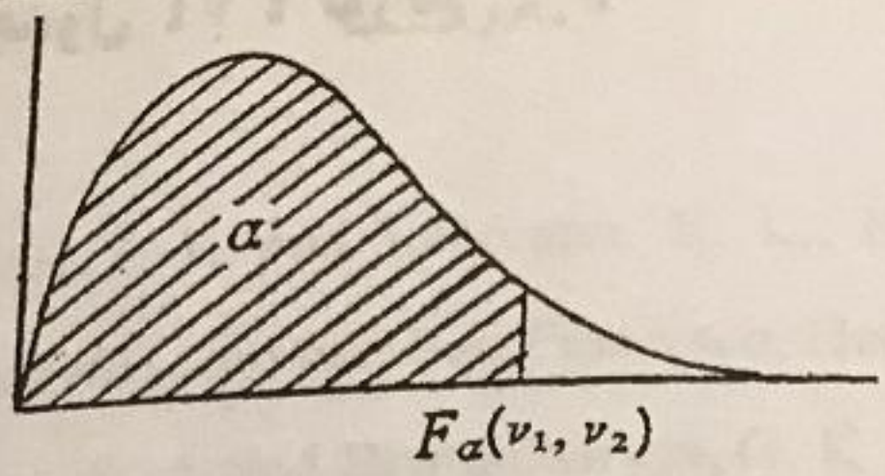


ν_2 \ ν_1	NUMERATOR DEGREES OF FREEDOM								
Denominator Degrees of Freedom	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.77
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

ادامه جدول VI : توزیع $F_{0/95}$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
	NUMERATOR DEGREES OF FREEDOM									
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Denominator Degrees of Freedom



NUMERATOR DEGREES OF FREEDOM

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	4,052	4,999.5	5,403	5.625	5,764	5,859	5,928	5,982	6,022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Denominator Degrees of Freedom

ν_2 \ ν_1	10
1	241.9
2	19.40
3	8.79
4	5.96
5	4.74
6	4.0
7	3.6
8	3.3
9	3.
10	2
11	2
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
40	
60	
120	
∞	

Denominator Degrees of Freedom

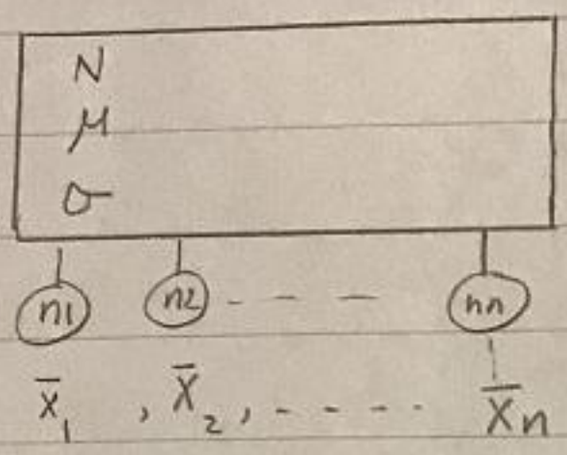
ادامه جدول VI : توزیع $F_{0.99}$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
	NUMERATOR DEGREES OF FREEDOM									
1	6,056	6,106	6,157	6.209	6.235	6,261	6.287	6,313	6,339	6,366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.0

Denominator Degrees of Freedom

توزیع میانگین نمونه‌ها

اگر از یک جامعه آماری مرکب از N عضو با میانگین μ و انحراف معیار σ همه نمونه‌ای n تایی ممکن را بطور تصادفی و بدون جایگزینی استخراج کرده و تغیر x را در آن‌ها اندازه گیری کنیم، می‌توان گفت که سری میانگین‌های نمونه‌های استخراجی یعنی $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ یک تغیر تصادفی جدید به نام $\bar{X}$ را تشکیل می‌دهند که از توزیع میانگین نمونه‌ها تبعیت می‌کند.



میانگین و واریانس توزیع میانگین نمونه‌ها (به ترتیب با $\mu_{\bar{x}}$ و $\sigma^2_{\bar{x}}$ نشان می‌دهند و بصورت ذیل محاسبه می‌گردد.

