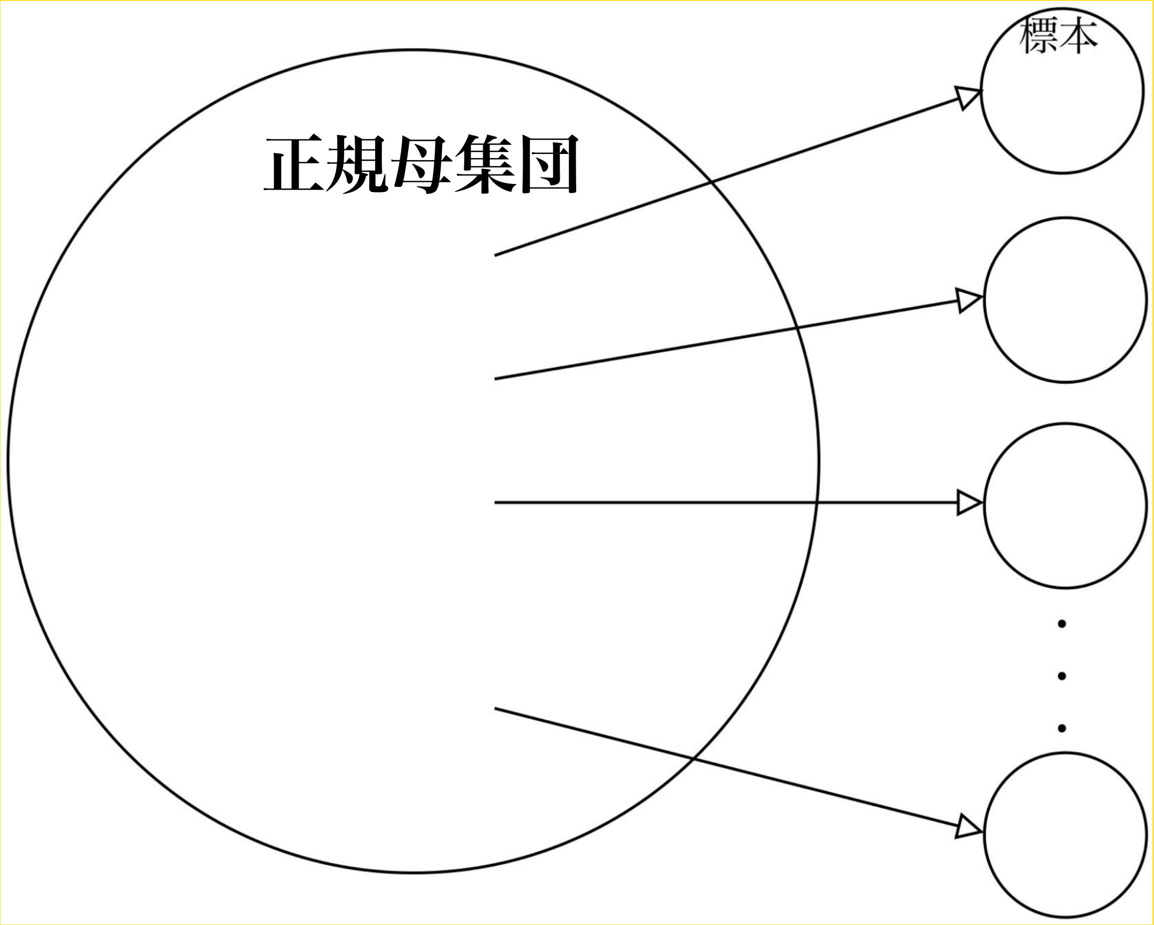
標本分布

静岡大学工学部機械工学科
岡本正芳

母集団と標本

- 標本統計量（標本平均、標本分散など）またはそれから構成される量は決まった量とはならず、確率に従って変化する。つまり、確率変数となる。
- 母集団の特質が決まると標本が従う確率分布（標本分布）も一意に決まる。

標本分布



左図のように正規母集団からその一部の標本を無作為に取り出す。この標本の集合はそれ自体が何らかの確率分布に従う。この確率分布は標本分布と呼ばれる。ここでは代表的な３つの代表的な標本分布であるχ^2分布、t分布、F分布を説明していく。

χ^2分布の解説

χ^2分布. 1

【定義】標準正規分布に従うn個の独立な確率変数を$u_1, u_2, \cdots, u_n$とするとき、それらの２乗和

$$x = u_1^2 + u_2^2 + \cdots u_n^2 = \sum_{i=1}^{n} u_i^2$$

を新たな確率変数とする標本分布を、自由度nのχ^2分布と呼ぶ。その確率密度関数は

$$f(x) = \frac{1}{2\Gamma(n/2)} e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{n}{2}-1}$$

で与えられる。

赤い部分は証明が必要

χ^2分布. 2

《証明》

xの従う確率分布の特性関数は以下で与えられる。

$$\tilde{f}(\xi) = E\left[e^{i\xi x}\right] = E\left[e^{i\xi\left(u_1^2 + u_2^2 + \cdots + u_n^2\right)}\right] = E\left[e^{i\xi u_1^2}\right] \times \cdots\cdots \times E\left[e^{i\xi u_n^2}\right]$$

となり、それぞれの確率変数u_iは独立で同一の標準正規分布なので

$$\tilde{f}(\xi) = \left\{E\left[e^{i\xi u_i^2}\right]\right\}^n$$

となり１つの確率変数の期待値を求めn乗すればよい。

χ^2分布. 3

《証明のつづき》

１つの確率変数の期待値

$$E\left[e^{i\xi u_i^2}\right]=\int_{-\infty}^{\infty}du_i e^{i\xi u_i^2}f\left(u_i\right)=\int_{-\infty}^{\infty}du_i e^{i\xi u_i^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u_i^2}$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}du_i e^{-\frac{1}{2}(1-2i\xi)u_i^2}$$

変数変換　$y=\sqrt{1-2i\xi}u_i$　$dy=\sqrt{1-2i\xi}du_i$

$$=\frac{1}{\sqrt{2\pi}\sqrt{1-2i\xi}}\int_{-\infty}^{\infty}dye^{-\frac{1}{2}y^2}=\frac{\sqrt{2\pi}}{\sqrt{2\pi}\sqrt{1-2i\xi}}=\left(1-2i\xi\right)^{-\frac{1}{2}}$$

χ^2分布. 4

《証明のつづき》

χ^2分布の特性関数

$$\tilde{f}(\xi) = (1 - 2i\xi)^{-\frac{n}{2}}$$

逆フーリエ変換

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\xi e^{-i\xi x}\tilde{f}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\xi e^{-i\xi x}(1 - 2i\xi)^{-\frac{n}{2}}$$

$$= \frac{1}{2\Gamma(n/2)} e^{-\frac{x}{2}}\left(\frac{x}{2}\right)^{\frac{n}{2}-1}$$

Q.E.D.

