



Universidade Estadual de Santa Cruz – UESC  
Pavilhão de Ciência Exatas e Computação – DEC  
Curso de Ciência da Computação

# Variáveis Aleatórias Contínuas

CET083 – Probabilidade e Estatística  
Professor: José Cláudio Faria  
Aluno: Wilson Santos Silva Filho  
2024.1



# Variáveis Aleatórias

Uma variável aleatória contínua tem a capacidade de tomar qualquer valor numérico dentro de um intervalo específico ou múltiplos intervalos. Essa variável pode representar valores em intervalos de números reais. Isso significa que entre dois pontos quaisquer, existem infinitas quantidades de valores possíveis, que são determinados pela precisão do instrumento de medição.

Experimento	Variável aleatória (Y)	Possíveis valores para a VAC
Trabalhar em um projeto	Percentual executado após 30 dias	$0 \leq Y \leq 100\%$
Observar um operador de máquina agrícola	Tempo ocioso em um dia	$0 \leq Y \leq 24 \text{ hs}$
Pulverizar 10.000 m <sup>2</sup> de uma área agrícola	Volume de água gasto	$0 \leq Y \leq \infty$

# Exemplo

No esporte olímpico do lançamento de dardo, a distância máxima que um dardo pode alcançar é de 90 metros. Para se qualificar para a competição, um atleta deve ser capaz de lançar o dardo pelo menos 65 metros. Isso cria um intervalo de lançamento, onde todos os lançamentos possíveis ocorrerão entre 65 e 90 metros.

Dentro deste intervalo, existem infinitas possibilidades para a distância que um lançamento pode alcançar. Por exemplo, um lançamento poderia atingir uma distância específica de 89,3438 metros.

# Função Densidade de Probabilidade

Em probabilidade e estatística, a função de densidade de probabilidade (FDP) de uma variável aleatória contínua é a função que quantifica a probabilidade relativa de uma variável aleatória assumir um valor específico.

A probabilidade de uma variável aleatória contínua assumir um valor dentro de um intervalo específico é calculada pela integral da função densidade de probabilidade sobre o intervalo.

# Função Densidade de Probabilidade

Seja  $Y$  uma VAC, a função densidade de probabilidade  $f(y)$  é uma função que deve satisfazer as seguintes condições:

$$a) f(y) \geq 0 \text{ para todo } y \in [a, b] \text{ com } a < b$$

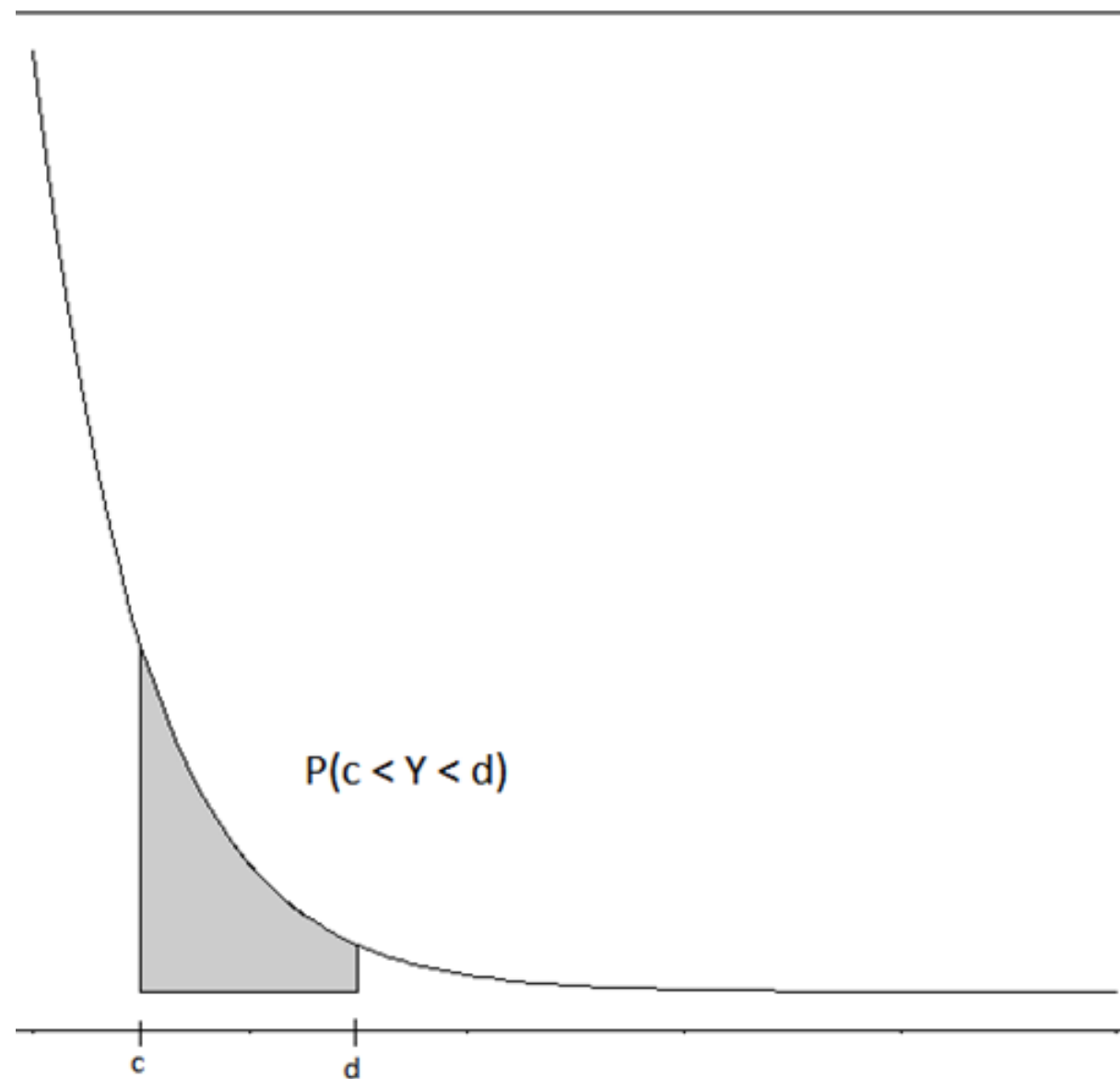
$$b) \int_a^b f(y) dy = 1$$

Observar: Deve-se notar que a função  $f(y)$ , da densidade de probabilidade, não é probabilidade! Apenas quando integramos entre dois extremos (limites) será produzido uma probabilidade.



# Função Densidade de Probabilidade

Além disso, define-se, para qualquer  $[c < d]$ , contido no intervalo  $[a, b]$



$$P(c < Y < d) = \int_c^d f(y) dy$$

Sendo  $P$  a área, a variável deve ser maior que o limite inferior e menor que o limite superior.

# Variáveis Aleatórias

A definição anteriormente citada mostra que a probabilidade de qualquer valor especificado de  $Y$ , por exemplo,  $y_0$ , tem  $P(Y = y_0) = 0$ , pois:

$$P(Y = y_0) = \int_{y_0}^{y_0} f(y)dy = 0$$

Assim, as probabilidades abaixo serão todas iguais, se  $Y$  for uma VAC:

$$P(a \leq Y \leq b) = P(a \leq Y < b) = P(a < Y \leq b) = P(a < Y < b)$$

# Exemplo

Seja  $X$  uma v.a.c com função de densidade dada por:

$$f(x) \begin{cases} \frac{c(4x - 2x^2)}{0}, & \text{se } x \in (0,2) \\ 0, & \text{se } x \notin (0,2) \end{cases}$$

- a. Qual o valor de  $c$ ?
- b. Calcule a  $P(X > 1)$



# Exemplo

Para  $f(x)$  ser uma função de densidade deve obedecer as exigências das propriedades a e b do [slide 5](#):

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Como a função está nos intervalos entre (0, 2):

$$= c \left( \int_0^2 4x dx - \int_0^2 2x^2 dx \right) = c \left( 8 - \frac{16}{3} \right) = \frac{8}{3}c$$

$$\mathbf{S} = c = \frac{3}{8}$$

# Exemplo

b. Calcule a  $P(X > 1)$

$$P(x > 1) = \int_1^{\infty} f(x) dx = \int_1^2 \frac{3}{8} (4x - 2x^2) dx = \frac{1}{2}$$

# Função Distribuição Acumulada

Em Estatística, a **função de distribuição acumulada (FDA)**, também conhecida como **função de distribuição**, para **variáveis aleatórias contínuas**, é uma ferramenta fundamental para calcular probabilidades relacionadas a eventos aleatórios. Ela representa a probabilidade de que a variável aleatória  $X$  assuma um valor menor ou igual a um valor específico  $x$ .

A FDA e a FDP estão intimamente relacionadas. A FDP representa a **densidade de probabilidade** em cada ponto da variável aleatória, enquanto a FDA fornece a **probabilidade acumulada** até um determinado valor.

A partir da FDP, podemos obter a FDA.

# Função de Distribuição Acumulada

**Definição:** A função de distribuição acumulada  $F(y)$  de uma v.a.c  $Y$  com densidade  $f(y)$  é:

$$F(y) = P(Y \leq y) = \int_{-\infty}^y f(t)dt, \text{ para } -\infty < y < \infty$$

## Propriedades:

A FDA possui algumas propriedades importantes que garantem sua utilidade na análise de probabilidades:

Monotonicidade não decrescente:  $F(x)$  é sempre não decrescente, o que significa que para qualquer  $x_1$  e  $x_2$ , onde  $x_1 < x_2$ , temos  $F(x_1) \leq F(x_2)$ .

Limites:  $F(x)$  tende a 0 quando  $x$  tende a menos infinito e tende a 1 quando  $x$  tende a infinito:

$$\lim_{x \rightarrow -\infty} F(x) = 0$$

$\lim_{x \rightarrow \infty} F(x) = 1$ ,  $F(x) = 1 - P(X > x)$ : A FDA é complementar à probabilidade de  $X$  assumir valores maiores que  $x$ .

# Distribuição Normal





# Distribuição Normal: ascensão

- Friederich Gauss, em meados do século XIX, durante seus estudos e pesquisas sobre eventos da natureza, observou que suas amostras possuíam comportamento padrão.
- Suas observações posteriormente foram base para criação da representação matemática para os eventos que ficavam em torno de um valor médio, com uma certa variabilidade, chamada: Curva de Gauss.





# Distribuição Normal:

- Na área de probabilidade e estatística, a distribuição normal se destaca como uma das distribuições probabilísticas mais aplicadas para representar eventos naturais. Ela serve como uma aproximação para a distribuição binomial quando o número de ensaios ( $n$ ) é grande. Além disso, as médias e proporções obtidas de amostras volumosas tendem a aderir à distribuição normal, o que facilita a análise e interpretação de dados em larga escala.

- **Onde:** A densidade da variável aleatória  $X$  normal com média  $\mu$  e  $\sigma$  o desvio padrão; moda, média e mediana coincidem, e é simétrico em relação à média.

NOTAÇÃO:  $X \sim N(\mu, \sigma^2)$

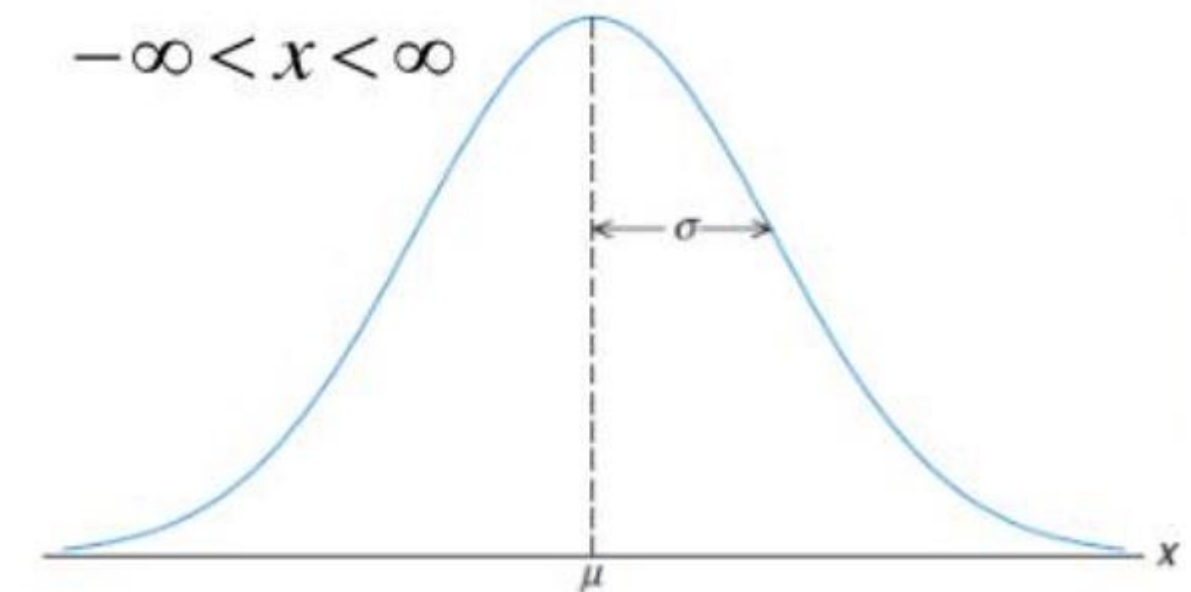


Figura I – Curva normal

- Média  $\mu$  e variância  $\sigma$  ao quadrado
- Desvio padrão é a raiz quadrada da variância

# Distribuição Normal:

- Então, seja uma variável aleatória contínua, dizemos que  $X \sim N(\mu, \sigma)$  se, e somente se sua FDP for dada por:

