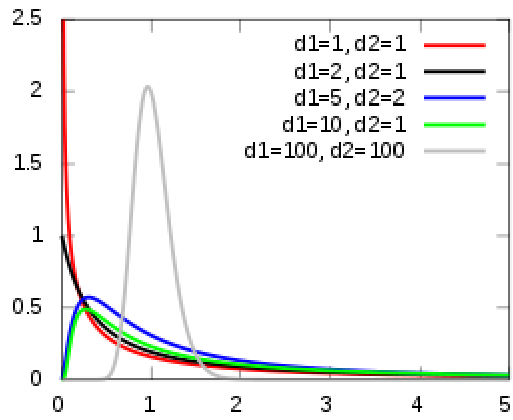
WIKIPÉDIA

Distribuição F de Fisher-Snedecor

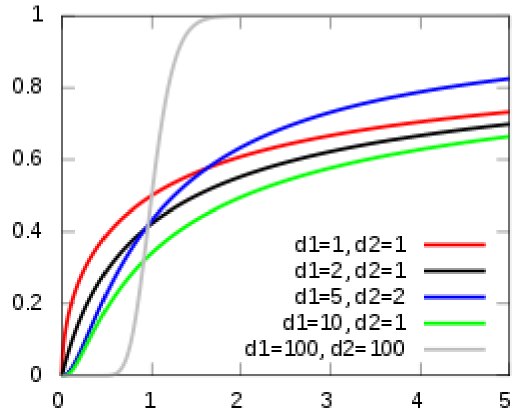
Origem: Wikipédia, a enciclopédia livre.

Em teoria das probabilidades e estatística, a **distribuição *F* de Fisher-Snedecor**, também conhecida como **distribuição *F***, **distribuição F de Fisher** e **distribuição F de Snedecor**, em homenagem ao biólogo e estatístico britânico Ronald Fisher e ao matemático norte-americano George Waddel Snedecor,[1] é uma distribuição de probabilidade contínua que surge frequentemente como a distribuição nula da estatística de um teste, mais notadamente na análise de variância, como no teste *F*.[2][3][4][5]

Distribuição F de Fisher-Snedecor



Função densidade de probabilidade



Função distribuição acumulada

Parâmetros	$d_1, d_2 > 0$ graus de liberdade
Suporte	$x \in [0, +\infty)$
f.d.p.	$\dfrac{\sqrt{\dfrac{(d_1\,x)^{d_1}\,\,d_2^{d_2}}{(d_1\,x+d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)}$
f.d.a.	$I_{\frac{d_1 x}{d_1 x+d_2}}\left(\frac{d_1}{2},\frac{d_2}{2}\right)$
Média	$\dfrac{d_2}{d_2-2}$ para $d_2 > 2$
Moda	$\dfrac{d_1-2}{d_1}\,\dfrac{d_2}{d_2+2}$

Índice

Definição

Caracterização

Equação diferencial

Propriedades e distribuições relacionadas

Ver também

Referências

Ligações externas

Definição

Se uma variável aleatória X tiver uma distribuição F com parâmetros d_1 e d_2, escrevemos $X \sim F(d_1, d_2)$. Então, a função densidade de probabilidade de X é dada por

	para $d_1 > 2$
Variância	$\frac{2\,d_2^2\,(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}$ para $d_2 > 4$
Obliquidade	$\frac{(2d_1+d_2-2)\sqrt{8(d_2-4)}}{(d_2-6)\sqrt{d_1(d_1+d_2-2)}}$ para $d_2 > 6$
Curtose	Definida no texto.
Entropia	$\ln\left(\sigma\sqrt{2\,\pi\,\mathrm{e}}\right)$
Função Geradora de Momentos	Não existe. Os momentos brutos estão definidos no texto.
Função Característica	$\frac{\Gamma(\frac{d_1+d_2}{2})}{\Gamma(\frac{d_2}{2})}U\left(\frac{d_1}{2},1-\frac{d_2}{2},-\frac{d_2}{d_1}\imath t\right)$ onde $U(a,b,z)$ é a função hipergeométrica confluente do segundo tipo

$$f(x;d_1,d_2)=\frac{\sqrt{\frac{(d_1\,x)^{d_1}\,\,d_2^{d_2}}{(d_1\,x+d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)}$$

$$=\frac{1}{\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)}\left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}}x^{\frac{d_1}{2}-1}\left(1+\frac{d_1}{d_2}\,x\right)^{-\frac{d_1+d_2}{2}}$$

para x real e maior que zero. Aqui, B é uma função beta. Em muitas aplicações, os parâmetros d_1 e d_2 são números inteiros positivos, mas a distribuição é bem definida para valores reais positivos destes parâmetros.