میانگین توزیع میانگین ← $\mu_{\bar{x}} = E(\bar{x}) = \sum_{i=1}^{n} \bar{x}_i \cdot Pr(\bar{X} = \bar{x}_i) = \mu$ ← میانگین جامعه اصلی

$$\mu_{\bar{x}} = E(\bar{x}) = \sum_{i=1}^{n} \bar{x}_i \cdot f(\bar{x}_i)$$

$$Var(\bar{x}) = \sigma^2_{\bar{x}} = E(\bar{x}^2) - E^2(\bar{x}) \xrightarrow{\text{بدان باشد}} E(\bar{x}^2) = \sum_{i=1}^{n} \bar{x}_i^2 \cdot P(\bar{x}_i)$$

انحراف از معیار جامعه ← $\sigma_{\bar{x}} = \sigma / \sqrt{n}$ ← انحراف از معیار توزیع میانگین

$\sigma^2_{\bar{x}} = \sigma^2 / n$ ← واریانس جامعه ← حجم نمونه

واریانس توزیع میانگین

البته در صورتی که نمونه‌گیری از یک جامعه محدود و یا متناهی باشد یا حجم نمونه بزرگتر از ۱۰٪ حجم جامعه باشد واریانس توزیع میانگین براساس این رابطه برآورد می‌شود:

$$\sigma^2_{\bar{x}} = \sigma^2 / n \left(\frac{N-n}{N-1} \right) \quad , \quad \sigma_{\bar{x}} = \sigma / \sqrt{n} \sqrt{\frac{N-n}{N-1}}$$

متغیر نرمال استاندارد شده یا Z هنگامیکه توزیع میانگین نمونه‌ها به سمت توزیع نرمال میل کند براساس رابطه ذیل برآورد می‌شود:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad \Rightarrow \quad \bar{x} \sim N(\mu, \sigma / \sqrt{n})$$

STAEDTLER®

● البته محاسبه احتمال وقوع متغیر تصادفی $\bar{X}$ با استفاده از توزیع احتمال نرمال و بکارگیری جدول Z بستگی به تعریف و قضیه حد مرکزی دارد. براساس این قضیه بیان داریم، توزیع میانگین نمونه‌های زمانی به سمت توزیع نرمال میل می‌کند یا توزیع جامعه اصلی نرمال باشد (بدون توجه به حجم نمونه: $n>30$ و یا $n<30$) و یا اگر توزیع جامعه اصلی نرمال نباشد به ازای $n>30$، توزیع میانگین نمونه‌ها تقریباً از توزیع نرمال تبعیت می‌کند بنابراین برای محاسبه احتمال وقوع متغیر تصادفی $\bar{X}$ از جدول Z استفاده می‌شود.

سؤال اساسی که در این مبحث مطرح می‌شود این است که اگر توزیع جامعه اصلی نرمال نباشد و حجم نمونه $n<30$ باشد برای

● محاسبه احتمال وقوع متغیر $\bar{X}$ چکار باید کرد. در ادامه درس با توزیع t-استیودنت آشنا شده و جواب سؤال خود را خواهید فهمید.

مثال: فرض کنید متوسط کیفیت درس با میانگین 15 و انحراف استاندارد 3 از توزیع نرمال تبعیت کند. احتمال اینکه در یک نمونه تصادفی و تصادفی انتخاب شده میانگین نمره حداقل 18 باشد چقدر است؟

$$P_r(\bar{X}\geq 18) \Rightarrow P_r\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{18-15}{3/\sqrt{9}}\right) = P_r(Z\geq 3) = 1-P_r(Z\leq 3)$$

$$P_r(Z\leq 3) = 1-0.9987 = 0.0013$$

● مثال: جامعه متغیر X دارای توزیع نرمال با $\mu=10$ و $\sigma^2=36$ باشد و از این جامعه نمونه تصادفی متشکل از 25 عضو انتخاب شود، مطلوبست:

الف) $P_r(\bar{X}\geq 12) = P_r\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{12-10}{6/\sqrt{25}}\right) = P_r(Z\geq 1.66) = 1-P_r(Z\leq 1.66)$

$$1-P_r(Z\leq 1.66) = 1-0.9515 = 0.0485$$

ب) $P_r(8\leq \bar{x}\leq 12) = P_r\left(\frac{8-10}{6/\sqrt{25}} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{12-10}{6/\sqrt{25}}\right) = P_r(-1.66\leq Z\leq 1.66)$

$$P_r(Z\leq 1.66) - P_r(Z\leq -1.66) = 0.9515 - 0.0485 = 0.903$$

STAEDTLER®

مثال: نمرات دانشجویان در درس آمار دارای توزیع نرمال با میانگین 80.83 و انحراف از معیار 4.08 اگر:

الف) هرگاه یک دانشجو بطور تصادفی انتخاب شود احتمال اینکه نمره آن بین 80 و 81.25 باشد چقدر؟

ب) هرگاه 100 دانشجو بطور تصادفی انتخاب شود احتمال اینکه میانگین نمرات آنان بین 80 و 81.25 باشد چقدر؟

الف) $P(80 < x < 81.25) = P(\frac{80 - 80.83}{4.08} < \frac{x-\mu}{\sigma} < \frac{81.25 - 80.83}{4.08})$