χ^2分布.5

〔証明で利用した公式〕

$$\int_{-\infty}^{\infty} dx e^{i\xi x} \frac{1}{2\Gamma(n/2)} e^{-\frac{x}{2}} \left(\frac{x}{2}\right)^{\frac{n}{2}-1} = \frac{1}{2\Gamma(n/2)} \int_{-\infty}^{\infty} dx e^{-\frac{x}{2}(1-2i\xi)} \left(\frac{x}{2}\right)^{\frac{n}{2}-1}$$

変数変換　$y = \frac{x}{2}(1-2i\xi)$　$dy = \frac{1-2i\xi}{2} dx$

$$= \frac{(1-2i\xi)^{-\frac{n}{2}}}{\Gamma(n/2)} \int_{-\infty}^{\infty} dy e^{-y} y^{\frac{n}{2}-1} = \frac{(1-2i\xi)^{-\frac{n}{2}}}{\Gamma(n/2)} \Gamma(n/2) = (1-2i\xi)^{-\frac{n}{2}}$$

χ^2分布. 6

〔ガンマ関数〕

$$\Gamma(m)=\int_0^\infty t^{m-1}e^{-t}\,dt$$

代表的な値

$$\Gamma(n+1)=n!,\ \ \Gamma\left(n+\frac{1}{2}\right)=\frac{(2n-1)!!}{2^n}\sqrt{\pi},\Gamma(m+1)=m\Gamma(m)$$

$$I_{2n}(a=2)=\int_{-\infty}^\infty x^{2n}e^{-x^2}\,dx=\int_0^\infty x^{2n}e^{-x^2}\,dx+\int_{-\infty}^0 x^{2n}e^{-x^2}\,dx$$

$$=2\int_0^\infty x^{2n}e^{-x^2}\,dx=\Gamma\left(n+\frac{1}{2}\right)$$

変数変換

$$x^2=t$$

χ^2分布. 7

χ^2分布の例

$$f_{n=1}(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}$$

$$f_{n=2}(x) = \frac{1}{2} e^{-\frac{x}{2}}$$

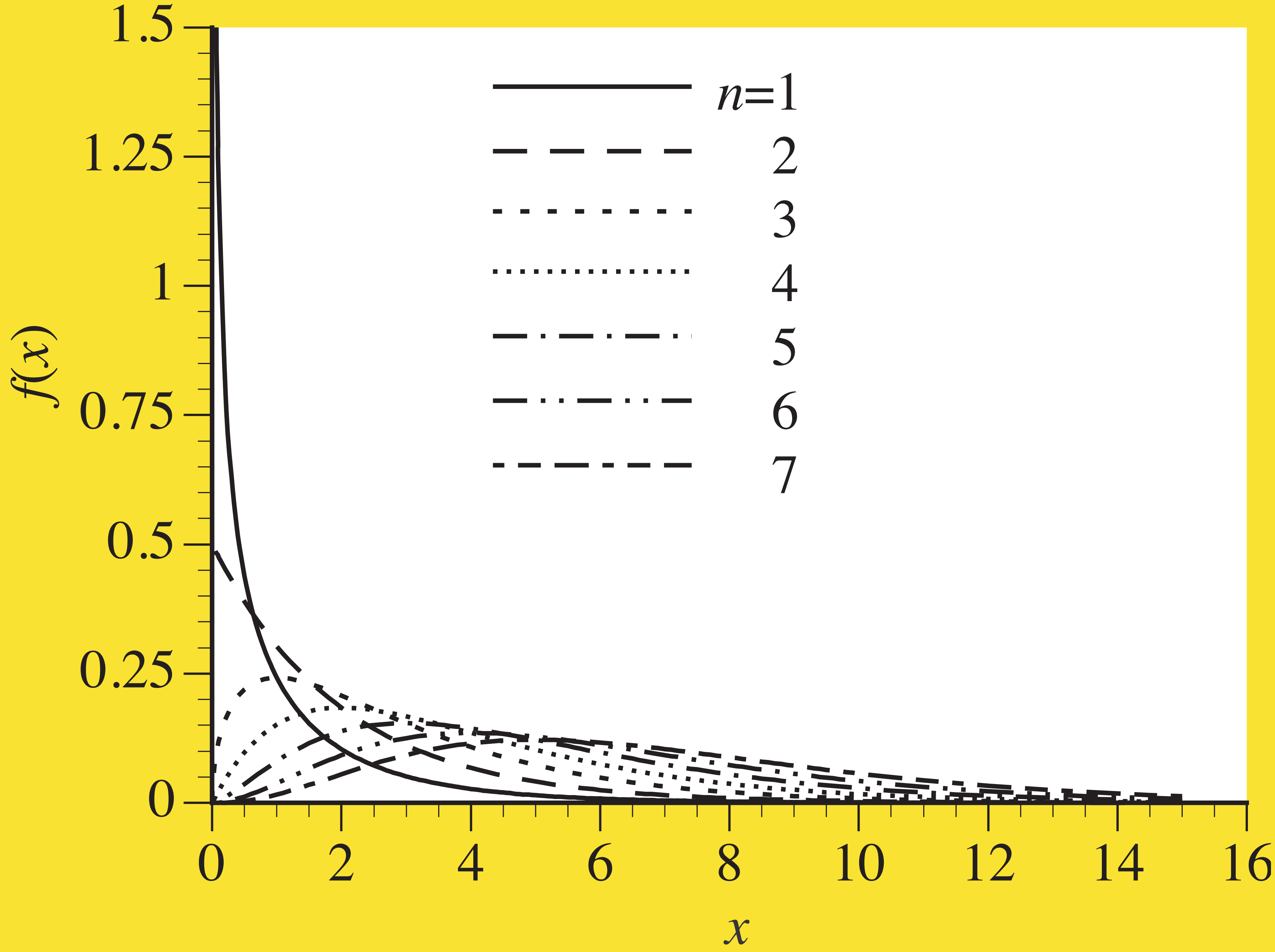
$$f_{n=3}(x) = \frac{1}{\sqrt{2\pi}} x^{\frac{1}{2}} e^{-\frac{x}{2}}$$

$$f_{n=4}(x) = \frac{1}{4} x e^{-\frac{x}{2}}$$

$$f_{n=5}(x) = \frac{1}{3\sqrt{2\pi}} x^{\frac{3}{2}} e^{-\frac{x}{2}}$$

$$f_{n=6}(x) = \frac{1}{16} x^2 e^{-\frac{x}{2}}$$

$$f_{n=7}(x) = \frac{1}{15\sqrt{2\pi}} x^{\frac{5}{2}} e^{-\frac{x}{2}}$$



χ^2分布. 8

χ^2分布のキュムラント展開

$$\tilde{f}(\xi)=(1-2i\xi)^{-\frac{n}{2}}=\exp\left\{-\frac{n}{2}\log(1-2i\xi)\right\}=\exp\left\{\frac{n}{2}\sum_{k=1}^{\infty}\frac{(2i\xi)^k}{k}\right\}$$