A função distribuição acumulada é

$$F(x;d_1,d_2)=I_{\frac{d_1x}{d_1x+d_2}}\left(\frac{d_1}{2},\frac{d_2}{2}\right),$$

em que I é a função beta incompleta regularizada.

O valor esperado, a variância e outros detalhes sobre $F(d_1,d_2)$ são dados na caixa ao lado. Para $d_2 > 8$, a curtose de excesso é

$$\gamma_2=12\frac{d_1(5d_2-22)(d_1+d_2-2)+(d_2-4)(d_2-2)^2}{d_1(d_2-6)(d_2-8)(d_1+d_2-2)}.$$

O k-ésimo momento de uma distribuição $F(d_1,d_2)$ existe e é finita somente quando $2k < d_2$ e é igual a[6]

$$\mu_X(k) = \left(\frac{d_2}{d_1}\right)^k \frac{\Gamma\left(\frac{d_1}{2}+k\right)}{\Gamma\left(\frac{d_1}{2}\right)} \frac{\Gamma\left(\frac{d_2}{2}-k\right)}{\Gamma\left(\frac{d_2}{2}\right)}$$

A distribuição *F* é uma parametrização particular da distribuição beta prima, também chamada de distribuição beta de segundo tipo.

A função característica é[7]

$$\varphi^F_{d_1,d_2}(s) = \frac{\Gamma(\frac{d_1+d_2}{2})}{\Gamma(\frac{d_2}{2})} U\left(\frac{d_1}{2}, 1-\frac{d_2}{2}, -\frac{d_2}{d_1}\imath s\right)$$

em que $U(a,b,z)$ é a função hipergeométrica confluente do segundo tipo.

Caracterização

O valor observado de uma variável aleatória de distribuição *F* com parâmetros d_1 e d_2 surge como a razão de dois valores observados de distribuição qui-quadrado apropriadamente escalados:[8]

$$X = \frac{U_1/d_1}{U_2/d_2}$$

em que

- U_1 e U_2 têm distribuições qui-quadrado com graus de liberdade d_1 e d_2 respectivamente e
- U_1 e U_2 são independentes.

Em instâncias em que a distribuição *F* é usada, por exemplo, na análise de variância, a independência de U_1 e U_2 pode ser demonstrada pela aplicação do teorema de Cochran.

Equivalentemente, a variável aleatória da distribuição *F* também pode ser escrita como

$$X = \frac{s_1^2}{\sigma_1^2} \;/\; \frac{s_2^2}{\sigma_2^2}$$

em que s_1^2 e s_2^2 são as somas dos quadrados S_1^2 e S_2^2 de dois processos normais com variâncias σ_1^2 e σ_2^2 divididas pelo número correspondente de χ^2 graus de liberdades. d_1 e d_2 são respectivamente $s_1^2 = \frac{S_1^2}{d_1}$ e $s_2^2 = \frac{S_2^2}{d_2}$.

Em um contexto frequencista, uma distribuição *F* escalada dá portanto a probabilidade $p(s_1^2/s_2^2|\sigma_1^2, \sigma_2^2)$, ela própria com distribuição F, sem qualquer escala, o que se aplica onde σ_1^2 é igual σ_2^2. Este é o contexto em que a distribuição *F* aparece de forma mais generalizada em testes *F*: em que a hipótese nula é de que duas variâncias normais independentes são iguais e as somas observadas de alguns quadrados apropriadamente selecionados são então examinadas a fim de verificar se sua razão é significantemente incompatível com esta hipótese nula.