$P(-0.2 < Z < 0.1) = P(Z < 0.1) - P(Z < -0.2) = 0.5398 - 0.4207 = 0.1191$

ب) $P(80 < \bar{x} < 81.25) = P(\frac{80 - 80.83}{\frac{4.08}{10}} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{81.25 - 80.83}{\frac{4.08}{10}})$

$P(-2.03 < Z < 1.02) = P(Z < 1.02) - P(Z < -2.03) = 0.8461 - 0.0212 = 0.8249$

توزیع t - استیودنت

می دانیم که نسبت انحراف متغیر تصادفی میانگین ($\bar{X}$) از امید ریاضی خود به خطای معیار آن در صورتیکه توزیع جامعه اصلی نرمال باشد برابر است با متغیر نرمال استاندارد شده یا Z

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

اگر محاسبه مقدار σ^2 امکان پذیر نباشد به جای آن باید از S^2 یا واریانس نمونه استفاده کرد. مقادیر حاصل از این جایگزاری دیگر از توزیع نرمال پیروی نخواهد کرد (البته در صورتی که حجم نمونه $n < 30$ باشد) یک از توزیع احتمال t - استیودنت پیروی خواهد کرد. در چنین صورتی متغیر تصادفی t براساس این رابطه بدست می آید:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \quad \begin{matrix} \sigma^2 \text{ نامعلوم} \\ n < 30 \end{matrix}$$

منحنی توزیع t استیودنت هم نیز منحنی نرمال زنگوله‌ای شکل می باشد و البته متقارن ولی پراکندگی در توزیع t بیشتر از توزیع نرمال می باشد.

۴.۹. آماره $\bar{X}_1 - \bar{X}_2$

در بسیاری از موارد لازم است که میانگین‌های دو جامعه با یکدیگر مقایسه شوند. دو جامعه با میانگین‌های μ_1 و μ_2 و واریانس‌های σ_1^2 و σ_2^2 را درنظر بگیرید. از جامعه اول یک نمونه تصادفی با حجم n_1 انتخاب کرده و میانگین نمونه را با $\bar{x}_1$ نشان می‌دهیم. به‌همین ترتیب از جامعه دوم یک نمونه تصادفی با حجم n_2 و مستقل از نمونه اول انتخاب نموده و میانگین نمونه را $\bar{x}_2$ می‌نامیم. حال آماره‌ای به‌صورت $\bar{X}_1 - \bar{X}_2$ تعریف می‌کنیم. این آماره را آماره **تفاضل میانگین** می‌نامیم. قضیه زیر توزیع این آماره را مشخص می‌کند. ($\bar{x}_1 - \bar{x}_2$ یک مقدار از آماره $\bar{X}_1 - \bar{X}_2$ است.)

قضیه ۷.۹. اگر نمونه‌های تصادفی مستقل با حجم‌های n_1 و n_2 از دو جامعه با میانگین‌های μ_1 و μ_2 و واریانس‌های σ_1^2 و σ_2^2 انتخاب شوند، توزیع نمونه تفاضل میانگین‌ها (توزیع $\bar{X}_1 - \bar{X}_2$) تقریباً توزیع نرمال با میانگین $\mu_{X_1 - X_2} = \mu_1 - \mu_2$ و واریانس $\sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ خواهد شد. به‌عبارت دیگر متغیر

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

دارای توزیع نرمال استاندارد است. ○

توجه: در قضیه بالا، اگر n_1 و $n_2 \geqslant ۳۰$ باشند، $\bar{X}_1 - \bar{X}_2$ بدون توجه به نوع توزیع دو جامعه، دارای توزیع نرمال خواهد شد و اگر n_1 و $n_2 < ۳۰$ باشند، در صورتی‌که توزیع دو جامعه نرمال باشند، توزیع $\bar{X}_1 - \bar{X}_2$ نیز نرمال خواهد شد.

مثال ۶.۹. یک نمونه تصادفی ۲۵ تایی از یک جامعه نرمال با میانگین ۸۰ و انحراف معیار ۵ و از جامعه نرمال دیگری با میانگین ۷۵ و انحراف معیار ۳ یک نمونه تصادفی ۳۶تایی انتخاب می‌کنیم. اگر $\bar{X}_1$ و $\bar{X}_2$ به‌ترتیب میانگین‌های نمونه جامعه اول و جامعه دوم باشند، مطلوب است محاسبه:

$$P(۳/۴ < \bar{X}_1 - \bar{X}_2 < ۸/۹)$$

در صورتی‌که: **الف)** میانگین‌ها تا هر میزان دقتی اندازه‌گیری شده باشند.

ب) میانگین‌ها تا یک‌دهم تقریب اندازه‌گیری شده باشند.