対数関数の無限級数展開公式

$$\log(1-x)=-\sum_{k=1}^{\infty}\frac{x^k}{k}$$

各オーダーのキュムラント

$$\kappa_k=2^{k-1}(k-1)!n$$

平均

$$E[x]=\kappa_1=n$$

分散

$$E\left[(x-\kappa_1)^2\right]=\kappa_2=2n$$

χ^2分布. 9

【別バージョンの定義】正規母集団N(μ,σ^2)に従うn個の独立な変数を$x_1, x_2, \cdots, x_n$とするとき、次の量

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

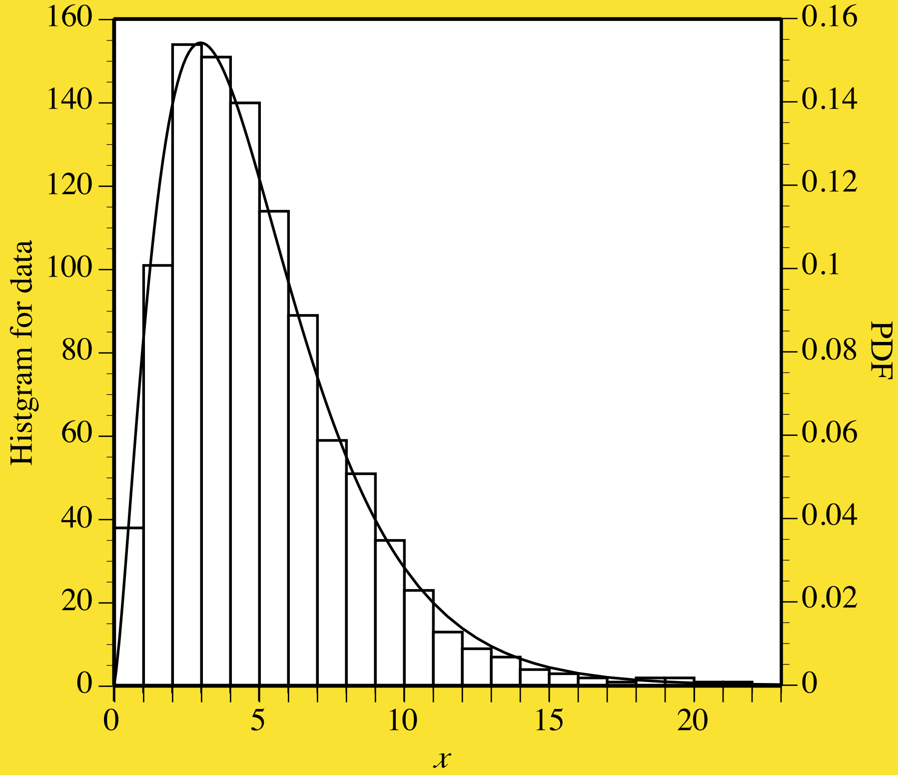
標準化処理

を新たな確率変数とする標本分布を、自由度nのχ^2分布と呼ぶ。

χ^2分布.１０

２０歳男子の身長データーから無作為に５０００人分（平均と分散を合わせた正規乱数）を取り出して、そのデーターを５人を一組として先の処理を施したデーター（１０００個）のヒストグラムと理論PDFを比較した。

$$f_{n=5}(x) = \frac{1}{3\sqrt{2\pi}} x^{\frac{3}{2}} e^{-\frac{x}{2}}$$



χ^2分布. １１

標本を取り扱う際には、多くの場合母平均μは不明である。

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

そこで母平均の代わりに標本平均を利用すると、次の定理が成立する。

χ^2分布.１２

正規母集団の標本$x_1, x_2, \cdots, x_n$が存在し、その２乗和を母分散で規格化したもの

$$\frac{S}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

を新たな確率変数とする標本分布を、自由度$n-1$のχ^2分布に従う。

自由度が減る理由 標本平均を固定したため

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

から１つの標本が他の修正する

χ^2分布.１３

S/σ^2の自由度$n-1$のχ^2分布では、平均と分散は

$$E\left[\frac{S}{\sigma^2}\right]=n-1 \qquad E\left[\left\{\frac{S}{\sigma^2}-(n-1)\right\}^2\right]=2(n-1)$$

となり、標本分散s^2の平均と分散は上式より

$$E\left[s^2\right]=E\left[\frac{S}{n-1}\right]=E\left[\frac{S}{\sigma^2}\frac{\sigma^2}{n-1}\right]=\frac{\sigma^2}{n-1}E\left[\frac{S}{\sigma^2}\right]=\sigma^2$$

このようにχ^2分布は標本分散s^2を評価する標本分布であり、これらの性質から母分散を推測することが可能である。

$$E\Big[\quad\cdots\quad -1)\Big\}^2\Big]$$

$$=\frac{\quad}{(n-1)^2}E\left[\left\{\frac{\quad}{\sigma^2}-(n-1)\right\}^2\right]=\frac{\quad}{n-1}$$

χ^2分布．１４

χ^2分布の加法定理

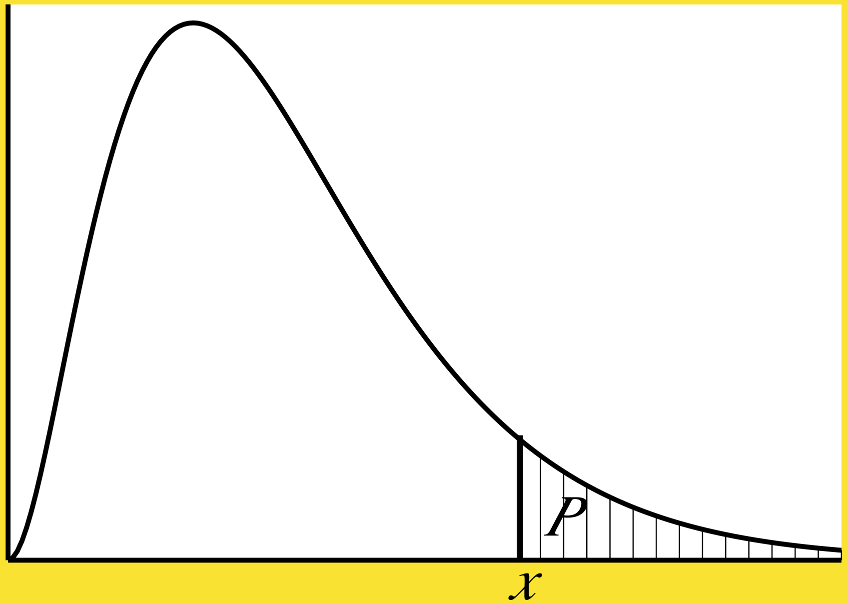
２つの独立なχ^2分布があるとする。一つは自由度m、もう一つは自由度nのとき、それぞれの確率変数がxとyとする。この和$x+y$は自由度$m+n$のχ^2分布に従う。

χ^2分布.１５（分布表）

χ^2分布表　　確率値

自由度

n \ P	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0.000	0.000	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	11.071	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	1.647	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



$$\chi^2(n,p)$$

$$\chi^2(20,0.05)=31.410$$

t分布の解説

t分布. 1

【定義】標準正規分布に従う変数をu、自由度nのχ^2分布に従う変数をvとするとき、変数

$$x = \frac{u}{\sqrt{v / n}}$$

を新たな確率変数とする標本分布を、自由度nのt分布と呼ぶ。その確率密度関数は

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

で与えられる。

赤い部分は証明が必要

t分布.2

《証明》

変数を変換しても同一範囲内における確率（積分値）は変わらないとは

$$f_{y_1,\cdots,y_n}\left(y_1,\cdots,y_n\right)=f_{x_1,\cdots,x_n}\left(x_1\left(y_1,\cdots,y_n\right),\cdots,x_n\left(y_1,\cdots,y_n\right)\right)|J|$$

$$J=\frac{\partial\left(x_1,\cdots,x_n\right)}{\partial\left(y_1,\cdots,y_n\right)}$$