A quantidade X tem a mesma distribuição na estatística bayesiana, se um método de Jeffreys não informativo, de rescalamento invariante for tomado para as probabilidades *a priori* de σ_1^2 e σ_2^2.[9] Neste contexto, uma distribuição *F* escalada dá assim a probabilidade *a posteriori* $p(\sigma_1^2, \sigma_2^2 | s_1^2/s_2^2)$, em que as somas agora observadas s_1^2 e s_2^2 são tomadas como conhecidas.

De forma geral, resumida e simplificada, a distribuição F tem como características básicas:

- É uma família de curvas, cada uma, determinada por dois tipos de graus de liberdade, os correspondentes à variância no numerador, e os que correspondem à variância no denominador.
- É uma distribuição positivamente assimétrica.
- A área total sob cada curva de uma distribuição F é igual a 1.
- Todos os valores de X são maiores ou iguais a 0.
- Para todas as distribuições F, o valor médio de X é aproximadamente igual a 1.[10]

Equação diferencial

A função densidade de probabilidade da distribuição *F* é uma solução da seguinte equação diferencial:

$$\left\{\begin{array}{l} 2x\left(d_1 x + d_2\right) f'(x) + \left(2 d_1 x + d_2 d_1 x - d_2 d_1 + 2 d_2\right) f(x) = 0, \\ \\ f(1) = \frac{d_1^{\frac{d_1}{2}} d_2^{\frac{d_2}{2}} \left(d_1 + d_2\right)^{\frac{1}{2}\left(-d_1 - d_2\right)}}{B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \end{array}\right\}$$

Propriedades e distribuições relacionadas

- Se $X \sim \chi^2_{d_1}$ e $Y \sim \chi^2_{d_2}$ forem independentes, então $\frac{X/d_1}{Y/d_2} \sim \mathrm{F}(d_1, d_2)$;
- Se $X_k \sim \Gamma(\alpha_k, \beta_k)$ forem independentes, então $\frac{\alpha_2 \beta_1 X_1}{\alpha_1 \beta_2 X_2} \sim \mathrm{F}(2\alpha_1, 2\alpha_2)$;
- Se $X \sim \mathrm{Beta}(d_1/2, d_2/2)$(distribuição beta), então $\frac{d_2 X}{d_1(1-X)} \sim \mathrm{F}(d_1, d_2)$;
- Equivalentemente, se $X \sim F(d_1, d_2)$, então $\frac{d_1 X/d_2}{1 + d_1 X/d_2} \sim \mathrm{Beta}(d_1/2, d_2/2)$;
- Se $X \sim F(d_1, d_2)$, então $Y = \lim_{d_2 \to \infty} d_1 X$ tem a distribuição qui-quadrado $\chi^2_{d_1}$;
- $F(d_1, d_2)$ é equivalente a distribuição T-quadrado de Hotelling escalada $\frac{d_2}{d_1(d_1 + d_2 - 1)} \mathrm{T}^2(d_1, d_1 + d_2 - 1)$;
- Se $X \sim F(d_1, d_2)$, então $X^{-1} \sim F(d_2, d_1)$;
- Se $X \sim t(n)$ (distribuição t de Student), então:

$$X^2 \sim \mathrm{F}(1, n)$$
$$X^{-2} \sim \mathrm{F}(n, 1)$$

- A distribuição *F* é um caso especial de distribuição de Pearson de tipo 6;
- Se X e Y forem independentes com $X, Y \sim \mathrm{Laplace}(\mu, b)$, então:

$$\frac{|X-\mu|}{|Y-\mu|} \sim \mathrm{F}(2,2);$$

- Se $X \sim F(n,m)$, então $\frac{\log X}{2} \sim \mathrm{FisherZ}(n,m)$ (distribuição z de Fisher);
- A distribuição F não central simplifica à distribuição *F* se $\lambda = 0$;
- A distribuição F não central dupla simplifica à distribuição *F* se $\lambda_1 = \lambda_2 = 0$;
- Se $\mathrm{Q}_X(p)$ for o quantil p para $X \sim F(d_1, d_2)$ e $\mathrm{Q}_Y(1-p)$ for o quantil $1-p$ para $Y \sim F(d_2, d_1)$, então