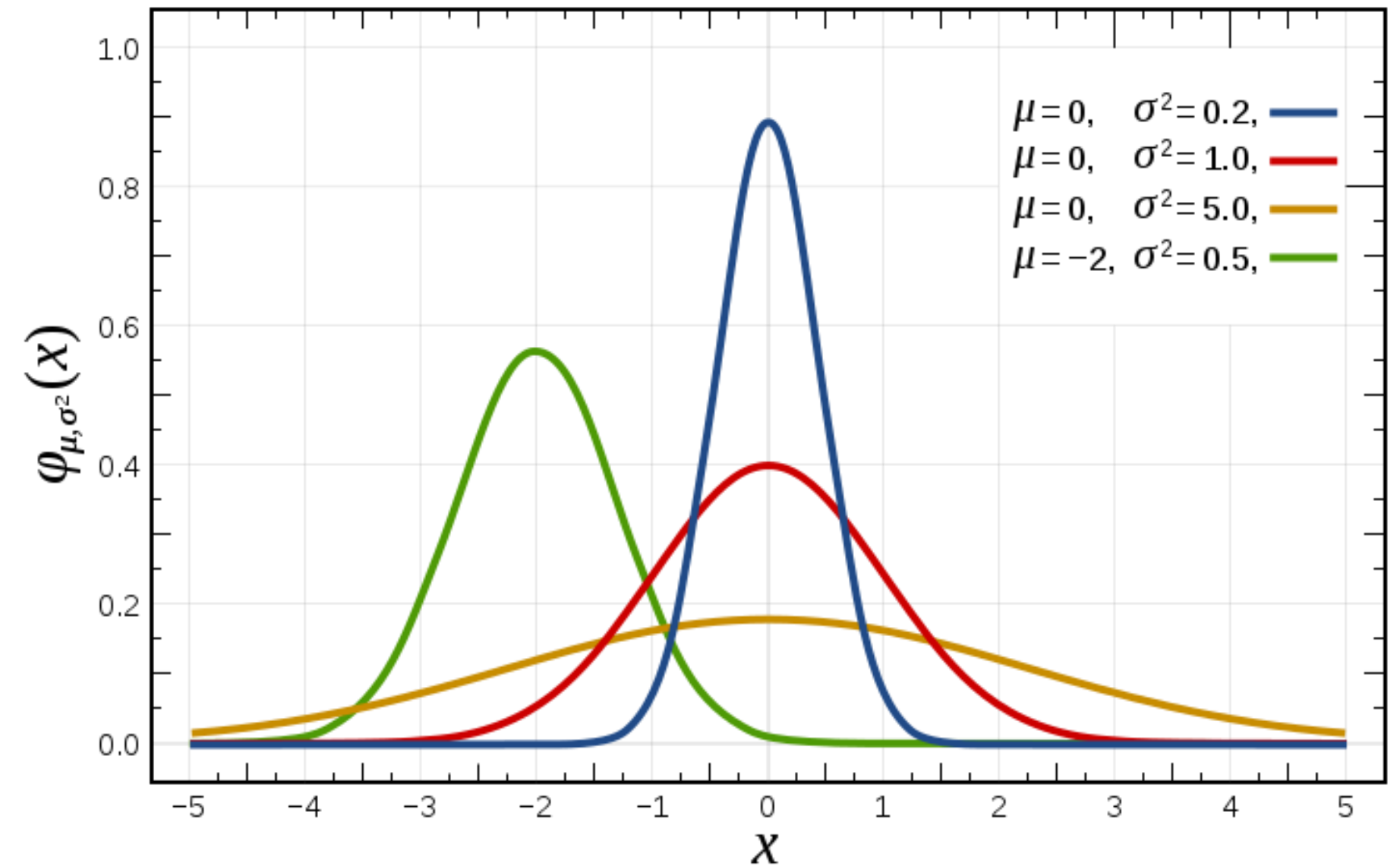
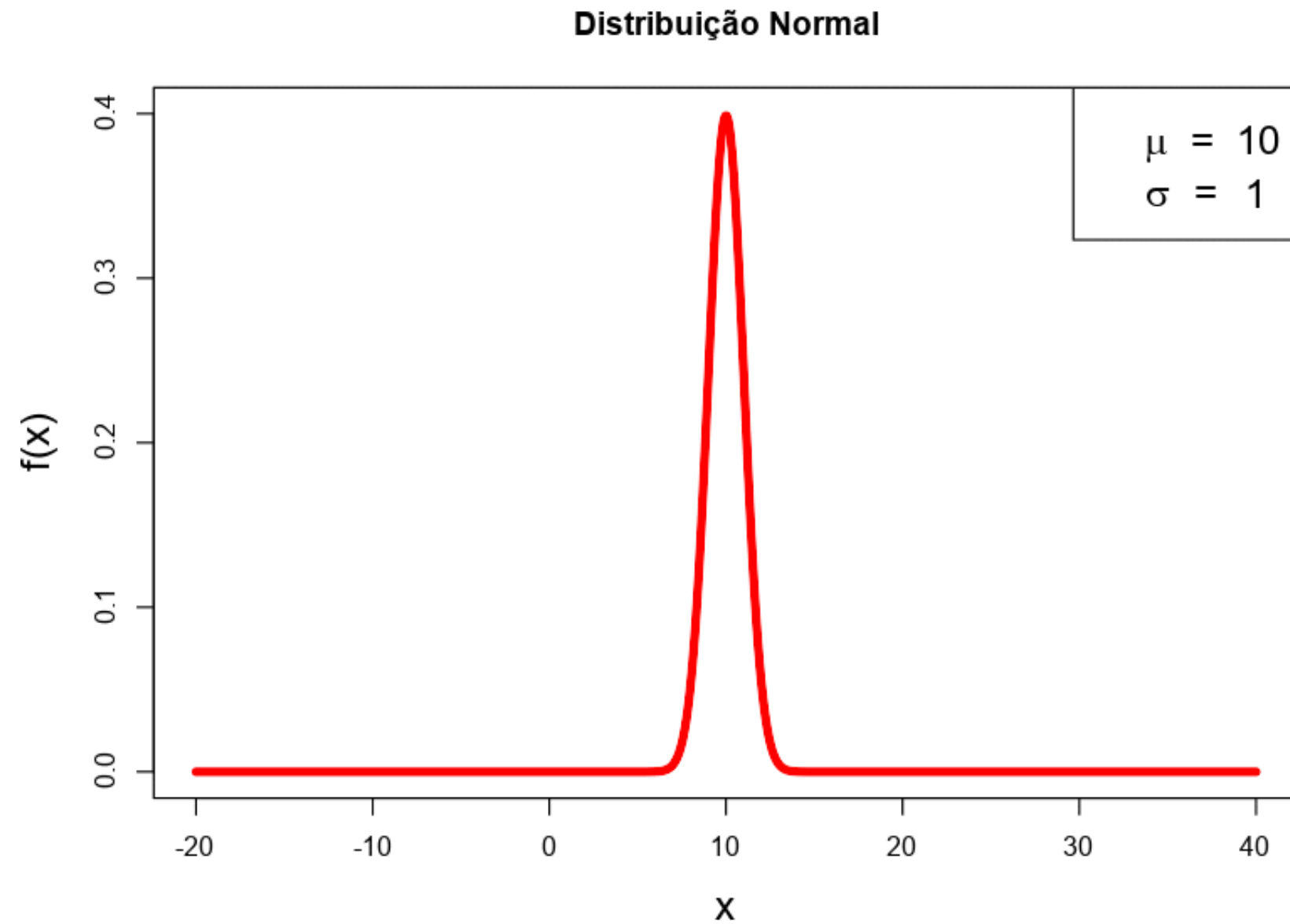
$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}, -\infty < x < \infty, \sigma > 0$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

- Graficamente temos:

# Distribuição Normal:

Gráfico de uma FDP

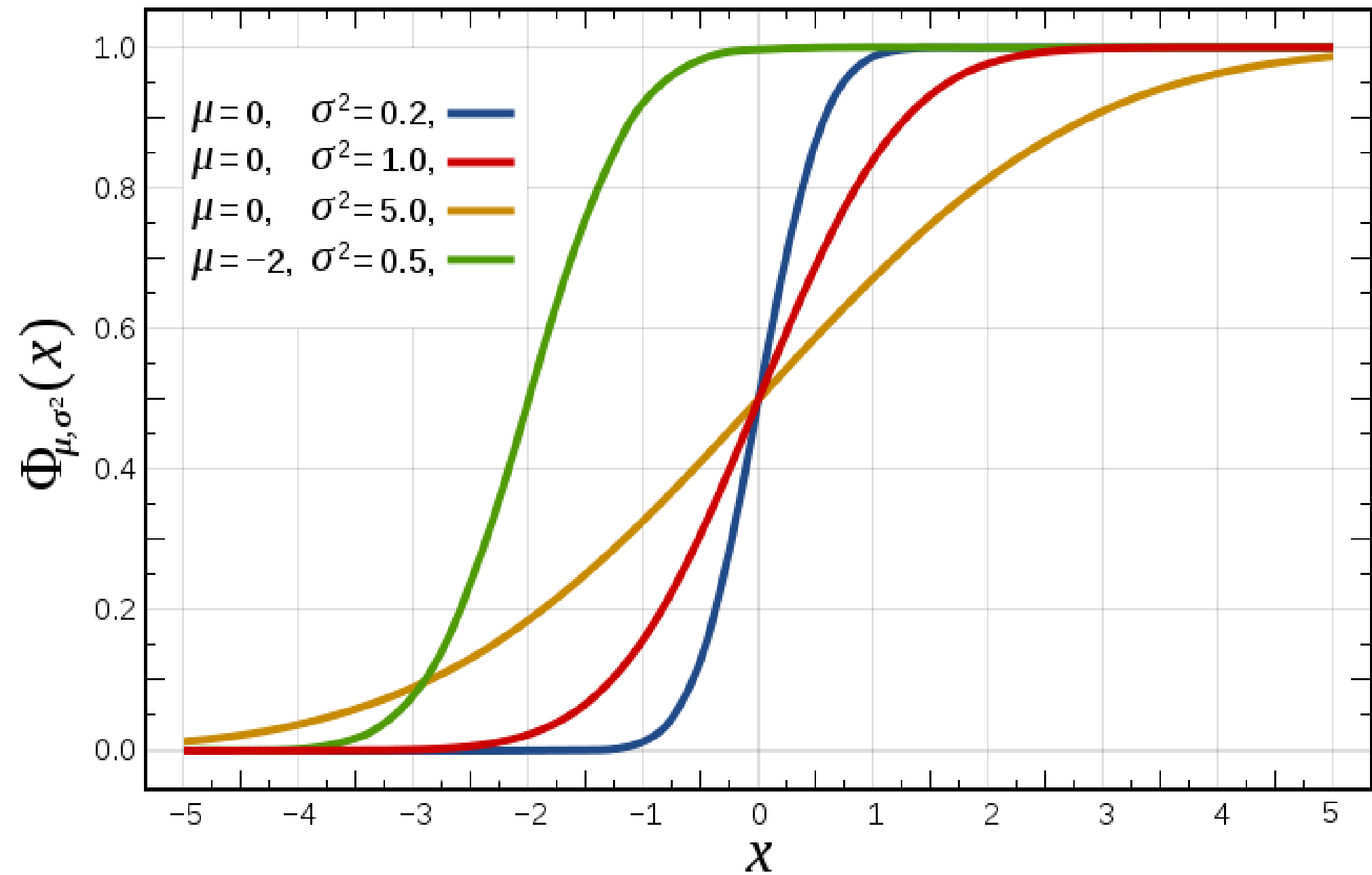


# Distribuição Normal:

## Gráfico de uma FDA

Esta equação representa a FDA de uma distribuição normal, onde “erf” denota a função erro, “x” é uma variável no eixo horizontal do gráfico, “μ” representa o valor médio e afeta o deslocamento horizontal neste eixo em relação ao ponto zero e “σ” representa o desvio padrão, que afeta o quão espalhadas ou acentuadamente curvadas são cada linha.

$$\frac{1}{2} \left( 1 + \operatorname{erf} \frac{x - \mu}{\sigma\sqrt{2}} \right)$$



# Distribuição Normal Padrão

Para o cálculo das probabilidades utilizando a função densidade de probabilidades surgiam dois problemas:

- I. Relativo a integração de  $f(y)$ , pois é necessário o desenvolvimento em séries, o que é um cálculo relativamente complexo.
- II. Tabelar todas as probabilidades considerando-se as várias combinações possíveis de  $\mu$  e  $\sigma$  acarretaria um grande trabalho, pois,  $f(y)$  depende dos parâmetros  $\mu$  e  $\sigma$ .

Esses problemas foram solucionados por meio de uma mudança de variável, obtendo-se, assim, a distribuição normal padronizada ou reduzida ( $\mu = 0$  e  $\sigma = 1$ ):



# Distribuição Normal padrão

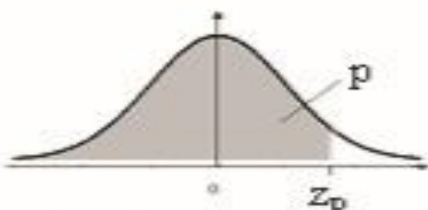
1. Para melhor compreensão, a primeira equação, é a fórmula que padroniza um ponto dado  $Y_i$ .

2. A segunda representa a função de densidade de probabilidade da distribuição normal padrão.

$$Z_i = \frac{Y_i - \mu}{\sigma} \quad \text{Equação 1}$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \text{Equação 2}$$

Tabela I: Distribuição Normal Padrão Acumulada



Fornece  $\Phi(z) = P(-\infty < Z \leq z)$ , para todo  $z$ , de 0,01 em 0,01, desde  $z = 0,00$  até  $z = 3,59$   
A distribuição de  $Z$  é  $\text{Normal}(0;1)$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817



# Distribuição Qui-quadrado



# Distribuição Qui-quadrado:

- É uma distribuição importantíssima para a teoria de inferência estatística.

Seja  $Y_1, Y_2, \dots, Y_p$  variáveis aleatórias independentes, normalmente distribuídas, com média zero e variância 1. Define-se variável aleatória com distribuição qui-quadrado, como:

$$\chi_p^2 = Y_1^2 + Y_2^2 + \dots + Y_p^2$$

- Onde  $p$  é o parâmetro da função de densidade, denominado grau de liberdade. (sendo indicado por  $\phi$  (lê-se fi)). Portanto, a curva de distribuição não é simétrica e para aproximar a simetria devemos aumentar os graus de liberdade, **fazer  $\phi$  crescer até certo ponto para ter confiança nos dados e aproximar essa distribuição de uma normal.**

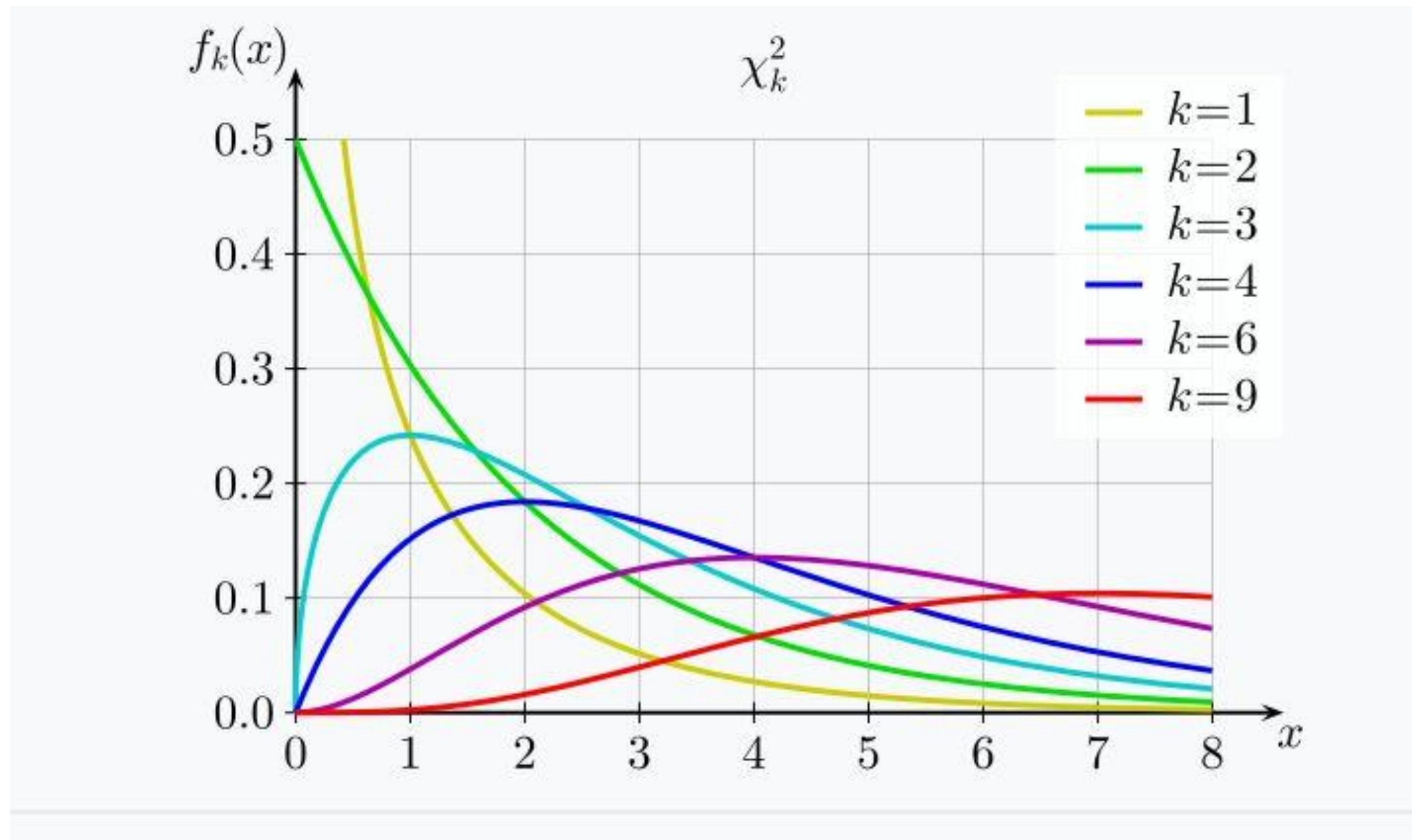
Pode-se demonstrar que a média de uma distribuição qui-quadrado é igual ao grau de liberdade, e que a variância é igual ao dobro do número de graus de liberdade. Assim:

$$E(\chi_\phi^2) = \mu(\chi_\phi^2) = \phi$$

$$Var(\chi_\phi^2) = \sigma^2(\chi_\phi^2) = 2\phi$$

# Distribuição Qui-quadrado:

- Quando temos número de graus de liberdade suficiente, a distribuição  $\chi^2$  representa bem uma distribuição normal
  - Grau de liberdade:  $n-1$  (amostras)



# Distribuição Qui-Quadrado ( $\chi^2$ )

FDP:

$$f(\chi^2, \varphi) = c \cdot \chi^{2\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{\chi^2}{2}} \quad c = \frac{1}{\Gamma\left(\frac{\varphi}{2}\right) \cdot 2^{\frac{\varphi}{2}}}$$

**c** é uma constante dependente de **φ** e determinada pela condição em que a área sob a curva de probabilidade é igual a um.