で表現される。以下ではこの変換公式を利用していく。

t分布.3

《証明のつづき》

確率変数u、vの確率密度関数

$$f_{u,v}(u,v)=f_u(u)f_v(v)=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\frac{1}{2\Gamma\left(\frac{n}{2}\right)}e^{-\frac{v}{2}}\left(\frac{v}{2}\right)^{\frac{n}{2}-1}$$

変数変換　$x=\dfrac{u}{\sqrt{v/n}}$　　$y=v$

$$\frac{\partial u}{\partial x}=\frac{1}{\sqrt{n}}\sqrt{y},\ \frac{\partial u}{\partial y}=\frac{1}{2\sqrt{n}}\frac{x}{\sqrt{y}},\ \frac{\partial v}{\partial x}=0,\ \frac{\partial v}{\partial y}=1$$

$$J=\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}-\frac{\partial u}{\partial y}\frac{\partial v}{\partial x}=\frac{1}{\sqrt{n}}\sqrt{y}$$

t分布.4

《証明のつづき》

確率変数x、yの確率密度関数

$$f_{x,y}(x,y)=f_{u,v}(u,v)J=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\frac{1}{2\Gamma\left(\frac{n}{2}\right)}e^{-\frac{v}{2}}\left(\frac{v}{2}\right)^{\frac{n}{2}-1}J$$

$$=\frac{1}{2\sqrt{2\pi}\Gamma\left(\frac{n}{2}\right)}e^{-\frac{x^2}{2n}y-\frac{y}{2}}\left(\frac{y}{2}\right)^{\frac{n}{2}-1}\frac{\sqrt{y}}{\sqrt{n}}=\frac{1}{2^{\frac{n}{2}}\sqrt{2n\pi}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n-1}{2}}e^{-\left(\frac{x^2}{2n}+\frac{1}{2}\right)y}$$

t分布.5

《証明のつづき》

確率変数yに関する積分を実行すると確率変数xに対する確率密度関数が求まる。

$$f(x)=\int_0^\infty dy f_{x,y}(x,y)=\frac{1}{2^{\frac{n}{2}}\sqrt{2n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty dy\, y^{\frac{n-1}{2}}e^{-\left(\frac{x^2}{2n}+\frac{1}{2}\right)y}$$

積分公式 $\int_0^\infty dy\, y^{\alpha}e^{-ay}=\frac{\Gamma(\alpha+1)}{a^{\alpha+1}}$

$$=\frac{1}{2^{\frac{n}{2}}\sqrt{2n\pi}\,\Gamma\left(\frac{n}{2}\right)}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\frac{x^2}{2n}+\frac{1}{2}\right)^{\frac{n+1}{2}}}=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

Q.E.D.

t分布.6

t分布の例

$$f_{n=1}(x)=\frac{1}{\pi}\left(1+x^2\right)^{-1}$$

$$f_{n=2}(x)=\frac{1}{2\sqrt{2}}\left(1+\frac{x^2}{2}\right)^{-\frac{3}{2}}$$

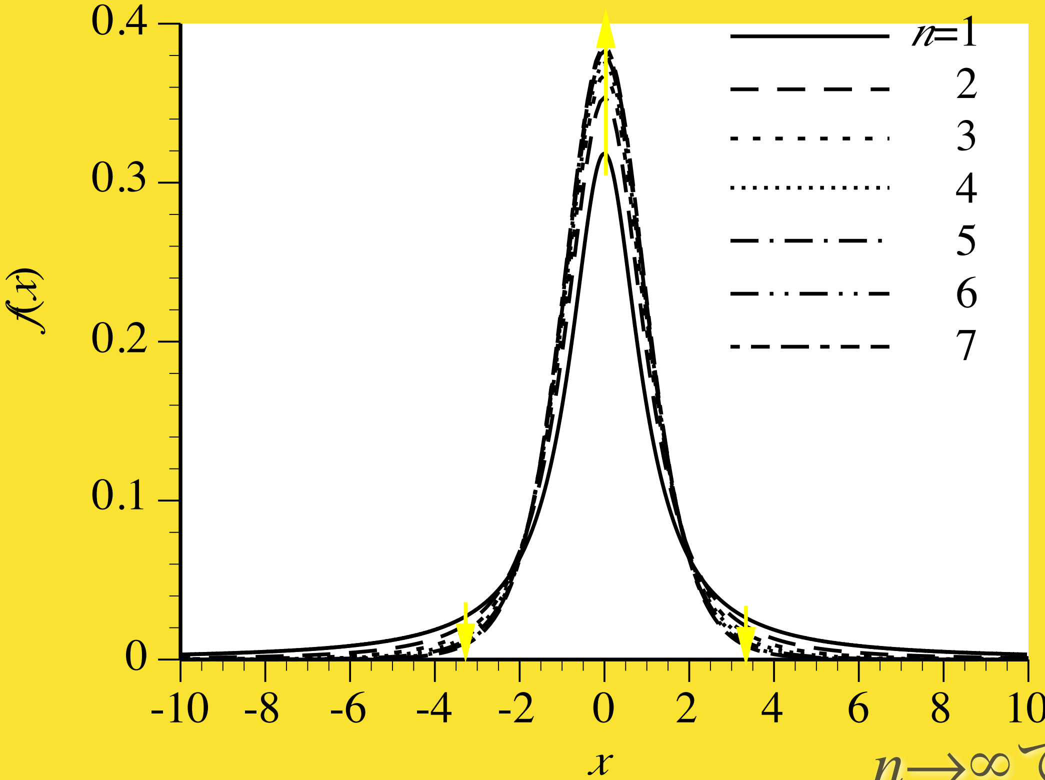
$$f_{n=3}(x)=\frac{2}{\sqrt{3}\pi}\left(1+\frac{x^2}{3}\right)^{-2}$$