$$\mathrm{Q}_X(p) = \frac{1}{\mathrm{Q}_Y(1-p)}.$$

Ver também

- Distribuição gama
- Qui-quadrado
- Teste de Chow

Referências

1. «Earliest Known Uses of Some of the Words of Mathematics (F)» (http://jeff560.tripod.com/f.html). *jeff560.tripod.com*. Consultado em 19 de junho de 2017
2. Johnson, Norman Lloyd; Kotz, Samuel; Balakrishnan, N. (8 de maio de 1995). *Continuous univariate distributions* (https://books.google.com.br/books?id=0QzvAAAAMAAJ&dq=Continuous+Univariate+Distributions&hl=pt-BR&sa=X&redir_esc=y) (em inglês). [S.l.]: Wiley & Sons. ISBN 9780471584940
3. Abramowitz, Milton; Stegun, Irene A. (30 de abril de 2012). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* (https://books.google.com.br/books?id=KiPCAgAAQBAJ&printsec=frontcover&dq=Handbook+of+Mathematical+Functions+with+Formulas,+Graphs,+and+Mathematical+Tables&hl=pt-BR&sa=X&redir_esc=y#v=onepage&q=Handbook%20of%20Mathematical%20Functions%20with%20Formulas,%20Graphs,%20and%20Mathematical%20Tables&f=false) (em inglês). [S.l.]: Courier Corporation. ISBN 9780486158242
4. «1.3.6.6.5. F Distribution» (http://www.itl.nist.gov/div898/handbook/eda/section3/eda3665.htm). *www.itl.nist.gov*. Consultado em 19 de junho de 2017
5. Mood, Alexander McFarlane; Graybill, Franklin A.; Boes, Duane C. (janeiro 1974). *Introduction to the Theory of Statistics* (https://books.google.com.br/books?id=Viu2AAAAIAAJ&q=Introduction+to+the+Theory+of+Statistics&dq=Introduction+to+the+Theory+of+Statistics&hl=pt-BR&sa=X&redir_esc=y) (em inglês). [S.l.]: McGraw-Hill. ISBN 9780070428645
6. «F distribution» (https://www.statlect.com/probability-distributions/F-distribution). *www.statlect.com*. Consultado em 19 de junho de 2017
7. Phillips, P. C. B. (1 de abril de 1982). «The true characteristic function of the F distribution» (https://academic.oup.com/biomet/article-abstract/69/1/261/243110/The-true-characteristic-function-of-the-F). *Biometrika*. **69** (1): 261–264. ISSN 0006-3444 (https://www.worldcat.org/issn/0006-3444). doi:10.1093/biomet/69.1.261 (https://dx.doi.org/10.1093%2Fbiomet%2F69.1.261)

8. DeGroot, Morris H.; Schervish, Mark J. (2002). *Probability and Statistics* (https://books.google.com.br/books?id=iH4ZAQAAIAAJ&q=de+groot+Probability+and+Statistics&dq=de+groot+Probability+and+Statistics&hl=pt-BR&sa=X&ved=0ahUKEwjgiK70nsrUAhVHUJAKHX16CBUQ6AEIJzAA) (em inglês). [S.l.]: Addison-Wesley. ISBN 9780201524888
9. Box, George E. P.; Tiao, George C. (25 de janeiro de 2011). *Bayesian Inference in Statistical Analysis* (https://books.google.com.br/books?id=T8Askeyk1k4C&printsec=frontcover&dq=Bayesian+Inference+in+Statistical+Analysis&hl=pt-BR&sa=X&redir_esc=y#v=onepage&q=Bayesian%20Inference%20in%20Statistical%20Analysis&f=false) (em inglês). [S.l.]: John Wiley & Sons. ISBN 9781118031445
10. LARSON, Ron; FARBER, Betsy (2016). *Estatística Aplicada*. São Paulo: PEARSON. 2 páginas

Ligações externas

- Tabela de valores críticos da distribuição *F* (em inglês) (http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm)
- Calculadora gratuita para teste *F* (em inglês) (http://www.waterlog.info/f-test.htm)