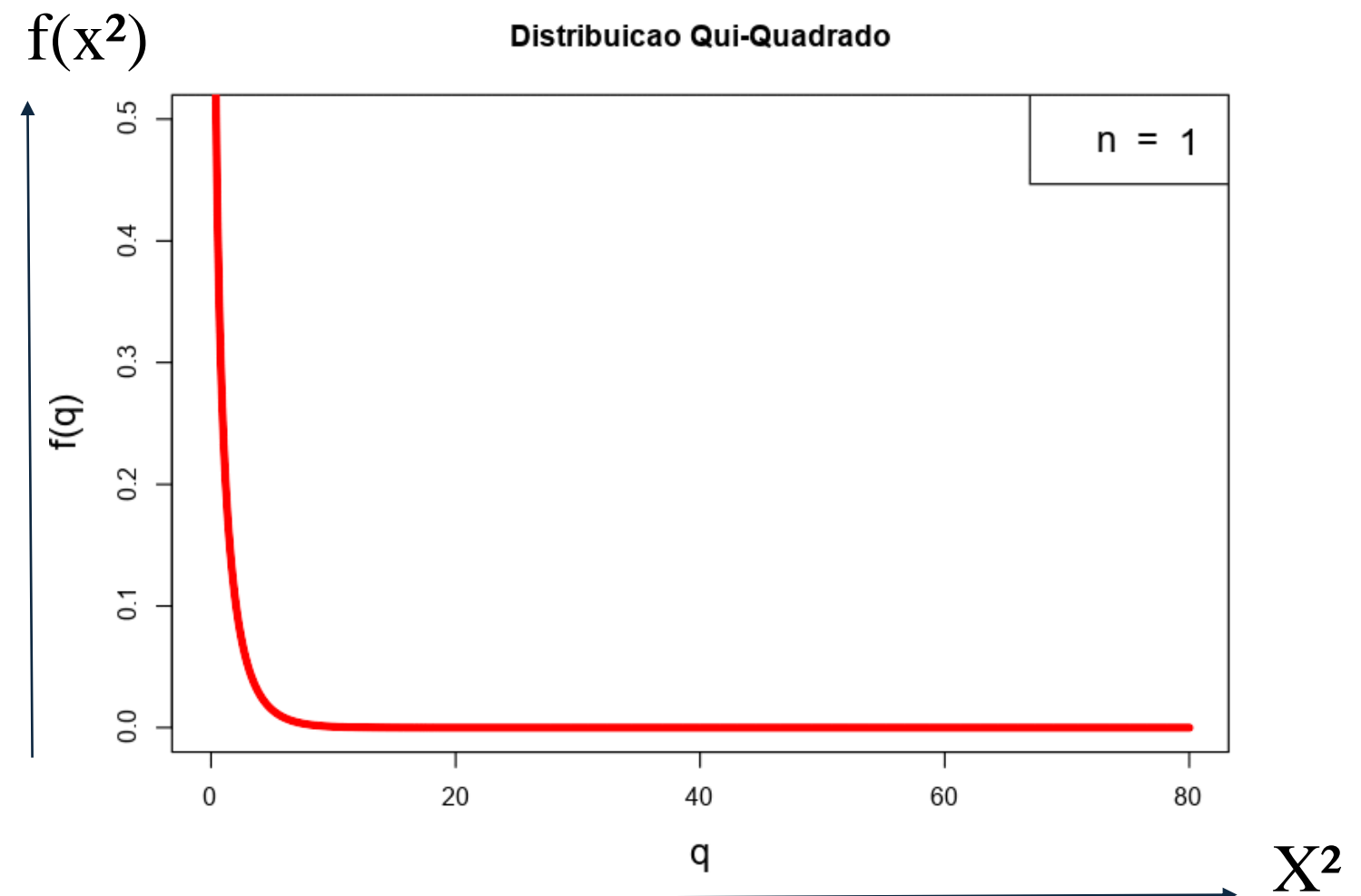
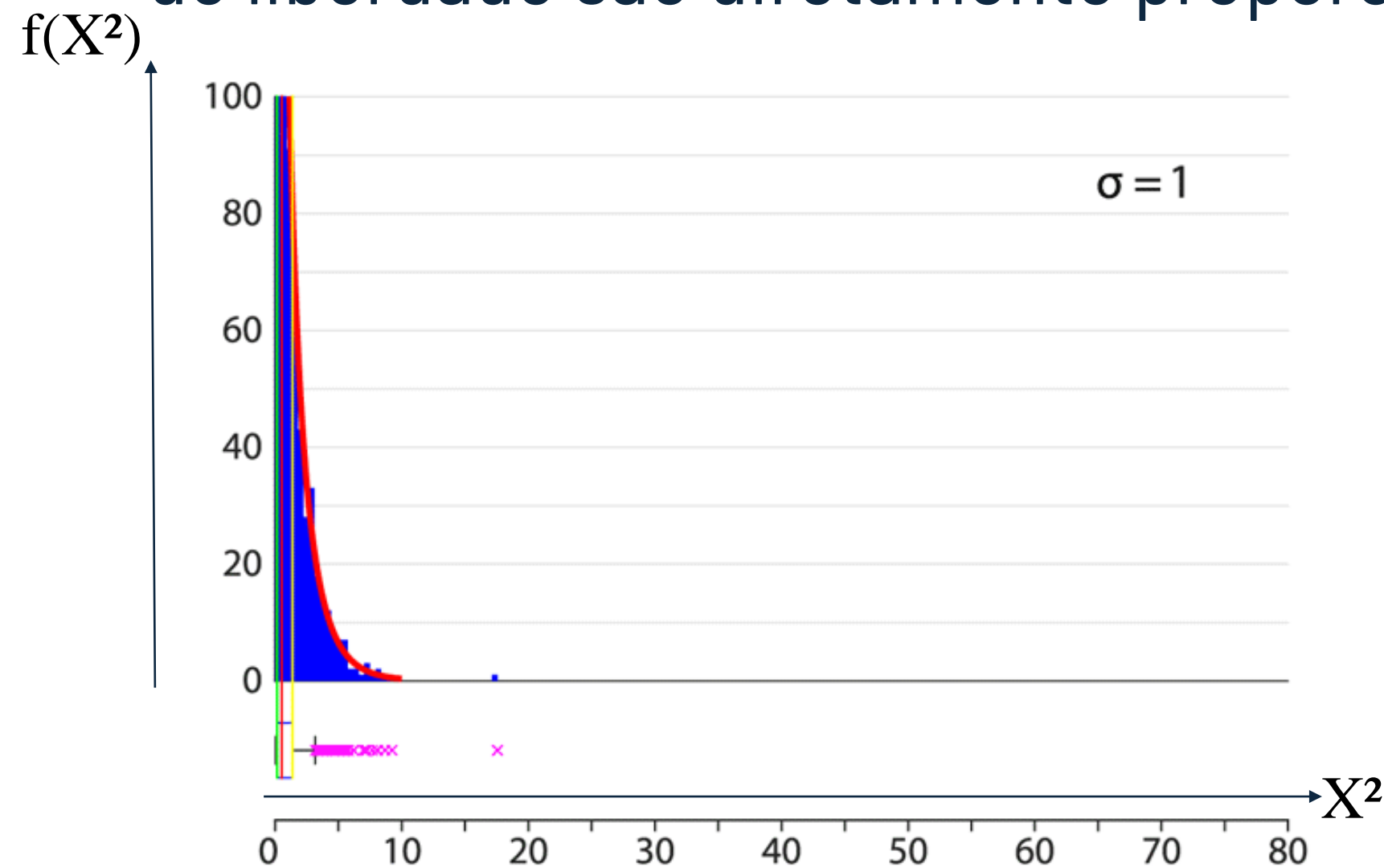
**φ** (lê-se fi) é um parâmetro da função densidade denominado grau de liberdade.

$$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, 0 < x < \infty$$



# Distribuição Qui-Quadrado ( $\chi^2$ )

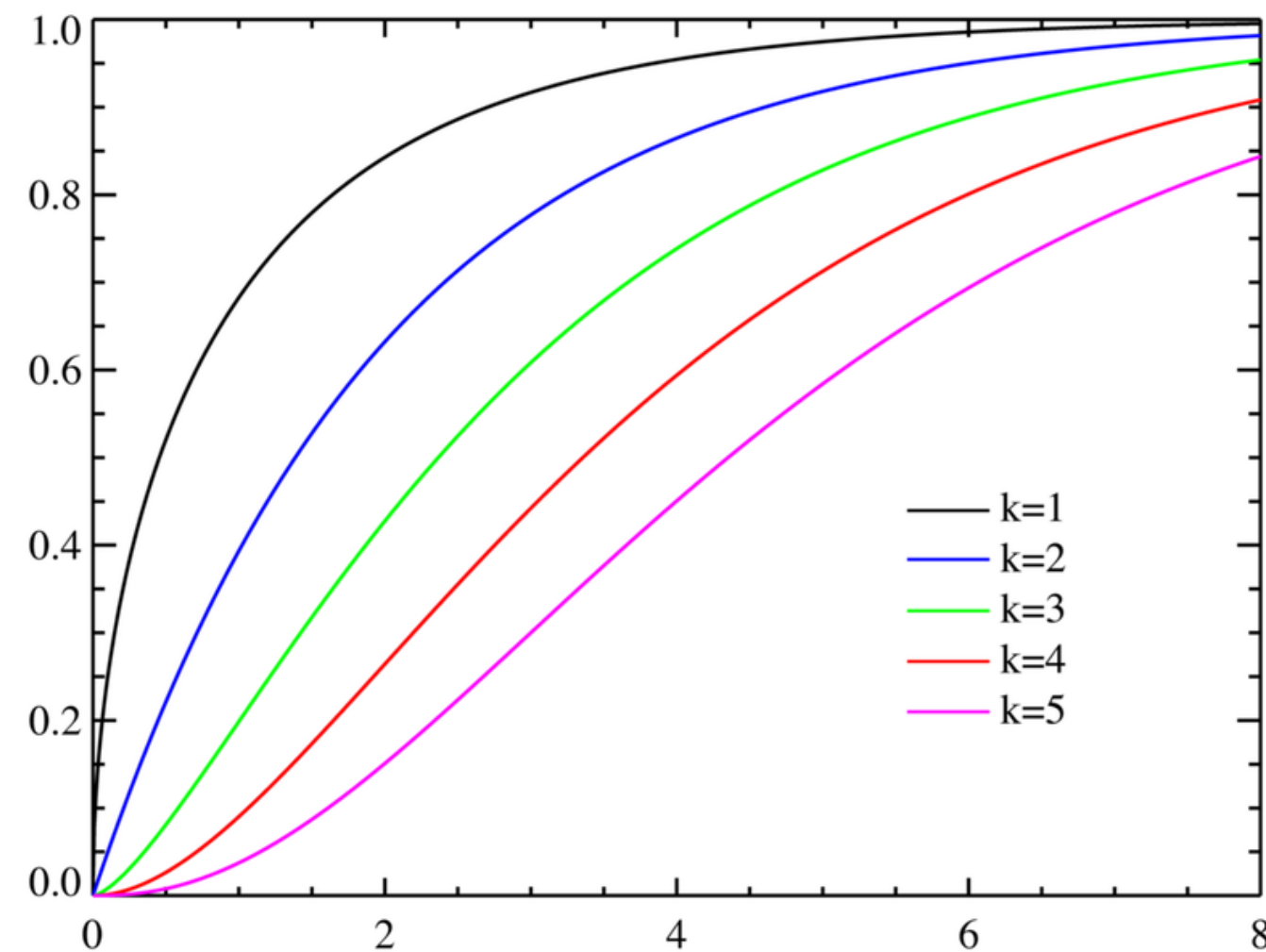
A probabilidade da distribuição qui-quadrado não é simétrica como a da distribuição normal, para aumentar seu estado de simetria é necessário aumentar o seu grau de liberdade, portanto a relação entre simetria e grau de liberdade são diretamente proporcionais.