$$f_{n=4}(x)=\frac{3}{8}\left(1+\frac{x^2}{4}\right)^{-\frac{5}{2}}$$

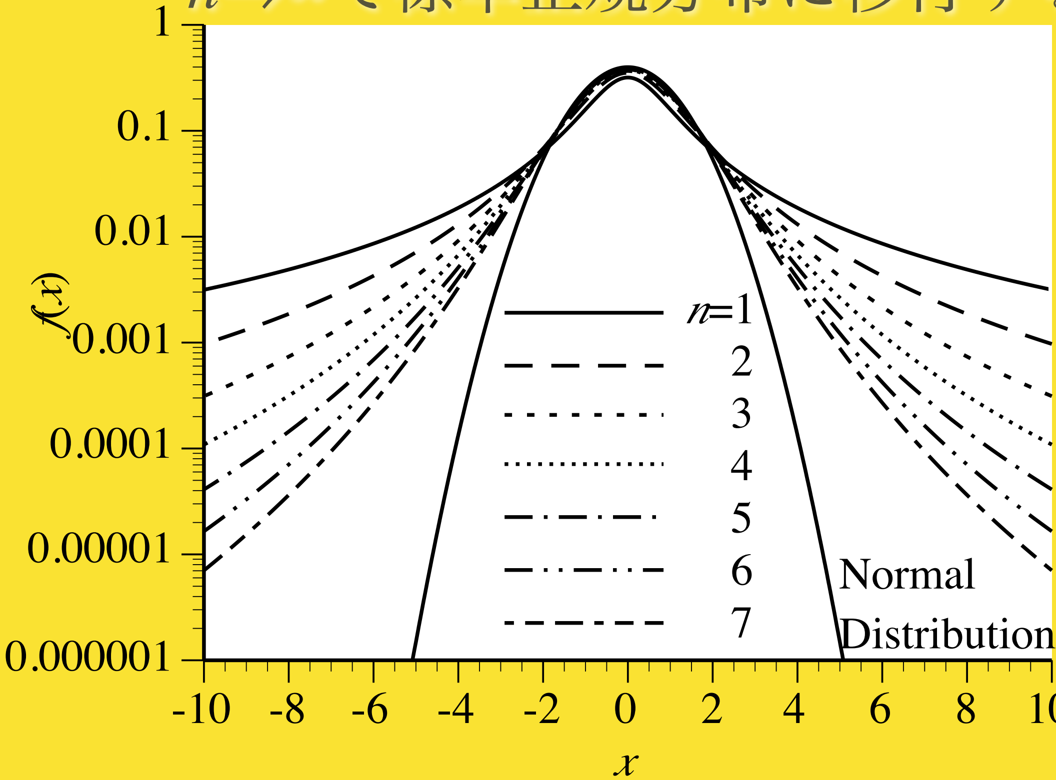
$$f_{n=5}(x)=\frac{8}{3\sqrt{5}\pi}\left(1+\frac{x^2}{5}\right)^{-3}$$

$$f_{n=6}(x)=\frac{15}{16\sqrt{6}}\left(1+\frac{x^2}{6}\right)^{-\frac{7}{2}}$$

$$f_{n=7}(x)=\frac{16}{5\sqrt{7}\pi}\left(1+\frac{x^2}{7}\right)^{-4}$$



$n\to\infty$で標準正規分布に移行する。



t分布.7

自由度無限大

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \overbrace{\frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\left(\frac{n}{2}\right)^{\frac{1}{2}}\Gamma\left(\frac{n}{2}\right)}}^{=1\quad (A.18)} \left\{ \overbrace{\left(1+\frac{1}{\frac{n}{x^2}}\right)^{\frac{n}{x^2}}}^{=e} \right\}^{-\frac{x^2}{2}\left(1+\overbrace{\frac{1}{n}}^{=0}\right)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

標準正規分布となる

t分布.8

平均の導出

$$E[x]=\int_{-\infty}^{\infty}dxxf(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\int_{-\infty}^{\infty}dxx\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

変数変換 $\quad y=\frac{x^2}{n} \qquad dy=\frac{2}{n}xdx$

$$=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\int_{\infty}^{\infty}dy\frac{n}{2}(1+y)^{-\frac{n+1}{2}}=0$$

t分布.9

分散の導出

$$E\left[x^2\right]=\int_{-\infty}^{\infty}dxx^2f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\int_{-\infty}^{\infty}dxx^2\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

$$=\frac{2\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\int_{0}^{\infty}dxx^2\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

変数変換 $\quad y=\dfrac{x^2}{n} \qquad dy=\dfrac{2}{n}xdx$

$$=\frac{n\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}\int_{0}^{\infty}dyy^{\frac{1}{2}}(1+y)^{-\frac{n+1}{2}}$$

t分布.１０

分散の導出のつづき

積分公式 $\int_0^\infty dy y^{\alpha-1}(1+y)^{-\alpha-\beta}=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\quad \alpha,\beta>0$

$$E\left[x^2\right]=\frac{n\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}\frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n}{2}-1\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

$$=\frac{n\frac{1}{2}\sqrt{\pi}\Gamma\left(\frac{n}{2}-1\right)}{\sqrt{\pi}\left(\frac{n}{2}-1\right)\Gamma\left(\frac{n}{2}-1\right)}=\frac{n}{n-2}$$

ただし、$n<3$では分散は発散し求めることができない。

t分布.１１

統計量

$$o_{2k+1} = m_{2k+1} = 0 \qquad o_{2k} = m_{2k} = \frac{n^k \prod_{i=1}^{k} \left(k + \frac{1}{2} - i \right)}{\prod_{i=1}^{k} \left(\frac{n}{2} - i \right)}$$

特性関数

$$\tilde{f}_n(\xi) = \frac{2^{-\frac{n}{2}+1}}{\Gamma\left(\frac{n}{2}\right)} \left| \sqrt{n}\xi \right|^{\frac{n}{2}} K_{n/2}\left(\sqrt{n} |\xi| \right)$$

第２種変形ベッセル関数

t分布.１２

【別バージョンの定義】正規母集団N(μ,σ^2)に従うn個の独立な変数を$x_1, x_2, \cdots, x_n$とするとき、中心極限定理よりこの変数を標本とする標本平均は正規分布N(μ,σ^2/n)に従う。この標本平均を標準化した変数$u=(\bar{x}-\mu)\sqrt{n/\sigma^2}$ と母平均からの差の２乗和を母分散で除した変数$v=\sum_{i=1}^{n}(x_i-\mu)^2/\sigma^2$ から構成される変数

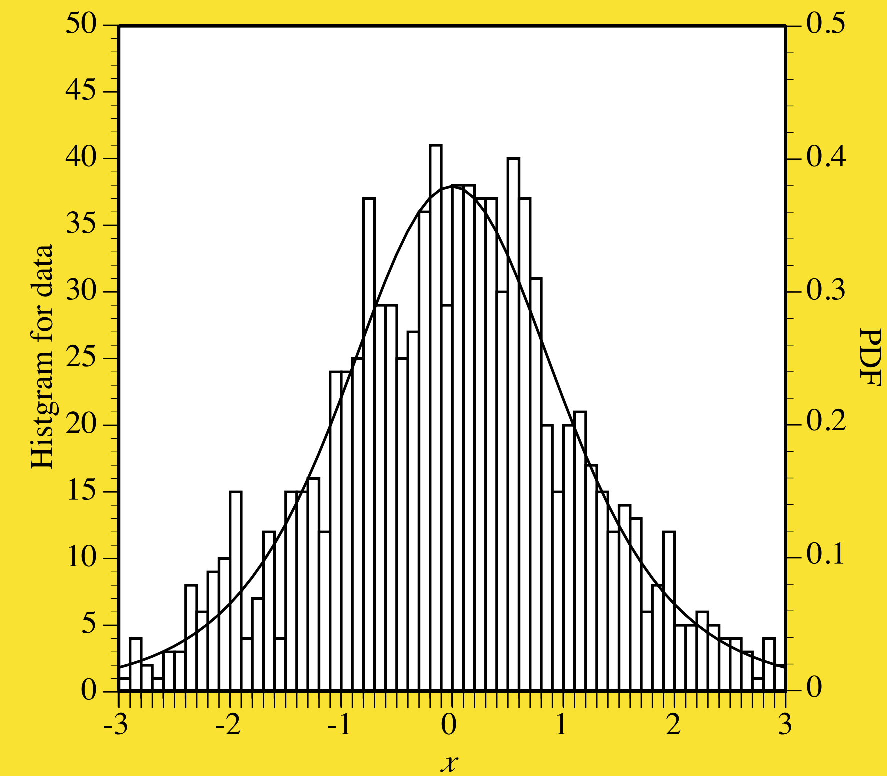
$$x=\frac{u}{\sqrt{v/n}}$$

を新たな確率変数とする標本分布を、自由度nのt分布と呼ぶ。

t分布.１３

２０歳男子の身長データーから無作為に５０００人分（平均と分散を合わせた正規乱数）を取り出して、そのデーターを５人を一組として先の処理を施したデーター（１０００個）のヒストグラムと理論PDFを比較した。