حل) با توجه به اطلاعات داده شده داریم:

جامعه اول
$\mu_1 = ۸۰$
$\sigma_1 = ۵$
$n_1 = ۲۵$

جامعه دوم
$\mu_2 = ۷۵$
$\sigma_2 = ۳$
$n_2 = ۳۶$

الف) با توجه به قضیه فوق $\bar{X}_1 - \bar{X}_2$ دارای توزیع نرمال با میانگین $\mu_{\bar{X}_1-\bar{X}_2} = ۸۰ - ۷۵ = ۵$ و واریانس $\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{۲۵}{۲۵} + \frac{۹}{۳۶} = ۱ + \frac{۱}{۴}$ است، بنابراین

$$P(۳/۴ < \bar{X}_1 - \bar{X}_2 < ۸/۹) = P\left(\frac{۳/۴ - ۵}{\sqrt{۱ + \frac{۱}{۴}}} < Z < \frac{۸/۹ - ۵}{\sqrt{۱ + \frac{۱}{۴}}}\right)$$

$$= P(-۱/۴۳ < Z < ۳/۴۸) = P(Z < ۳/۴۸) - P(Z < -۱/۴۳)$$

$$= ۱ - ۰/۰۷۶۴ = ۰/۹۲۳۶$$

ب) چون میانگین‌ها تا یک‌دهم تقریب محاسبه شده‌اند، لذا باید احتمال زیر را محاسبه کرد.

$$P(۳/۴۵ < \bar{X}_1 - \bar{X}_2 < ۸/۸۵) = P\left(\frac{۳/۴۵ - ۵}{\sqrt{۱ + \frac{۱}{۴}}} < Z < \frac{۸/۸۵ - ۵}{\sqrt{۱ + \frac{۱}{۴}}}\right)$$

$$= P(-۱/۳۸ < Z < ۳/۴۴) = P(Z < ۳/۴۴) - P(Z < -۱/۳۸)$$

$$= ۱ - ۰/۰۸۳۸ = ۰/۹۱۶۲$$

□

مثال ۷.۹. میانگین نمرات تست هوش دانشجویان سال اول یک دانشکده ۵۴۰ و انحراف معیار آن ۵۰ است. دو نمونه تصادفی با حجم $n_1 = ۳۲$ و $n_2 = ۵۰$ انتخاب می‌کنیم، احتمال این که تفاضل میانگین نمرات این دو نمونه

الف) بیش از ۲۰ باشد **ب)** بین ۵ و ۱۰ باشد

چقدر است.

(فرض کنید میانگین‌ها تا هر میزان دقتی می‌توانند اندازه‌گیری شوند.)

حل: اگر میانگین نمونه، نمرات نمونه‌های ۳۲ نفری را با $\bar{X}_1$ و میانگین نمونه، نمونه‌های ۵۰ نفری را با $\bar{X}_2$ نشان دهیم، باید احتمال‌های زیر را محاسبه کنیم.

$$P(|\bar{X}_1 - \bar{X}_2| > ۲۰) \quad \text{و} \quad P(۵ < |\bar{X}_1 - \bar{X}_2| < ۱۰)$$

با توجه به قضیه ۷.۹. $\bar{X}_1 - \bar{X}_2$ دارای توزیع نرمال با میانگین $\mu_{\bar{X}_1-\bar{X}_2} = ۵۴۰ - ۵۴۰ = ۰$ و واریانس $\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}$ است، بنابراین

$$P(|\bar{X}_۱ - \bar{X}_۲| > ۲۰) = ۱ - P(-۲۰ < \bar{X}_۱ - \bar{X}_۲ < ۲۰)$$

$$= ۱ - P\left(\frac{-۲۰ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}} < Z < \frac{۲۰ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}}\right)$$

$$= ۱ - P(Z < ۱٫۷۷) + P(Z < -۱٫۷۷)$$

$$= ۲P(Z < -۱٫۷۷) = ۲ \times ۰٫۰۳۸۴ = ۰٫۰۷۶۸$$

و به‌همین ترتیب

$$P(۵ < |\bar{X}_۱ - \bar{X}_۲| < ۱۰) = P(|\bar{X}_۱ - \bar{X}_۲| < ۱۰) - P(|\bar{X}_۱ - \bar{X}_۲| < ۵)$$

$$= P(-۱۰ < \bar{X}_۱ - \bar{X}_۲ < ۱۰) - P(-۵ < \bar{X}_۱ - \bar{X}_۲ < ۵)$$

$$= P\left(\frac{-۱۰ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}} < Z < \frac{۱۰ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}}\right) - P\left(\frac{-۵ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}} < Z < \frac{۵ - ۰}{\sqrt{\frac{۲۵۰۰}{۳۲} + \frac{۲۵۰۰}{۵۰}}}\right)$$

$$= P(-۰٫۸۸ < Z < ۰٫۸۸) - P(-۰٫۴۴ < Z < ۰٫۴۴) = ۰٫۲۸۱۲$$

□