# Distribuição Qui-Quadrado ( $\chi^2$ )

Gráfico da FDA:

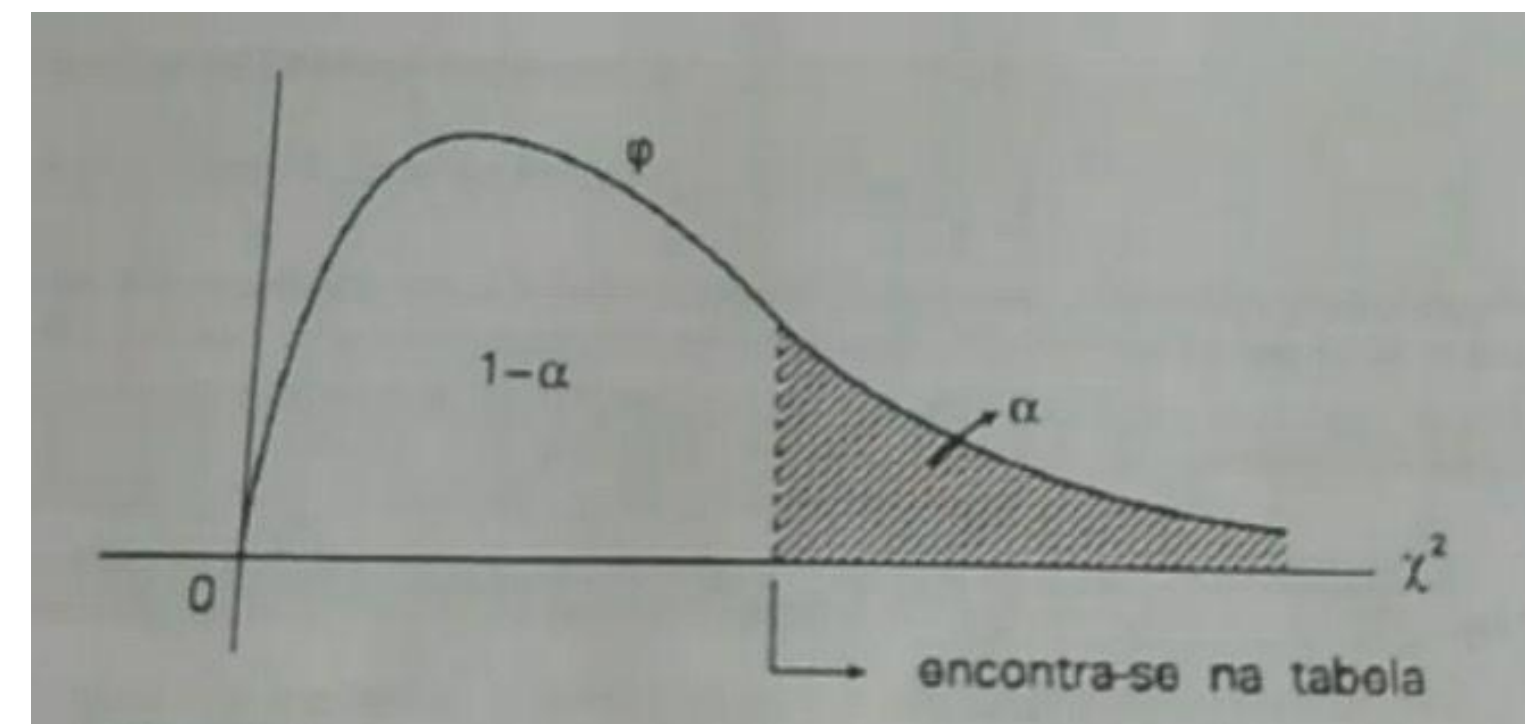
$$\frac{1}{\Gamma(k/2)} \gamma\left(\frac{k}{2}, \frac{x}{2}\right)$$





# DISTRIBUIÇÃO QUI-QUADRADO

- Curva é assimétrica;
- Gráfico começa no 0, tem um pico;
- Decresce a curva com cauda à direita;
- Área abaixo do  $\phi$  sempre vai ser de uma unidade;
- Queremos encontrar área à direita da cauda;
- O que calculamos é o  $\alpha$  que encontramos usando uma tabela;
- Vamos utilizar a distribuição  $\chi^2$  com cauda à direita;



# DISTRIBUIÇÃO QUI- QUADRADO

- Queremos procurar na tabela os valores de  $\phi$ [linha] e  $\alpha$ [coluna], sendo que  $\alpha$  é a área à direita do ponto, procuramos um valor entre  $\phi$  e  $\alpha$ , e por fim, a interseção entre as duas determina o valor da tabela que precisamos.

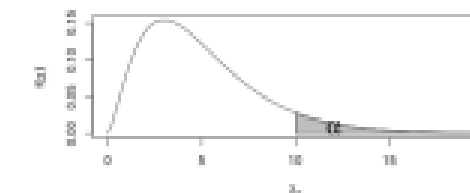


Tabela 1 Tabela de qui-quadrado considerando  $P(\chi^2 > \chi_\alpha^2)$

v	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.001
1	1.3233	1.6424	2.0723	2.7055	3.8415	5.0239	6.6349	7.8794	10.8276
2	2.7726	3.2189	3.7942	4.6052	5.9915	7.3778	9.2103	10.5966	13.8155
3	4.1083	4.6416	5.3170	6.2514	7.8147	9.3484	11.3449	12.8382	16.2662
4	5.3853	5.9886	6.7449	7.7794	9.4877	11.1433	13.2767	14.8603	18.4668
5	6.6257	7.2893	8.1152	9.2364	11.0705	12.8325	15.0863	16.7496	20.5150
6	7.8408	8.5581	9.4461	10.6446	12.5916	14.4494	16.8119	18.5476	22.4577
7	9.0371	9.8032	10.7479	12.0170	14.0671	16.0128	18.4753	20.2777	24.3219
8	10.2189	11.0301	12.0271	13.3616	15.5073	17.5345	20.0902	21.9550	26.1245
9	11.3888	12.2421	13.2880	14.6837	16.9190	19.0228	21.6660	23.5894	27.8772
10	12.5489	13.4420	14.5339	15.9872	18.3070	20.4832	23.2093	25.1882	29.5883
11	13.7007	14.6314	15.7671	17.2750	19.6751	21.9200	24.7250	26.7568	31.2641
12	14.8454	15.8120	16.9893	18.5493	21.0261	23.3367	26.2170	28.2995	32.9095
13	15.9839	16.9848	18.2020	19.8119	22.3620	24.7356	27.6882	29.8195	34.5282
14	17.1169	18.1508	19.4062	21.0641	23.6848	26.1189	29.1412	31.3193	36.1233
15	18.2451	19.3107	20.6030	22.3071	24.9958	27.4884	30.5779	32.8013	37.6973
16	19.3689	20.4651	21.7931	23.5418	26.2962	28.8454	31.9999	34.2672	39.2524
17	20.4887	21.6146	22.9770	24.7690	27.5871	30.1910	33.4087	35.7185	40.7902
18	21.6049	22.7595	24.1555	25.9894	28.8693	31.5264	34.8053	37.1565	42.3124
19	22.7178	23.9004	25.3289	27.2036	30.1435	32.8523	36.1909	38.5823	43.8202
20	23.8277	25.0375	26.4976	28.4120	31.4104	34.1696	37.5662	39.9968	45.3147
21	24.9348	26.1711	27.6620	29.6151	32.6706	35.4789	38.9322	41.4011	46.7970
22	26.0393	27.3015	28.8225	30.8133	33.9244	36.7807	40.2894	42.7957	48.2679
23	27.1413	28.4288	29.9792	32.0069	35.1725	38.0756	41.6384	44.1813	49.7282
24	28.2412	29.5533	31.1325	33.1962	36.4150	39.3641	42.9798	45.5585	51.1786
25	29.3389	30.6752	32.2825	34.3816	37.6525	40.6465	44.3141	46.9279	52.6197
26	30.4346	31.7946	33.4295	35.5632	38.8851	41.9232	45.6417	48.2899	54.0520
27	31.5284	32.9117	34.5736	36.7412	40.1133	43.1945	46.9629	49.6449	55.4760
28	32.6205	34.0266	35.7150	37.9159	41.3371	44.4608	48.2782	50.9934	56.8923
29	33.7109	35.1394	36.8538	39.0875	42.5570	45.7223	49.5879	52.3356	58.3012
30	34.7997	36.2502	37.9903	40.2560	43.7730	46.9792	50.8922	53.6720	59.7031
31	35.8871	37.3591	39.1244	41.4217	44.9853	48.2319	52.1914	55.0027	61.0983
32	36.9730	38.4663	40.2563	42.5847	46.1943	49.4804	53.4858	56.3281	62.4872
33	38.0575	39.5718	41.3861	43.7452	47.3999	50.7251	54.7755	57.6484	63.8701
34	39.1408	40.6756	42.5140	44.9032	48.6024	51.9660	56.0609	58.9639	65.2472

# TESTE $\chi^2$

- Teste não paramétrico
- Comparação entre frequências
- Dados nominais, ordinais e intervalares
- Amostras independentes

# TESTE $\chi^2$

- Vamos supor que voce esteja organizando um ciclo de palestras e gostaria de saber quais temas mais interessariam as pessoas. Uma pesquisa com 100 sujeitos resultou nos dados abaixo:

Teste Qui-quadrado – 1 critério

Temas	Frequência
A	38
B	25
C	16
D	12
E	9

- Estatisticamente, essa diferença de frequência entre temas é significativa, podemos afirmar que o tema predomina os demais?
- Qual seria a frequência esperada?

# TESTE $\chi^2$

Teste Qui-quadrado – 1 critério

Temas	Freq ( $f_o$ )	Frq esp ( $f_e$ )	$F_o - f_e$	$(F_o - f_e)^2$	$(F_o - f_e)^2 / f_e$
A	38	20	18	324	16,2
B	25	20	5	25	1,25
C	16	20	-4	16	0,8
D	12	20	-8	64	3,2
E	9	20	-11	121	6,05

- $\chi^2 = 27,6$
- O que esse número significa?
- Para isso, precisamos de outro cálculo, o grau de liberdade.

# TESTE $\chi^2$

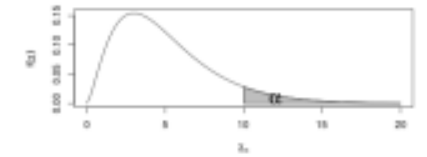


Tabela 1 Tabela de qui-quadrado considerando  $P(\chi^2 > \chi^2_{\alpha})$

- gl = k-1 (sendo k o número de categorias na distribuição de frequências observada);
- gl = 5-1=4
- $\chi^2$  crítico = 9,487 (sig = 0,05) → valor tabelado;

Então, o que faço com o valor crítico e com o valor de 27,5 calculado?

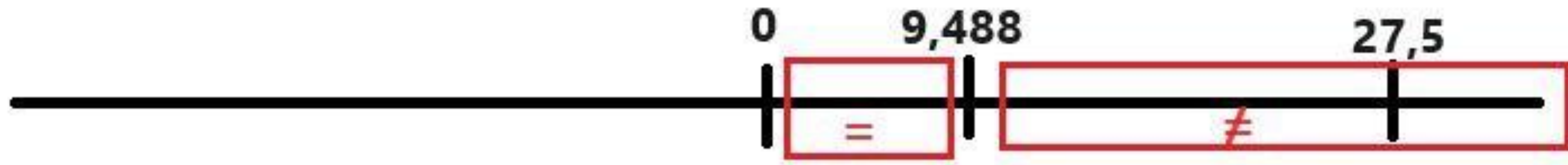
v	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.001
1	1.3233	1.6424	2.0723	2.7055	3.8415	5.0239	6.6349	7.8794	10.8276
2	2.7726	3.2189	3.7942	4.6052	5.9915	7.3778	9.2103	10.5966	13.8155
3	4.1083	4.6416	5.3170	6.2514	7.8147	9.3484	11.3449	12.8382	16.2662
4	5.3853	5.9886	6.7449	7.7794	9.4877	11.1433	13.2767	14.8603	18.4668
5	6.6257	7.2893	8.1152	9.2364	11.0705	12.8325	15.0863	16.7496	20.5150
6	7.8408	8.5581	9.4461	10.6446	12.5916	14.4494	16.8119	18.5476	22.4577
7	9.0371	9.8032	10.7479	12.0170	14.0671	16.0128	18.4753	20.2777	24.3219
8	10.2189	11.0301	12.0271	13.3616	15.5073	17.5345	20.0902	21.9550	26.1245
9	11.3888	12.2421	13.2880	14.6837	16.9190	19.0228	21.6660	23.5894	27.8772
10	12.5489	13.4420	14.5339	15.9872	18.3070	20.4832	23.2093	25.1882	29.5883
11	13.7007	14.6314	15.7671	17.2750	19.6751	21.9200	24.7250	26.7568	31.2641
12	14.8454	15.8120	16.9893	18.5493	21.0261	23.3367	26.2170	28.2995	32.9095
13	15.9839	16.9848	18.2020	19.8119	22.3620	24.7356	27.6882	29.8195	34.5282
14	17.1169	18.1508	19.4062	21.0641	23.6848	26.1189	29.1412	31.3193	36.1233
15	18.2451	19.3107	20.6030	22.3071	24.9958	27.4884	30.5779	32.8013	37.6973
16	19.3689	20.4651	21.7931	23.5418	26.2962	28.8454	31.9999	34.2672	39.2524
17	20.4887	21.6146	22.9770	24.7690	27.5871	30.1910	33.4087	35.7185	40.7902
18	21.6049	22.7595	24.1555	25.9894	28.8693	31.5264	34.8053	37.1565	42.3124
19	22.7178	23.9004	25.3289	27.2036	30.1435	32.8523	36.1909	38.5823	43.8202
20	23.8277	25.0375	26.4976	28.4120	31.4104	34.1696	37.5662	39.9968	45.3147
21	24.9348	26.1711	27.6620	29.6151	32.6706	35.4789	38.9322	41.4011	46.7970
22	26.0393	27.3015	28.8225	30.8133	33.9244	36.7807	40.2894	42.7957	48.2679
23	27.1413	28.4288	29.9792	32.0069	35.1725	38.0756	41.6384	44.1813	49.7282
24	28.2412	29.5533	31.1325	33.1962	36.4150	39.3641	42.9798	45.5585	51.1786
25	29.3389	30.6752	32.2825	34.3816	37.6525	40.6465	44.3141	46.9279	52.6197
26	30.4346	31.7946	33.4295	35.5632	38.8851	41.9232	45.6417	48.2899	54.0520
27	31.5284	32.9117	34.5736	36.7412	40.1133	43.1945	46.9629	49.6449	55.4760
28	32.6205	34.0266	35.7150	37.9159	41.3371	44.4608	48.2782	50.9934	56.8923
29	33.7109	35.1394	36.8538	39.0875	42.5570	45.7223	49.5879	52.3356	58.3012
30	34.7997	36.2502	37.9903	40.2560	43.7730	46.9792	50.8922	53.6720	59.7031
31	35.8871	37.3591	39.1244	41.4217	44.9853	48.2319	52.1914	55.0027	61.0983
32	36.9730	38.4663	40.2563	42.5847	46.1943	49.4804	53.4858	56.3281	62.4872
33	38.0575	39.5718	41.3861	43.7452	47.3999	50.7251	54.7755	57.6484	63.8701
34	39.1408	40.6756	42.5140	44.9032	48.6024	51.9660	56.0609	58.9639	65.2472



# TESTE $\chi^2$

Existe diferença estatisticamente significativa entre as frequências para cada um dos temas;

- Se o  $\chi^2 = 5$ , estatisticamente falando, diferenças aparentes na amostra não seriam relevantes



# Distribuição T de Student



# A distribuição T de Student

A distribuição T foi introduzida por William Sealy Gosset sob o pseudônimo “Student” em 1908. Ele trabalhava na empresa de cervejaria Guinness e desenvolveu esta distribuição para lidar com problemas estatísticos envolvendo pequenas amostras.