$$f_{n=5}(x) = \frac{8}{3\sqrt{5}\pi}\left(1+\frac{x^2}{5}\right)^{-3}$$



t分布.１４

標本を取り扱う際には、多くの場合母平均μは不明である。そこで母平均の代わりに標本平均を利用すると、次の定理が成立する。

正規母集団の標本$x_1, x_2, \cdots, x_n$が存在し、標準正規分布に従う変数を$u=(\bar{x}-\mu)\sqrt{n/\sigma^2}$、自由度$n-1$の

結論としては標本分散で標準化された標本平均はt分布に従うことを意味しており、この分布は標本平均を評価するのに利用でき、母平均の推測に活用できる。

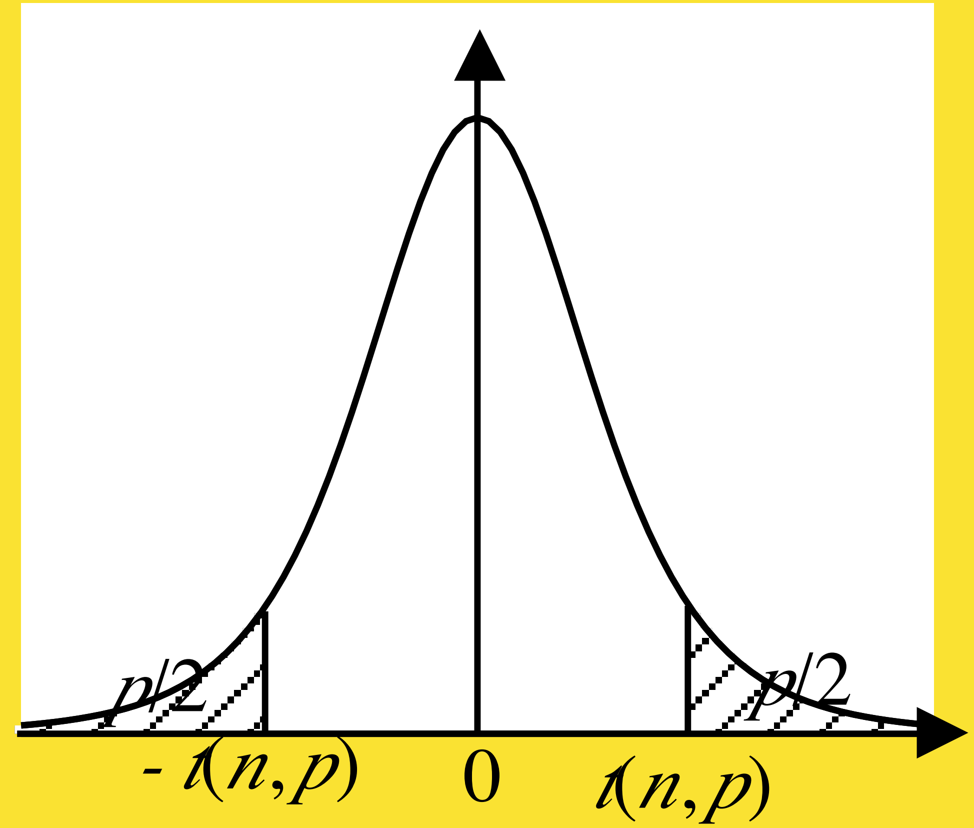
は自由度$n-1$のt分布に従う。

t分布.１５（分布表）

t分布表　　　　確率値

自由度

n　P	0.2	0.1	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.768
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
50	1.299	1.676	2.009	2.403	2.678	3.496
60	1.296	1.671	2.000	2.390	2.660	3.460
70	1.294	1.667	1.994	2.381	2.648	3.435
80	1.292	1.664	1.990	2.374	2.639	3.416
90	1.291	1.662	1.987	2.369	2.632	3.402
100	1.290	1.660	1.984	2.364	2.626	3.390



$$t(20, 0.05) = 2.086$$

F分布の解説

F分布. 1

【定義】自由度n_1のχ^2分布に従う変数v_1と自由度n_2のχ^2分布に従う変数v_2とするとき、変数

$$x=\frac{n_2 v_1}{n_1 v_2}$$

を新たな確率変数とする標本分布を自由度(n_1, n_2)のF分布と呼ぶ。その確率密度関数は

$$f(x)=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}x^{\frac{n_1}{2}-1}\left(\frac{n_1}{n_2}x+1\right)^{-\frac{n_1+n_2}{2}}$$

で与えられる。

赤い部分は証明が必要

F分布.2

《証明のつづき》

確率変数v_1、v_2の確率密度関数

$$f_{v_1,v_2}(v_1,v_2)=f_{v_1}(v_1)f_{v_2}(v_2)=\frac{e^{-\frac{v_1}{2}-\frac{v_2}{2}}}{4\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{v_1}{2}\right)^{\frac{n_1}{2}-1}\left(\frac{v_2}{2}\right)^{\frac{n_2}{2}-1}$$

変数変換 $x=\dfrac{n_2v_1}{n_1v_2}$ $\quad y=v_2$

$$\frac{\partial v_1}{\partial x}=\frac{n_1y}{n_2},\ \frac{\partial v_1}{\partial y}=\frac{n_1x}{n_2},\ \frac{\partial v_2}{\partial x}=0,\ \frac{\partial v_2}{\partial y}=1$$

$$J=\frac{\partial v_1}{\partial x}\frac{\partial v_2}{\partial y}-\frac{\partial v_1}{\partial y}\frac{\partial v_2}{\partial x}=\frac{n_1y}{n_2}$$

F分布.3

《証明のつづき》

確率変数x、yの確率密度関数

$$f_{x,y}(x,y)=\frac{e^{-\frac{v_1}{2}-\frac{v_2}{2}}\left(\frac{v_1}{2}\right)^{\frac{n_1}{2}-1}\left(\frac{v_2}{2}\right)^{\frac{n_2}{2}-1}}{4\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}J=\frac{e^{-\frac{n_1xy}{2n_2}-\frac{y}{2}}\left(\frac{n_1xy}{2n_2}\right)^{\frac{n_1}{2}-1}\left(\frac{y}{2}\right)^{\frac{n_2}{2}-1}\frac{n_1y}{n_2}}{4\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}$$

$$=\frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{2^{\frac{n_1}{2}+\frac{n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}x^{\frac{n_1}{2}-1}y^{\frac{n_1}{2}+\frac{n_2}{2}-1}e^{-\left(\frac{n_1}{2n_2}x+\frac{1}{2}\right)y}$$

F分布.4

《証明のつづき》

確率変数yに関する積分を実行すると確率変数xに対する確率密度関数が求まる。

$$f(x)=\int_0^\infty dy f_{x,y}(x,y)=\frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}}{2^{\frac{n_1}{2}+\frac{n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\int_0^\infty dy y^{\frac{n_1}{2}+\frac{n_2}{2}-1} e^{-\left(\frac{n_1}{2n_2}x+\frac{1}{2}\right)y}$$

$$=\frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}\Gamma\left(\frac{n_1}{2}+\frac{n_2}{2}\right)}{2^{\frac{n_1}{2}+\frac{n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\left(\frac{n_1}{2n_2}x+\frac{1}{2}\right)^{\frac{n_1}{2}+\frac{n_2}{2}}}=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}\left(\frac{n_1}{n_2}x+1\right)^{-\frac{n_1+n_2}{2}}$$

Q.E.D.