William Sealy Gosset. [Wikimedia](#)

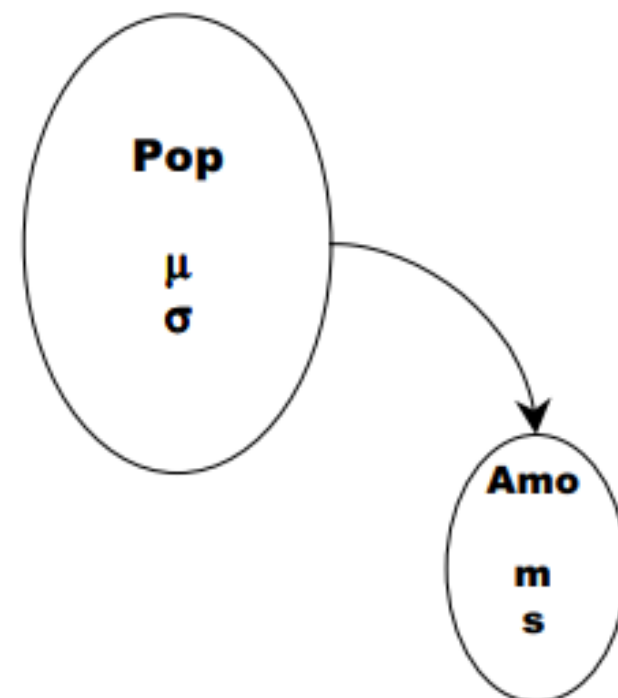


# Distribuição T de Student

## Quando usar a distribuição T?

Se os valores da média e desvio padrão,  $\mu$  e  $\sigma$ , são conhecidos, utiliza-se a distribuição normal padrão.

Entretanto se os valores da média e desvio padrão,  $\mu$  e  $\sigma$ , não são conhecidos, e fazemos inferências sobre uma população a partir das estimativas da média e do desvio padrão, ou seja, obtidas nas amostras, utiliza-se a distribuição “t”.



**m: estimador do parâmetro  $\mu$**

**s: estimador do parâmetro  $\sigma$**

**$\mu = m \pm$  erro de amostragem**

# Distribuição T de Student

Suponha  $Z$ , uma variável aleatória de **distribuição normal padrão**, ou seja, com **média 0 e variância 1**, e  $V$ , uma variável aleatória com **distribuição Qui-quadrado** com  $\nu$  graus de liberdade. Se  $Z$  e  $V$  são independentes, então a distribuição da variável aleatória  $t$  será:

$$t = \frac{Z}{\sqrt{V/\nu}}$$

Grau de liberdade, também pode ser encontrado como  $\phi$   
 $\phi = n - 1$ , sendo  $n$  o tamanho da amostra

Uma variável aleatória contínua  $t$  tem distribuição t de Student com  $\phi$  graus de liberdade, denotada por  $t_\phi$ , se sua função densidade for dada por:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

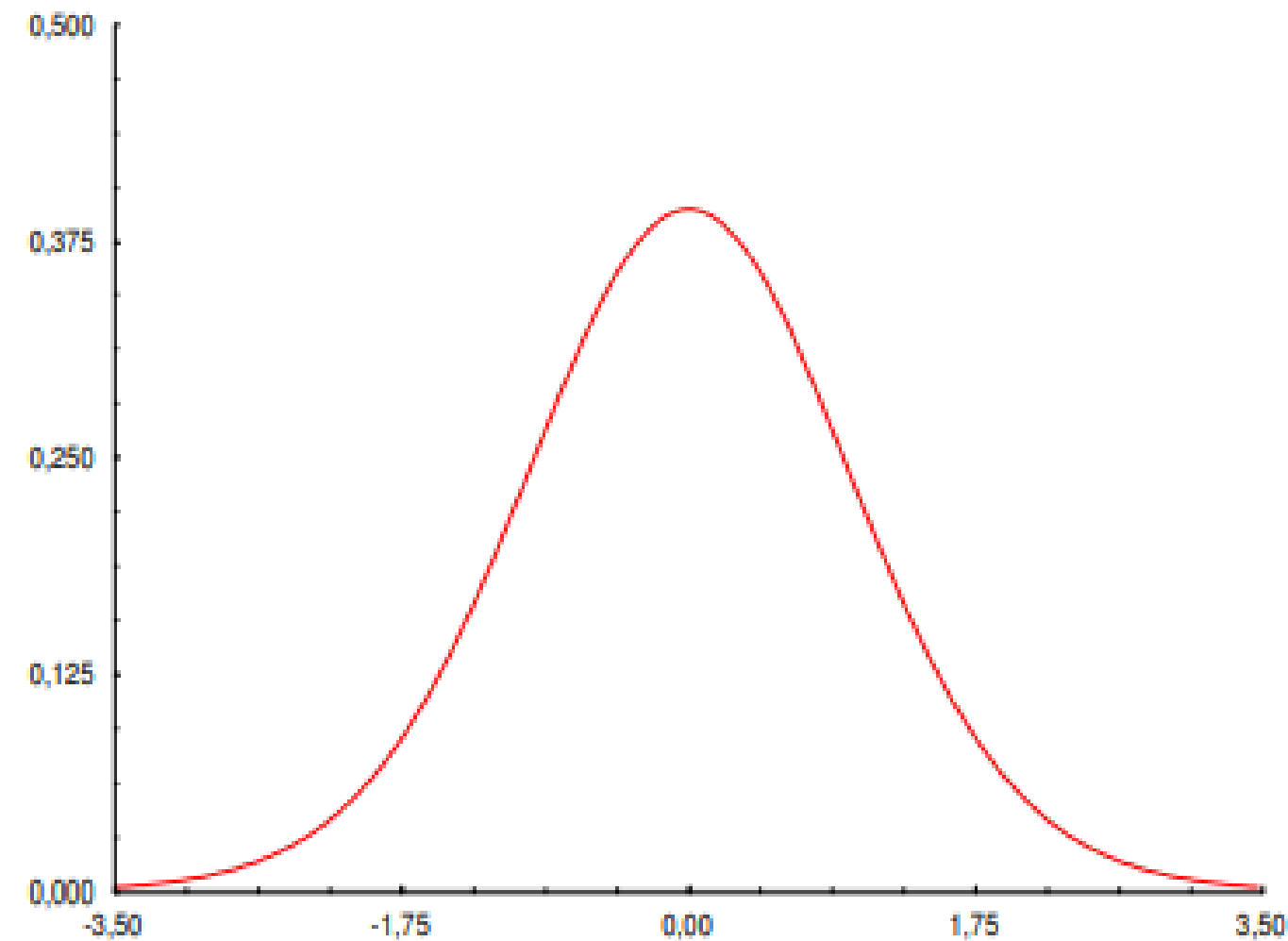
Função Gama

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



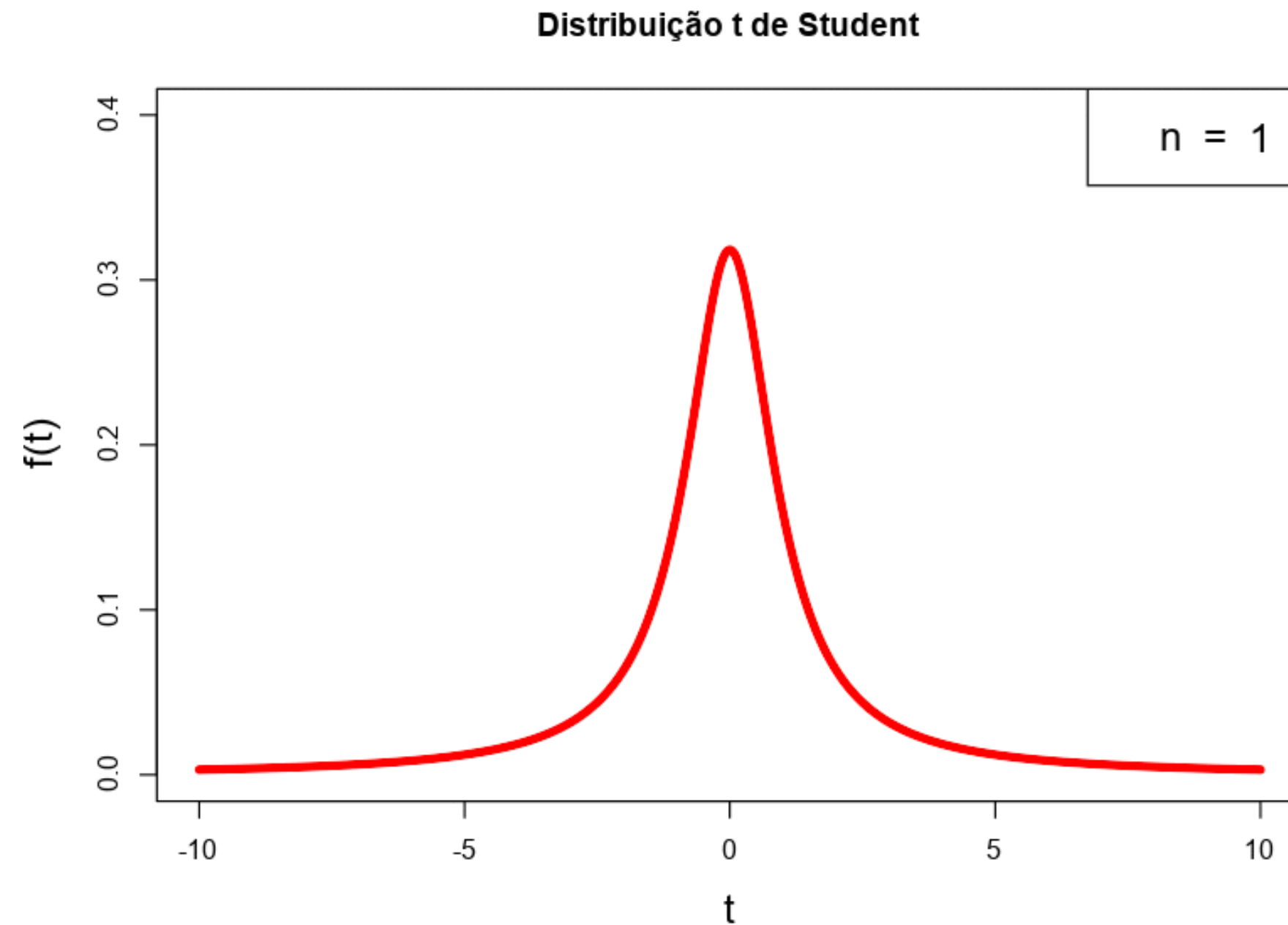
# Distribuição T

Graficamente temos que a área total é necessariamente 1 (100%), quando aplicamos a media, moda e mediana da distribuição T ambas tenderão a 0.



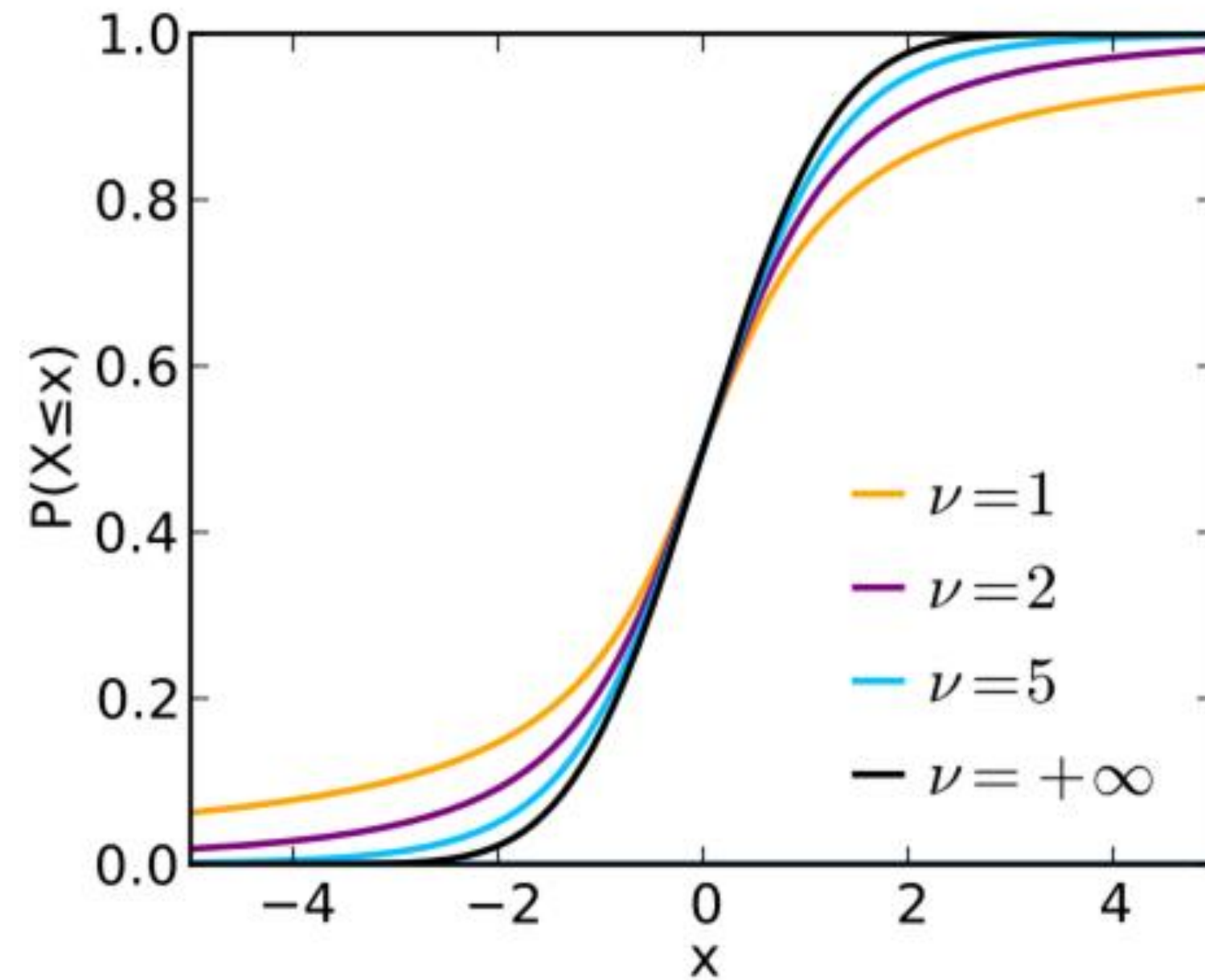
# Distribuição T

Grafico da FDP:



# Distribuição T

Gráfico da FDA:



# Distribuição T de Student

## Graus de Liberdade:

Assim como Guinness desenvolveu, é indicado usar a distribuição “T de Student” para amostras com tamanho inferior a 30.