F分布. 5

F分布の例

$$f_{n_1=1,n_2=1}(x)=\frac{x^{-\frac{1}{2}}}{\pi}(x+1)^{-1}$$

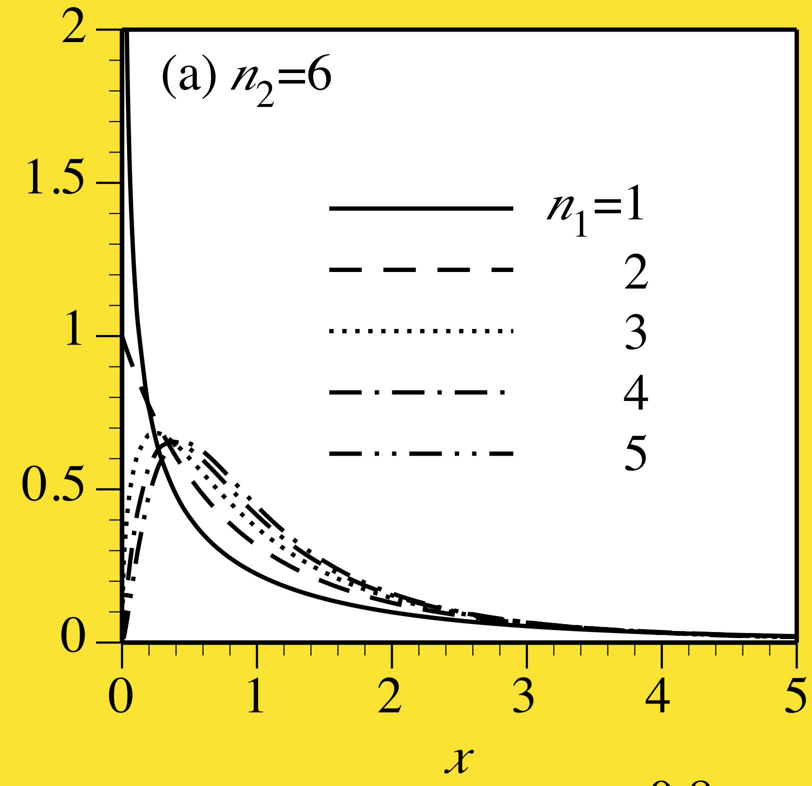
$$f_{n_1=2,n_2=1}(x)=(2x+1)^{-\frac{3}{2}}$$

$$f_{n_1=1,n_2=2}(x)=\frac{x^{-\frac{1}{2}}}{2\sqrt{2}}\left(\frac{x}{2}+1\right)^{-\frac{3}{2}}$$

$$f_{n_1=3,n_2=1}(x)=\frac{6\sqrt{3}x^{\frac{1}{2}}}{\pi}(3x+1)^{-2}$$

$$f_{n_1=2,n_2=2}(x)=(x+1)^{-2}$$

$$f_{n_1=1,n_2=3}(x)=\frac{2x^{-\frac{1}{2}}}{\sqrt{3}\pi}\left(\frac{x}{3}+1\right)^{-2}$$

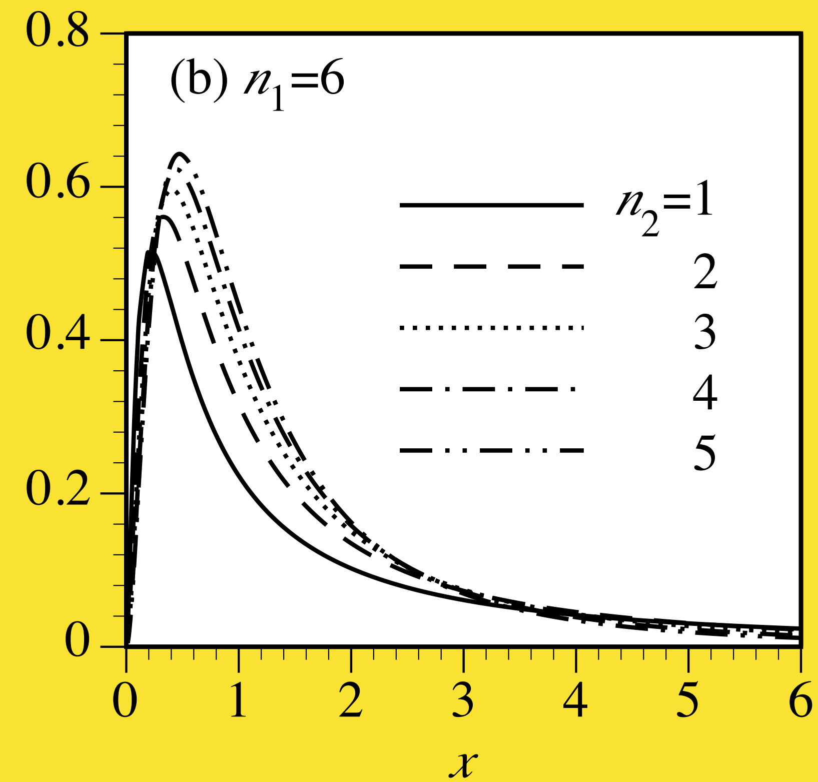


$$f_{1,6}(x)=\frac{5\sqrt{6}}{32}x^{-\frac{1}{2}}\left(\frac{x}{6}+1\right)^{-\frac{7}{2}}$$

$$f_{2,6}(x)=\left(\frac{x}{3}+1\right)^{-4}$$

$$f_{6,1}(x)=\frac{405}{2}x^{2}(6x+1)^{-\frac{7}{2}}$$

$$f_{6,2}(x)=81x^{2}(3x+1)^{-4}$$



F分布. 6

平均

$$\mu = \frac{n_2}{n_2 - 2}$$

自由度n_1には依存しない。

$n_2 > 2$でしか存在しない。

分散

$$\sigma^2 = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}$$

$n_2 > 4$でしか存在しない。

F分布.7

F分布の公式

$$F\left(n_1,n_2;p\right)=\frac{1}{F\left(n_2,n_1;1-p\right)}$$

《証明》

$$p=\int_{x_A}^{\infty}dxf_{n_1,n_2}\left(x\right) \qquad x_A=F\left(n_1,n_2;p\right)$$

とおくと、上公式が成立するということは

$$p=\int_0^{x_A^{-1}}dxf_{n_2,n_1}\left(x\right)$$

が証明されればよい。

F分布. 8

《証明のつづき》

$$p=\int_0^{x_A^{-1}} dx f_{n_2,n_1}(x)=\int_0^{x_A^{-1}} dx \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_2}{n_1}\right)^{\frac{n_2}{2}} x^{\frac{n_2}{2}-1}\left(\frac{n_2}{n_1}x+1\right)^{-\frac{n_1+n_2}{2}}$$

変数変換　$y=\dfrac{1}{x}$　　$dy=-\dfrac{1}{x^2}dx$

$$=-\int_\infty^{x_A} dy \frac{1}{y^2}\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\frac{n_2}{n_1}\right)^{\frac{n_2}{2}} y^{1-\frac{n_2}{2}}\left(\frac{n_2}{n_1}\frac{1}{y}+1\right)^{-\frac{n_1+n_2}{2}}$$

F分布. 9

《証明のつづき》

$$=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\int_{x_A}^{\infty}dy\left(\frac{n_2}{n_1}\right)^{\frac{n_2}{2}}y^{-1-\frac{n_2}{2}}\left(\frac{n_2}{n_1}\frac{1}{y}\right)^{-\frac{n_1+n_2}{2}}\left(1+\frac{n_2}{n_1}y\right)^{-\frac{n_1+n_2}{2}}$$

$$=\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\int_{x_A}^{\infty}dy\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}y^{\frac{n_1}{2}-1}\left(1+\frac{n_2}{n_1}y\right)^{-\frac{n_1+n_2}{2}}=\int_{x_A}^{\infty}dx f_{n_1,n_2}(x)$$

Q.E.D.

F分布. １０

２つの正規母集団N(μ_1,σ_1^2)とN(μ_2,σ_2^2)に従う分布が存在し、それぞれから大きさn_1とn_2の標本を選び出し、それぞれの２乗和を母分散で規格化した変数を考慮する。

$$S_1 \quad (\ \ -1)\, s_1^2 \qquad S_2 \quad (\ \ -1)\, s_2^2$$

このようにF分布は標本分散の比率を評価する標本分布であり、これらの性質から母分散比率を推測することが可能である。

$$\overline{v_2/\left(n_2-1\right)} = \overline{s_2^2\sigma_1^2}$$

は自由度(n_1−1, n_2−1)のF分布に従う。

F分布.１１（分布表）

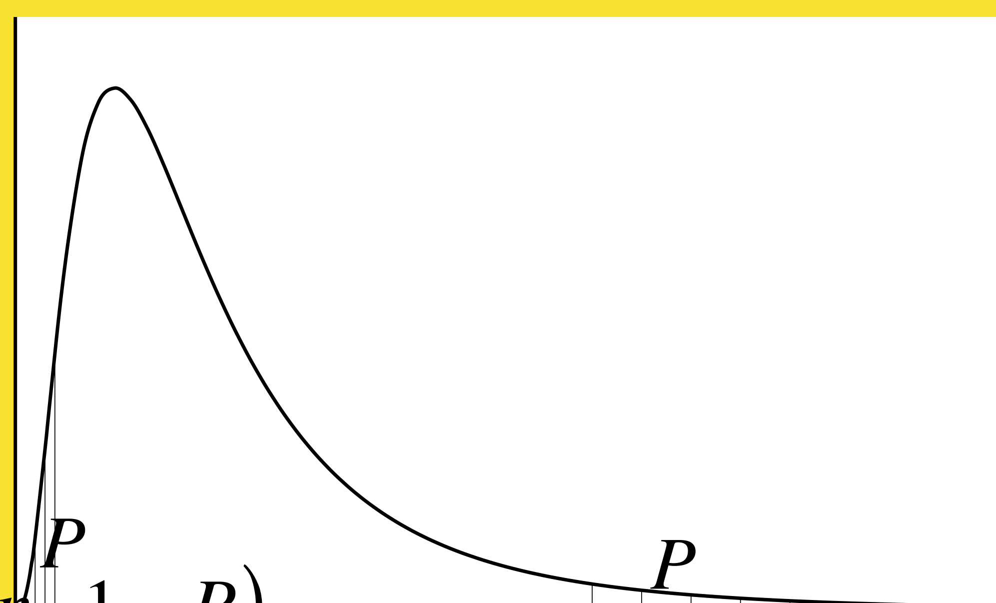
確率値

自由度１

自由度２

F分布表 (2.5%)

n_2 \ n_1	1	2	3	4	5	6	10	11	12	13	14	15	20	30	50	100
1	648	800	864	900	922	937	969	973	977	980	983	985	993	1001	1008	1013
2	38.5	39	39.2	39.3	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
3	17.4	16	15.4	15.1	14.9	14.7	14.4	14.4	14.3	14.3	14.3	14.3	14.2	14.1	14	14
4	12.2	10.7	9.98	9.6	9.36	9.2	8.84	8.79	8.75	8.71	8.68	8.66	8.56	8.46	8.38	8.32
5	10	8.43	7.76	7.39	7.15	6.98	6.62	6.57	6.52	6.49	6.46	6.43	6.33	6.23	6.14	6.08
6	8.81	7.26	6.6	6.23	5.99	5.82	5.46	5.41	5.37	5.33	5.3	5.27	5.17	5.07	4.98	4.92
10	6.94	5.46	4.83	4.47	4.24	4.07	3.72	3.66	3.62	3.58	3.55	3.52	3.42	3.31	3.22	3.15
11	6.72	5.26	4.63	4.28	4.04	3.88	3.53	3.47	3.43	3.39	3.36	3.33	3.23	3.12	3.03	2.96
12	6.55	5.1	4.47	4.12	3.89	3.73	3.37	3.32	3.28	3.24	3.21	3.18	3.07	2.96	2.87	2.8
13	6.41	4.97	4.35	4	3.77	3.6	3.25	3.2	3.15	3.12	3.08	3.05	2.95	2.84	2.74	2.67
14	6.3	4.86	4.24	3.89	3.66	3.5	3.15	3.09	3.05	3.01	2.98	2.95	2.84	2.73	2.64	2.56
15	6.2	4.77	4.15	3.8	3.58	3.41	3.06	3.01	2.96	2.92	2.89	2.86	2.76	2.64	2.55	2.47
20	5.87	4.46	3.86	3.51	3.29	3.13	2.77	2.72	2.68	2.64	2.6	2.57	2.46	2.35	2.25	2.17
30	5.57	4.18	3.59	3.25	3.03	2.87	2.51	2.46	2.41	2.37	2.34	2.31	2.2	2.07	1.97	1.88
50	5.34	3.97	3.39	3.05	2.83	2.67	2.32	2.26	2.22	2.18	2.14	2.11	1.99	1.87	1.75	1.66
100	5.18	3.83	3.25	2.92	2.7	2.54	2.18	2.12	2.08	2.04	2	1.97	1.85	1.71	1.59	1.48



$$F(n_1, n_2, 1-P) = 1/F(n_2, n_1, P)$$

$$F(n_1, n_2, P)$$

$$F(20,10,0.025) = 3.42$$

$$F(20,10,0.975) = \frac{1}{F(10,20,0.025)} = \frac{1}{2.77} = 0.361$$