Necessitamos de alguns cálculos básicos para trabalhar com a distribuição:

## Cálculo da Variância $v = \varphi$

$\frac{\nu}{\nu-2}$	se $\nu > 2$
$\infty$	se $1 < \nu \leq 2$
Indefinida	se $0 < \nu \leq 1$

O desvio padrão é a raiz da variância, então:

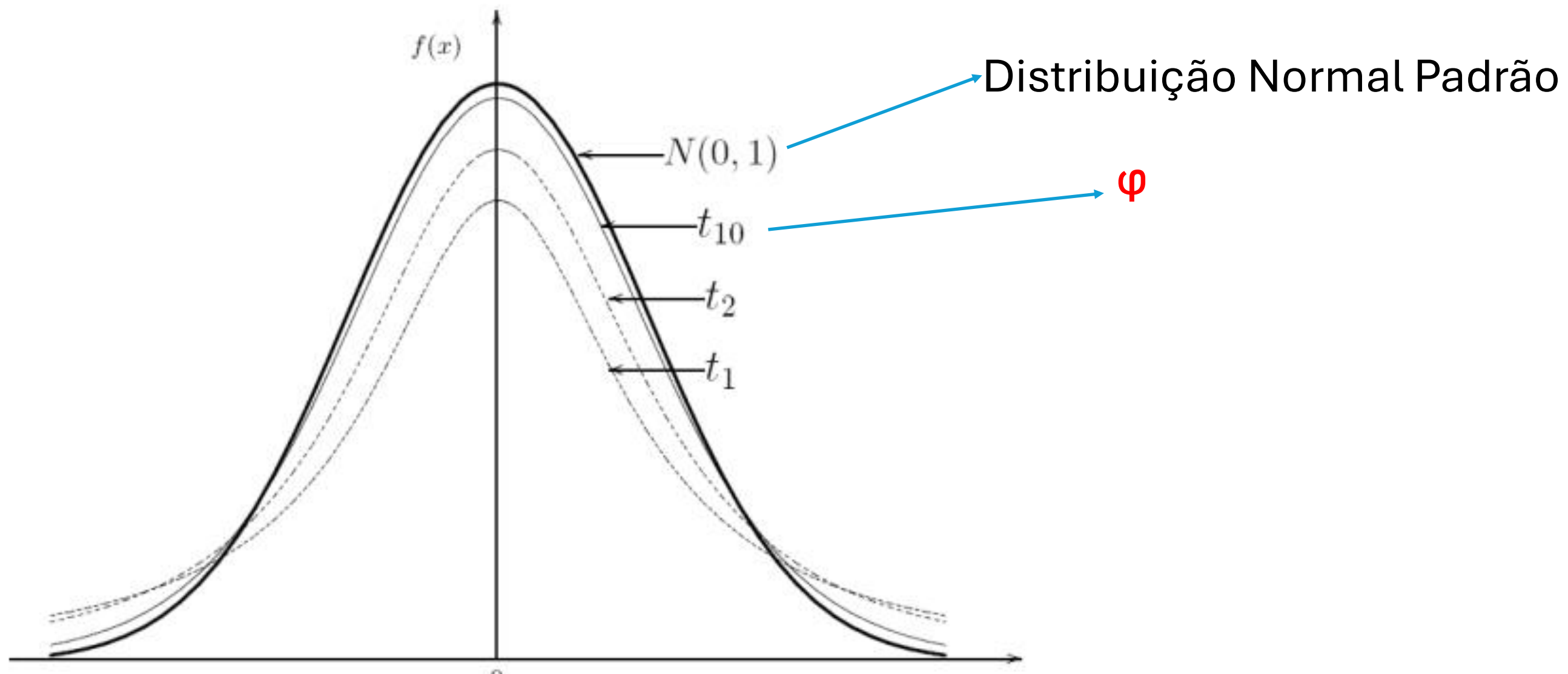
$$\sigma(t_4) = \sqrt{\frac{4}{4-2}} = 1,41 \quad \sigma(t_{35}) = \sqrt{\frac{35}{35-2}} = 1,03 \quad \sigma(t_{60}) = \sqrt{\frac{60}{60-2}} = 1,02$$

$v = t_n$

$\varphi$

# Distribuição T

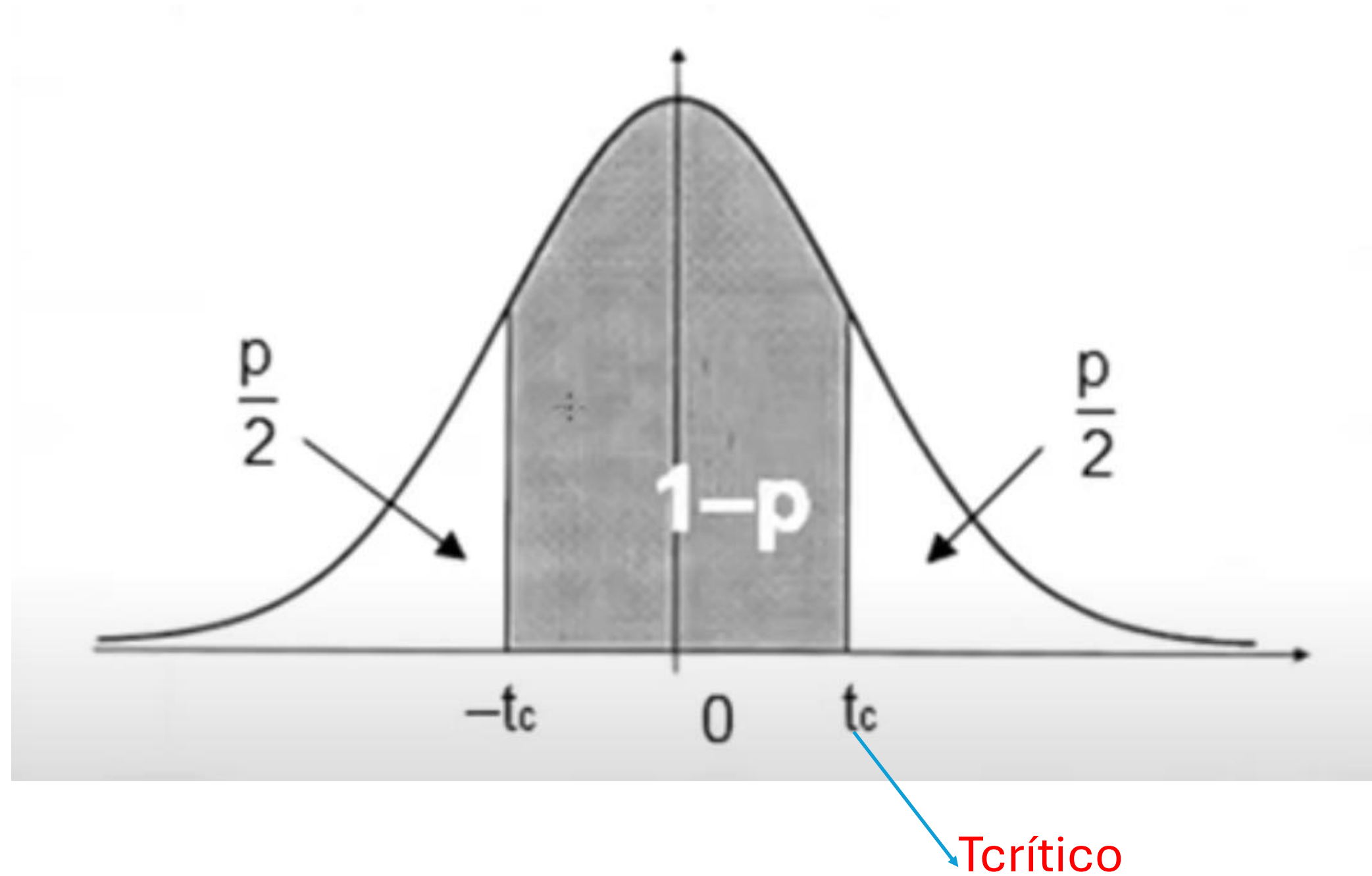
Graficamente temos uma comparação entre  $\varphi$  e DNP





# DISTRIBUIÇÃO T DE STUDENT

Para Entender a tabela:





# DISTRIBUIÇÃO T DE STUDENT

Distribuição <i>t</i> -Student : Valores $t_c$ tais que $P(-t_c \leq t \leq t_c) = 1 - p$																	
	p->90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	2%	1%	0.2%	0.1%		
Graus de liberdade	1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,894	31,821	63,656	318,289	636,578	1
	2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925	22,328	31,600	2
	3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841	10,214	12,924	3
	4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,999	3,747	4,604	7,173	8,610	4
	5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,757	3,365	4,032	5,894	6,869	5
	6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707	5,208	5,959	6
	7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499	4,785	5,408	7
	8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355	4,501	5,041	8
	9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250	4,297	4,781	9
	10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169	4,144	4,587	10
	11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106	4,025	4,437	11
	12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055	3,930	4,318	12
	13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012	3,852	4,221	13
	14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977	3,787	4,140	14
	15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,249	2,602	2,947	3,733	4,073	15
	16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921	3,686	4,015	16
	17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898	3,646	3,965	17
	18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878	3,610	3,922	18
	19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861	3,579	3,883	19
	20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845	3,552	3,850	20
	21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831	3,527	3,819	21
	22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819	3,505	3,792	22
	23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807	3,485	3,768	23
	24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797	3,467	3,745	24
	25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,167	2,485	2,787	3,450	3,725	25
	26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779	3,435	3,707	26
	27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,158	2,473	2,771	3,421	3,689	27
	28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763	3,408	3,674	28
	29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756	3,396	3,660	29
	30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750	3,385	3,646	30
	35	0,127	0,255	0,388	0,529	0,682	0,852	1,052	1,306	1,690	2,030	2,133	2,438	2,724	3,340	3,591	35
	40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,123	2,423	2,704	3,307	3,551	40
	50	0,126	0,255	0,388	0,528	0,679	0,849	1,047	1,299	1,676	2,009	2,109	2,403	2,678	3,261	3,496	50
	60	0,126	0,254	0,387	0,527	0,679	0,848	1,045	1,296	1,671	2,000	2,099	2,390	2,660	3,232	3,460	60
	120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,076	2,358	2,617	3,160	3,373	120
	∞	0,126	0,253	0,385	0,524	0,675	0,842	1,036	1,282	1,645	1,960	2,054	2,327	2,576	3,091	3,291	∞



# DISTRIBUIÇÃO T DE STUDENT

Descobrimos o intervalo de confiança, vamos precisar:

Média amostral

$$m = \frac{\sum_{i=1}^n y_i}{n}$$

Quantidade de elementos na amostra

Cada elemento da amostra

Variância amostral

$$s^2 = \frac{\sum (Y_i - m)^2}{n-1}$$

Desvio padrão amostral

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

Intervalo de confiança

$$I.C(\mu) = m \pm t_0 \cdot \frac{s}{\sqrt{n}}$$

$T_c = t_0$  que descobrimos na tabela, tendo o erro e o Grau de liberdade

# DISTRIBUIÇÃO T DE STUDENT

Descobrimos o intervalo de confiança, exemplo:

A amostra 9, 8, 12, 7, 9, 6, 11, 6, 10, 9 foi extraída de uma população Normal. Construa um intervalo de confiança para a média ao nível de 95%.

Média amostral:

$$\frac{9 + 8 + 12 + 7 + 9 + 6 + 11 + 6 + 10 + 9}{10} = 8,7$$

Variância amostral:

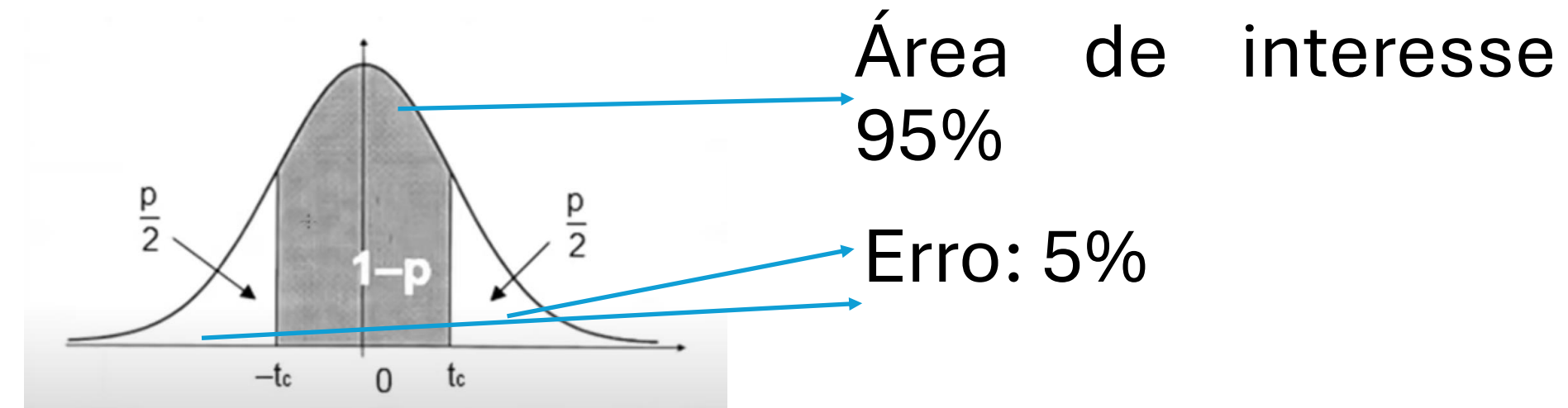
$$\frac{(9-8,7)^2 + (8-8,7)^2 + (12-8,7)^2 + (7-8,7)^2 + (9-8,7)^2 + (6-8,7)^2 + (11-8,7)^2 + (6-8,7)^2 + (10-8,7)^2 + (9-8,7)^2}{10-1} = 4$$

Desvio padrão amostral:

$$\sqrt{4} = 2$$

Grau de liberdade:

$$\varphi = 10 - 1$$



Tendo o Erro de 5% e o Grau de liberdade 9 podemos procurar o  $T_c$  na tabela:

$$T_c = 2,262$$

Intervalo de confiança

$$8,7 \pm 2,262 * \frac{2}{\sqrt{10}} = 8,7 \pm 1,42$$

7,27  
10,12

Isso significa que a probabilidade da media dos dados da população ( $\mu$ ) estar entre [7,27 ; 10,12] é de 95%

# Distribuição F de Snedecor





# Distribuição F de Snedecor e Fisher

A distribuição F de Snedecor foi criada em homenagem ao biólogo e estatístico britânico Ronald Fisher e ao matemático norte-americano George Waddel Snedecor.



Também conhecida como distribuição F, é uma distribuição de probabilidade contínua que surge frequentemente em análises estatísticas, particularmente na análise de variância (ANOVA) e em testes F.

# DISTRIBUIÇÃO F

## Função densidade de probabilidade:

Distribuição F é a razão entre duas variáveis aleatórias independentes com distribuição qui-quadrado, Assim, uma distribuição F com  $\phi_1$  graus de liberdade no numerador, e  $\phi_2$  graus de liberdade no denominador é expressa por:

$$F(\phi_1, \phi_2) = \frac{\frac{\chi_{\phi_1}^2}{\phi_1}}{\frac{\chi_{\phi_2}^2}{\phi_2}}$$

Duas variáveis aleatórias independentes com distribuição qui-quadrado

ou

$$F_{n,m} = \frac{\frac{X}{n}}{\frac{Y}{m}} \quad \begin{matrix} X = \chi_n^2 \\ Y = \chi_m^2 \end{matrix}$$

Uma variável aleatória contínua F tem distribuição F de Snedecor com  $\phi_1$  e  $\phi_2$  graus de liberdade, denotada por  $F_{\phi_1 \phi_2}$ , se sua função densidade for dada por:

$$f(F, \phi_1 : \phi_2) = c \cdot \left(\frac{\phi_1}{\phi_2}\right)^{\frac{\phi_1}{2}} \cdot F^{\left(\frac{\phi_1-1}{2}\right)} \cdot \left(1 + \frac{\phi_1}{\phi_2} \cdot F\right)^{-\left(\frac{\phi_1+\phi_2}{2}\right)}$$

$$\phi_2 = n_2 - 1$$

$$\phi_1 = n_1 - 1$$

$$c = \frac{\Gamma\left(\frac{\phi_1 + \phi_2}{2}\right)}{\Gamma\left(\frac{\phi_1}{2}\right) \cdot \Gamma\left(\frac{\phi_2}{2}\right)}$$

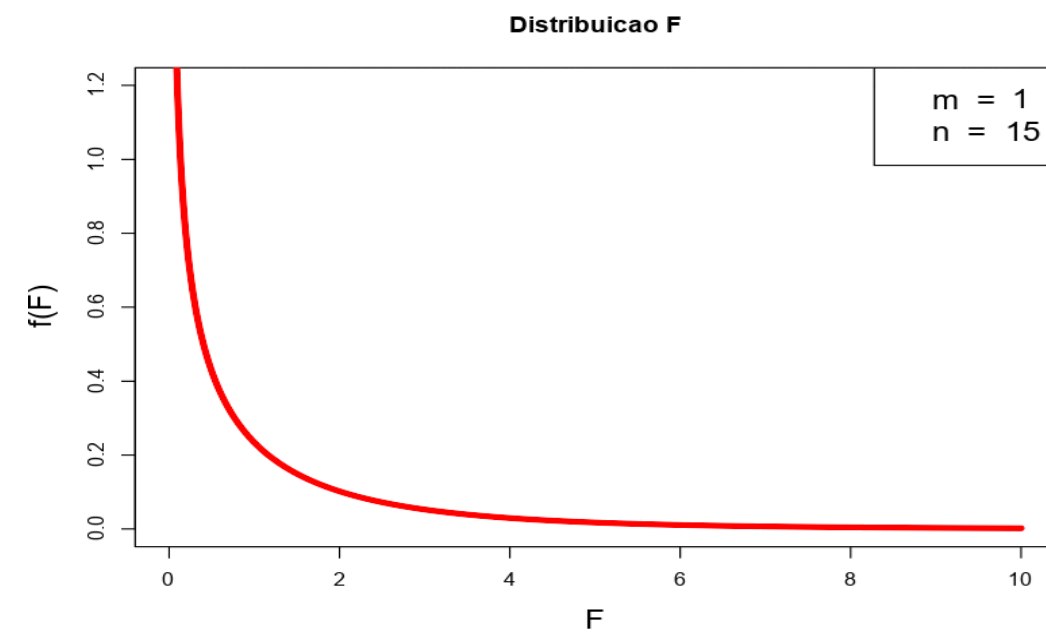
Função gama →

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

$$\text{ou} \quad f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\left(\frac{m}{2}\right)} \frac{x^{\left(\frac{m-2}{2}\right)}}{\left([1 + \left(\frac{m}{n}\right)x]\right)^{\left(\frac{m+n}{2}\right)}}, x \geq 0$$

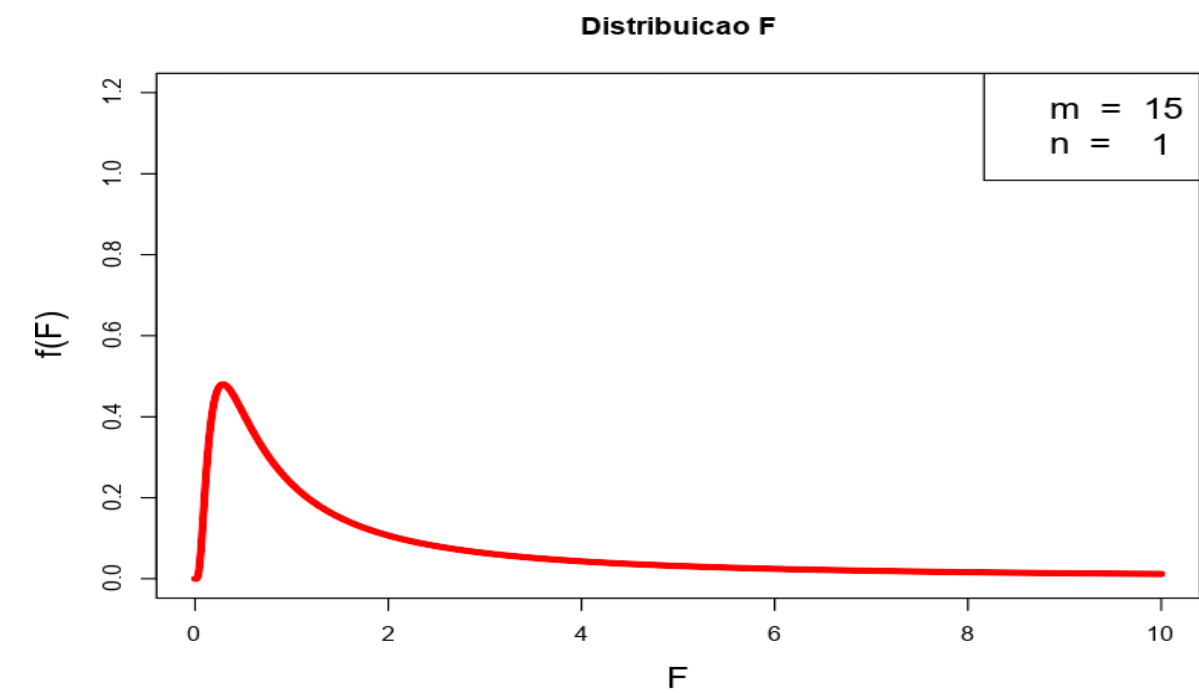
# Distribuição F de Snedecor

Distribuição F com m crescendo  
crescendo



$$F_{n,m} = \frac{\frac{X}{n}}{\frac{Y}{m}}$$

Distribuição F com n



# Distribuição F de Snedecor

Gráfico da FDP:

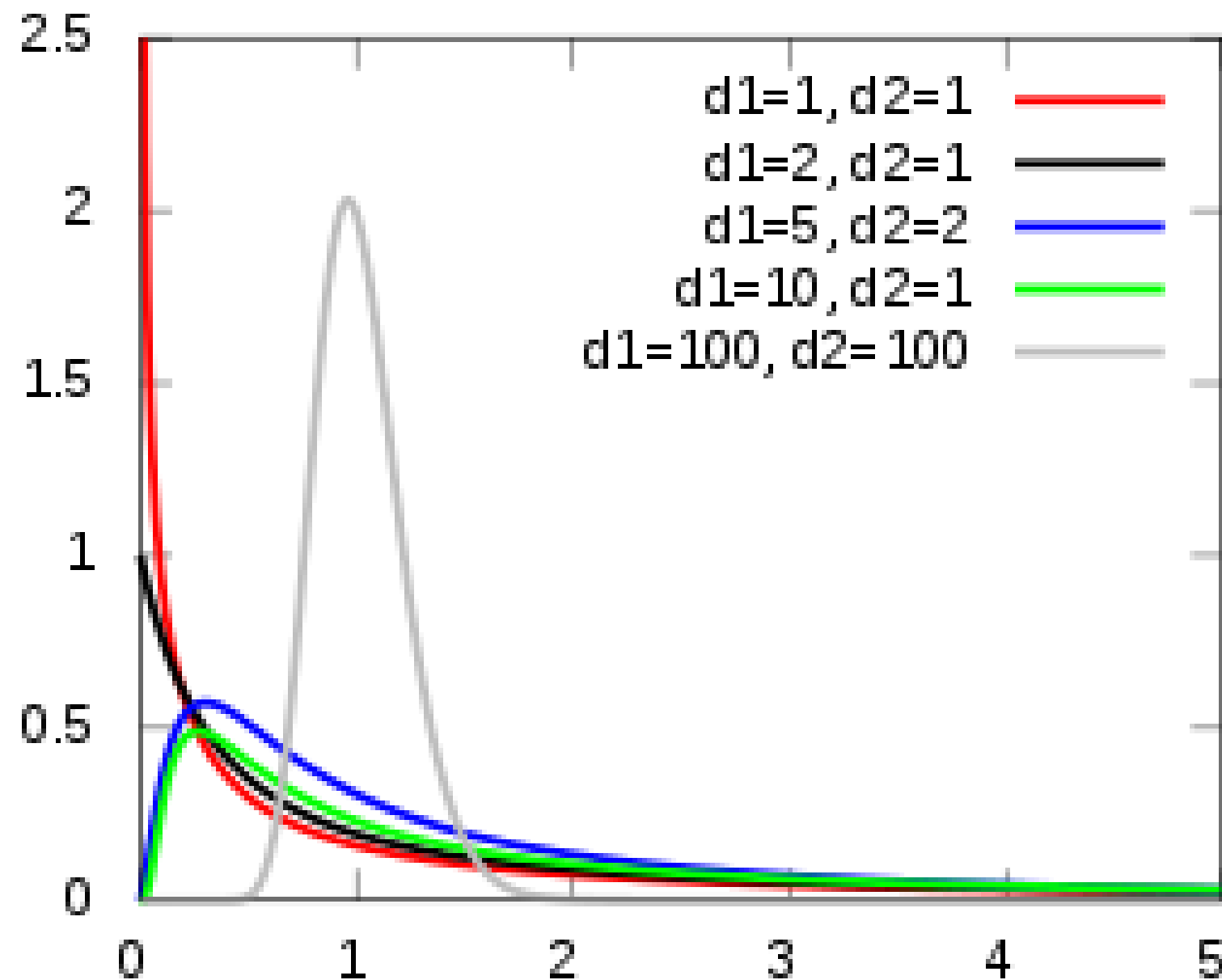
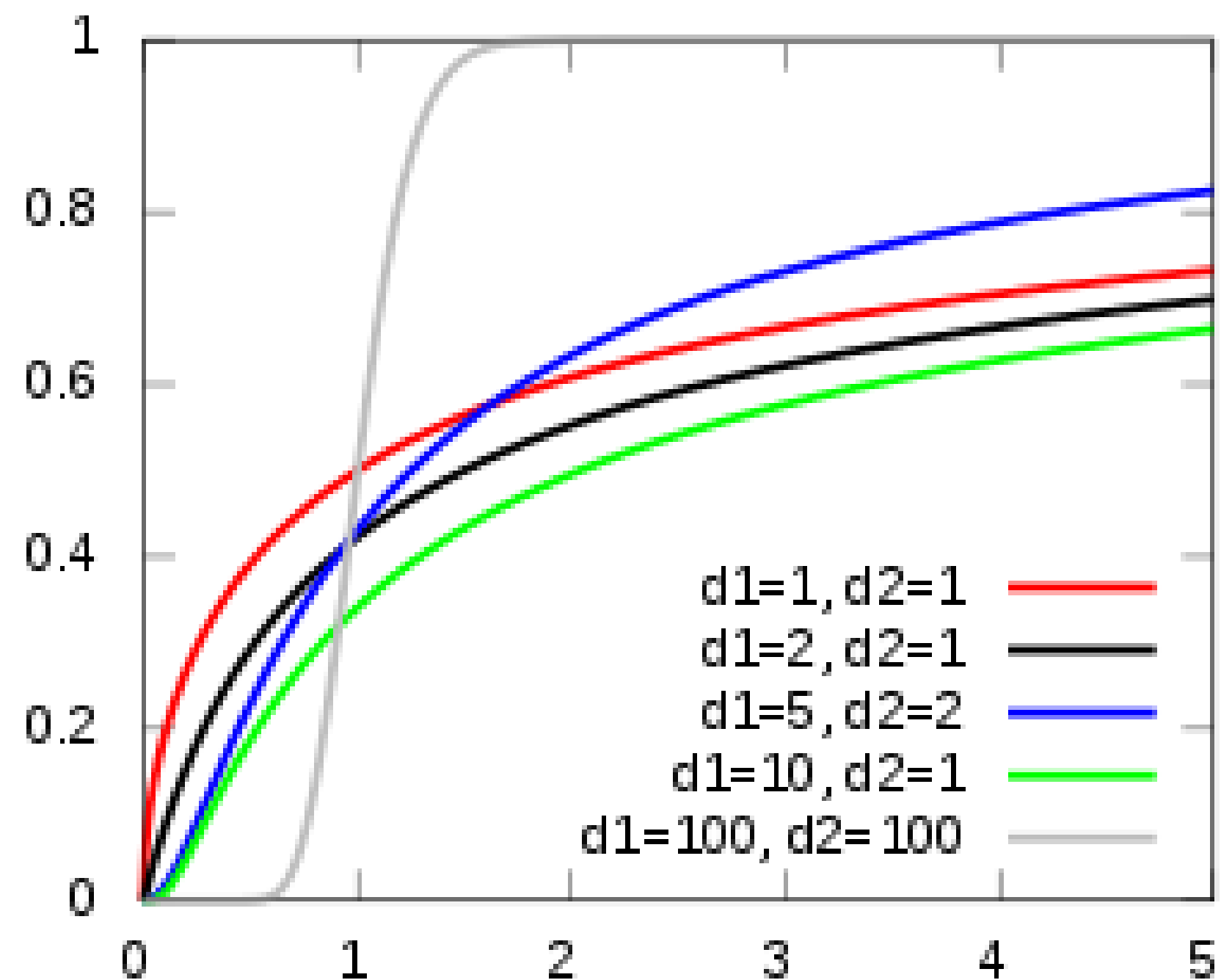


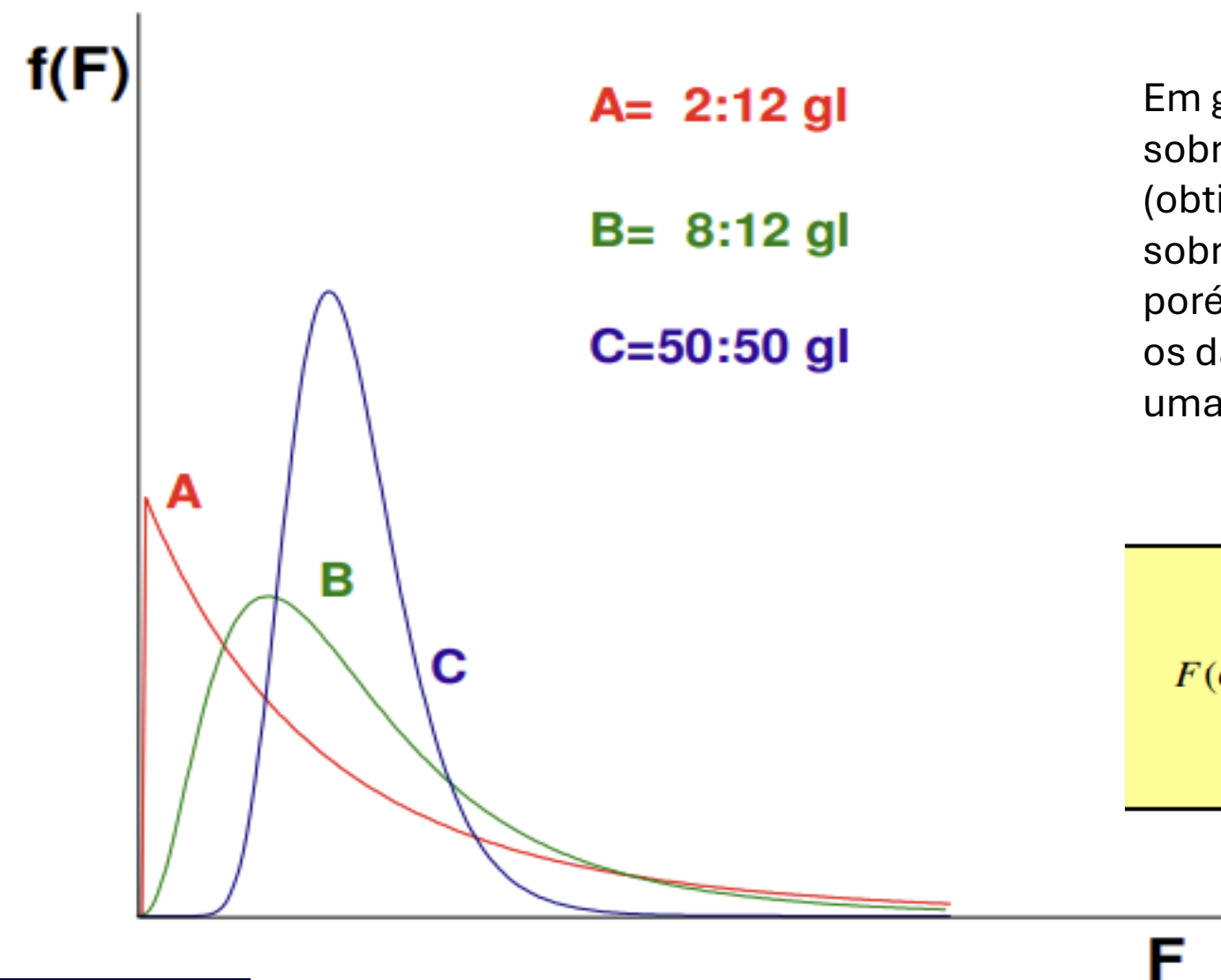
Gráfico da FDA:  $I_{\frac{d_1 x}{d_1 x + d_2}} \left( \frac{d_1}{2}, \frac{d_2}{2} \right)$



# DISTRIBUIÇÃO F

## Primeiros passos e utilização da Distribuição F

Possuindo dois parâmetros: graus de liberdade do numerador e grau de liberdade no denominador, que são denominados, comumente, por  $\phi_1$  e  $\phi_2$  respectivamente, ela encontra-se tabelada para as probabilidades mais utilizadas nos testes de hipóteses: 1%, 5% e 10%. Tal como a distribuição  $\chi^2$ , esta distribuição de probabilidades não apresenta uma forma fixa, mas sim variável de acordo com os graus de liberdade envolvidos:

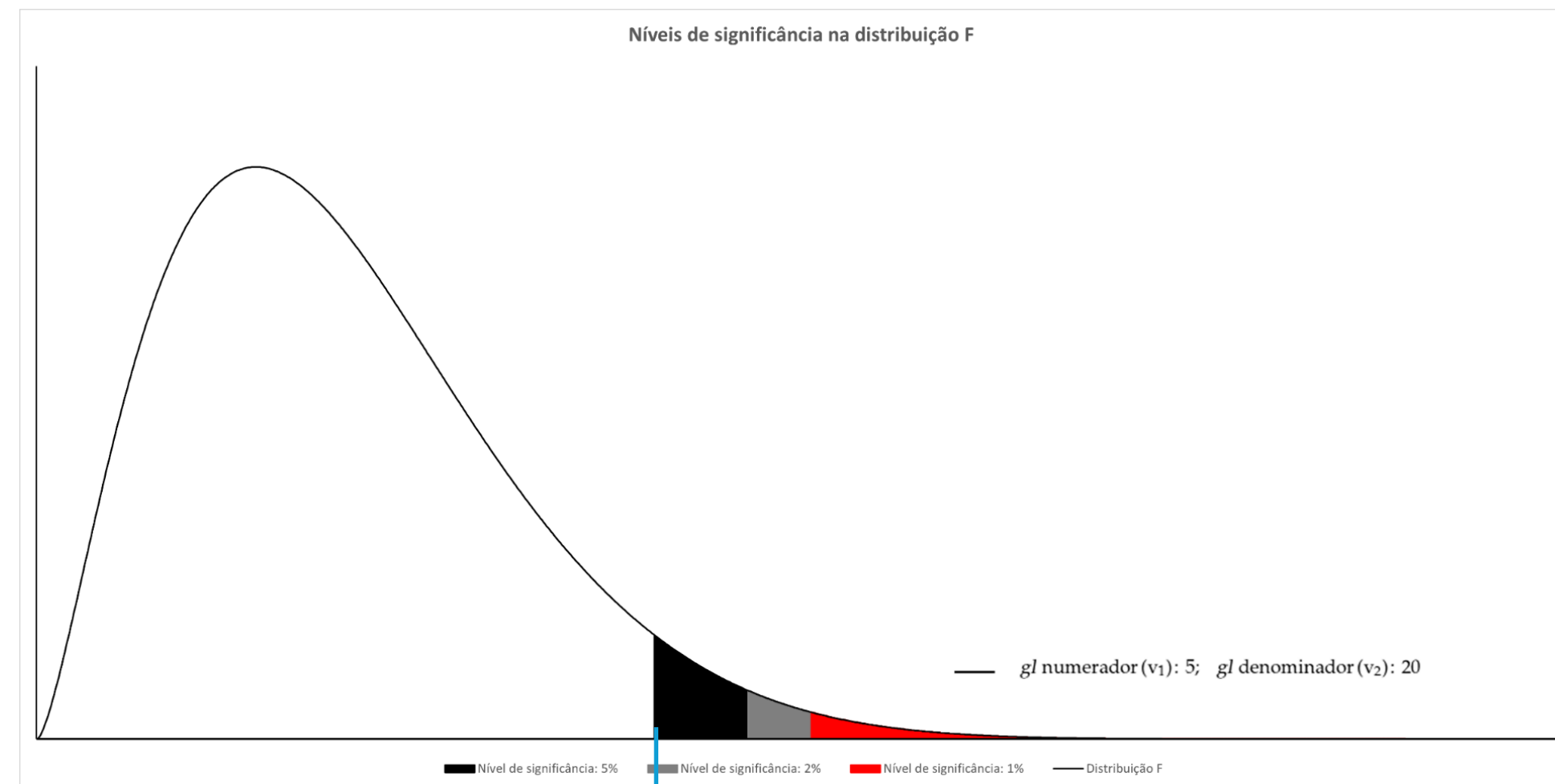


Em geral, utiliza-se a distribuição F para se tomar decisões sobre as populações a partir das estimativas das variâncias (obtidas das amostras) quando se testa hipóteses (inferências sobre as populações). As hipóteses são as mais diversas, porém, em geral, esta distribuição é utilizada para se decidir se os dados podem ser considerados como advindos, ou não, de uma mesma população básica.

$$F(\phi_1, \phi_2) = \frac{\frac{\chi^2_{\phi_1}}{\phi_1}}{\frac{\chi^2_{\phi_2}}{\phi_2}}$$

# DISTRIBUIÇÃO F

Sobre a tabela:



Limite unilateral – valor que será encontrado na tabela.



# Quando usar a Distribuição F?

Em geral, utiliza-se a distribuição F para se tomar decisões sobre as populações a partir das estimativas das variâncias (obtidas das amostras) quando se testa hipóteses (inferências sobre as populações).

As hipóteses são as mais diversas, porém, em geral, esta distribuição é utilizada para se decidir se os dados podem ser considerados como advindos, ou não, de uma mesma população básica.

Tabela 1 Tabela F considerando  $P(F > F_{0.005}) = 0.5\%$

v2	v1									
	1	2	3	4	5	6	7	8	9	10
1	16210.723	19999.500	21614.741	22499.583	23055.798	23437.111	23714.566	23925.406	24091.004	24224.487
2	198.501	199.000	199.166	199.250	199.300	199.333	199.357	199.375	199.388	199.400
3	55.552	49.799	47.467	46.195	45.392	44.838	44.434	44.126	43.882	43.686
4	31.333	26.284	24.259	23.155	22.456	21.975	21.622	21.352	21.139	20.967
5	22.785	18.314	16.530	15.556	14.940	14.513	14.200	13.961	13.772	13.618
6	18.635	14.544	12.917	12.028	11.464	11.073	10.786	10.566	10.391	10.250
7	16.236	12.404	10.882	10.050	9.522	9.155	8.885	8.678	8.514	8.380
8	14.688	11.042	9.596	8.805	8.302	7.952	7.694	7.496	7.339	7.211
9	13.614	10.107	8.717	7.956	7.471	7.134	6.885	6.693	6.541	6.417
10	12.826	9.427	8.081	7.343	6.872	6.545	6.302	6.116	5.968	5.847
11	12.226	8.912	7.600	6.881	6.422	6.102	5.865	5.682	5.537	5.418
12	11.754	8.510	7.226	6.521	6.071	5.757	5.525	5.345	5.202	5.085
13	11.374	8.186	6.926	6.233	5.791	5.482	5.253	5.076	4.935	4.820
14	11.060	7.922	6.680	5.998	5.562	5.257	5.031	4.857	4.717	4.603
15	10.798	7.701	6.476	5.803	5.372	5.071	4.847	4.674	4.536	4.424
16	10.575	7.514	6.303	5.638	5.212	4.913	4.692	4.521	4.384	4.272
17	10.384	7.354	6.156	5.497	5.075	4.779	4.559	4.389	4.254	4.142
18	10.218	7.215	6.028	5.375	4.956	4.663	4.445	4.276	4.141	4.030
19	10.073	7.093	5.916	5.268	4.853	4.561	4.345	4.177	4.043	3.933
20	9.944	6.986	5.818	5.174	4.762	4.472	4.257	4.090	3.956	3.847
21	9.830	6.891	5.730	5.091	4.681	4.393	4.179	4.013	3.880	3.771
22	9.727	6.806	5.652	5.017	4.609	4.322	4.109	3.944	3.812	3.703
23	9.635	6.730	5.582	4.950	4.544	4.259	4.047	3.882	3.750	3.642
24	9.551	6.661	5.519	4.890	4.486	4.202	3.991	3.826	3.695	3.587
25	9.475	6.598	5.462	4.835	4.433	4.150	3.939	3.776	3.645	3.537
26	9.406	6.541	5.409	4.785	4.384	4.103	3.893	3.730	3.599	3.492
27	9.342	6.489	5.361	4.740	4.340	4.059	3.850	3.687	3.557	3.450
28	9.284	6.440	5.317	4.698	4.300	4.020	3.811	3.649	3.519	3.412
29	9.230	6.396	5.276	4.659	4.262	3.983	3.775	3.613	3.483	3.377
30	9.180	6.355	5.239	4.623	4.228	3.949	3.742	3.580	3.450	3.344
31	9.133	6.317	5.204	4.590	4.196	3.918	3.711	3.549	3.420	3.314
32	9.090	6.281	5.171	4.559	4.166	3.889	3.682	3.521	3.392	3.286
33	9.050	6.248	5.141	4.531	4.138	3.861	3.655	3.495	3.366	3.260
34	9.012	6.217	5.113	4.504	4.112	3.836	3.630	3.470	3.341	3.235
35	8.976	6.188	5.086	4.479	4.088	3.812	3.607	3.447	3.318	3.212
36	8.943	6.161	5.062	4.455	4.065	3.790	3.585	3.425	3.296	3.191
37	8.912	6.135	5.038	4.433	4.043	3.769	3.564	3.404	3.276	3.171
38	8.882	6.111	5.016	4.412	4.023	3.749	3.545	3.385	3.257	3.152
39	8.854	6.088	4.995	4.392	4.004	3.731	3.526	3.367	3.239	3.134
40	8.828	6.066	4.976	4.374	3.986	3.713	3.509	3.350	3.222	3.117

Tabela 2 Tabela F considerando  $P(F > F_{0.01}) = 1\%$

v2	v1									
	1	2	3	4	5	6	7	8	9	10
1	4052.181	4999.500	5403.352	5624.583	5763.650	5858.986	5928.356	5981.070	6022.473	6055.847
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814
9	10.561	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398	3.310
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299	3.211
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217	3.129
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149	3.062
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092	3.005
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979
31	7.530	5.362	4.484	3.993	3.675	3.449	3.281	3.149	3.043	2.955
32	7.499	5.336	4.459	3.969	3.652	3.427	3.258	3.127	3.021	2.934
33	7.471	5.312	4.437	3.948	3.630	3.406	3.238	3.106	3.000	2.913
34	7.444	5.289	4.416	3.927	3.611	3.386	3.218	3.087	2.981	2.894
35	7.419	5.268	4.396	3.908	3.592	3.368	3.200	3.069	2.963	2.876
36	7.396	5.248	4.377	3.890	3.574	3.351	3.183	3.052	2.946	2.859
37	7.373	5.229	4.360	3.873	3.558	3.334	3.167	3.036	2.930	2.843
38	7.353	5.211	4.343	3.858	3.542	3.319	3.152	3.021	2.915	2.828
39	7.333	5.194	4.327	3.843	3.528	3.305	3.137	3.006	2.901	2.814
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801

Tabela 4 Tabela F considerando  $P(F > F_{0.05}) = 5\%$ 

v2	v1									
	1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209	2.153	2.106
37	4.105	3.252	2.859	2.626	2.470	2.356	2.270	2.201	2.145	2.098
38	4.098	3.245	2.852	2.619	2.463	2.349	2.262	2.194	2.138	2.091
39	4.091	3.238	2.845	2.612	2.456	2.342	2.255	2.187	2.131	2.084
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077

# DISTRIBUIÇÃO F

## Exemplo de aplicação:

Notações:

m = Média

s = Desvio padrão

$s^2$  = Variância

gl = Graus de

Liberdade

n = Qtd. Elementos

Dois métodos de determinação da CTC do solo são usados em uma amostra de controle e fornecem os resultados da Tabela 14.1.

Tabela 14.1 – Resultados da determinação da capacidade de troca catiônica (cmol<sub>c</sub>/kg) de dois métodos, UESC, BA – março 2009

	r <sub>1</sub>	r <sub>2</sub>	r <sub>3</sub>	r <sub>4</sub>	r <sub>5</sub>	r <sub>6</sub>	r <sub>7</sub>	r <sub>8</sub>	r <sub>9</sub>	r <sub>10</sub>	n	gl	m	<b>s<sup>2</sup></b>	s
<b>A</b>	10,2	8,7	9,5	12,0	9,0	11,2	12,5	10,9	8,9	10,6	10	9	10,35	<b>1,76</b>	1,33
<b>B</b>	9,9	9,2	10,4	10,5	11,0	11,3	9,6	9,4	10,0	10,4	10	9	10,17	<b>0,46</b>	0,68

A questão a ser investigada é se é possível, ou não, considerar as precisões dos dois métodos (população de resultados gerados por cada método) estatisticamente iguais.



# DISTRIBUIÇÃO F DE SNEDECOR

## Testes de hipóteses

$$H_0 : \sigma_A^2 = \sigma_B^2$$

→ precisão igual - população única

$$H_1 : \sigma_A^2 > \sigma_B^2$$

→ precisões distintas - populações distintas

O teste faz uso da razão entre duas estimativas da variância, e como o teste é unilateral à direita  $\sigma_A^2 > \sigma_B^2$  o maior valor ocupa o numerador:

$$F_{cal} = \frac{s_A^2}{s_B^2} \text{ sendo } s_A^2 \geq s_B^2$$



# DISTRIBUIÇÃO F DE SNEDECOR

## Testes de hipóteses

Tabela 14.1 – Resultados da determinação da capacidade de troca catiônica (cmol<sub>c</sub>/kg) de dois métodos, UESC, BA – março 2009

	r <sub>1</sub>	r <sub>2</sub>	r <sub>3</sub>	r <sub>4</sub>	r <sub>5</sub>	r <sub>6</sub>	r <sub>7</sub>	r <sub>8</sub>	r <sub>9</sub>	r <sub>10</sub>	n	gl	m	s <sup>2</sup>	s
<b>A</b>	10,2	8,7	9,5	12,0	9,0	11,2	12,5	10,9	8,9	10,6	10	9	10,35	<b>1,76</b>	1,33
<b>B</b>	9,9	9,2	10,4	10,5	11,0	11,3	9,6	9,4	10,0	10,4	10	9	10,17	<b>0,46</b>	0,68

Erro = 5%

$$1 - \int_0^{3,18} f(F) dF = 1 - 0,95 = 0,05 = 5\%$$

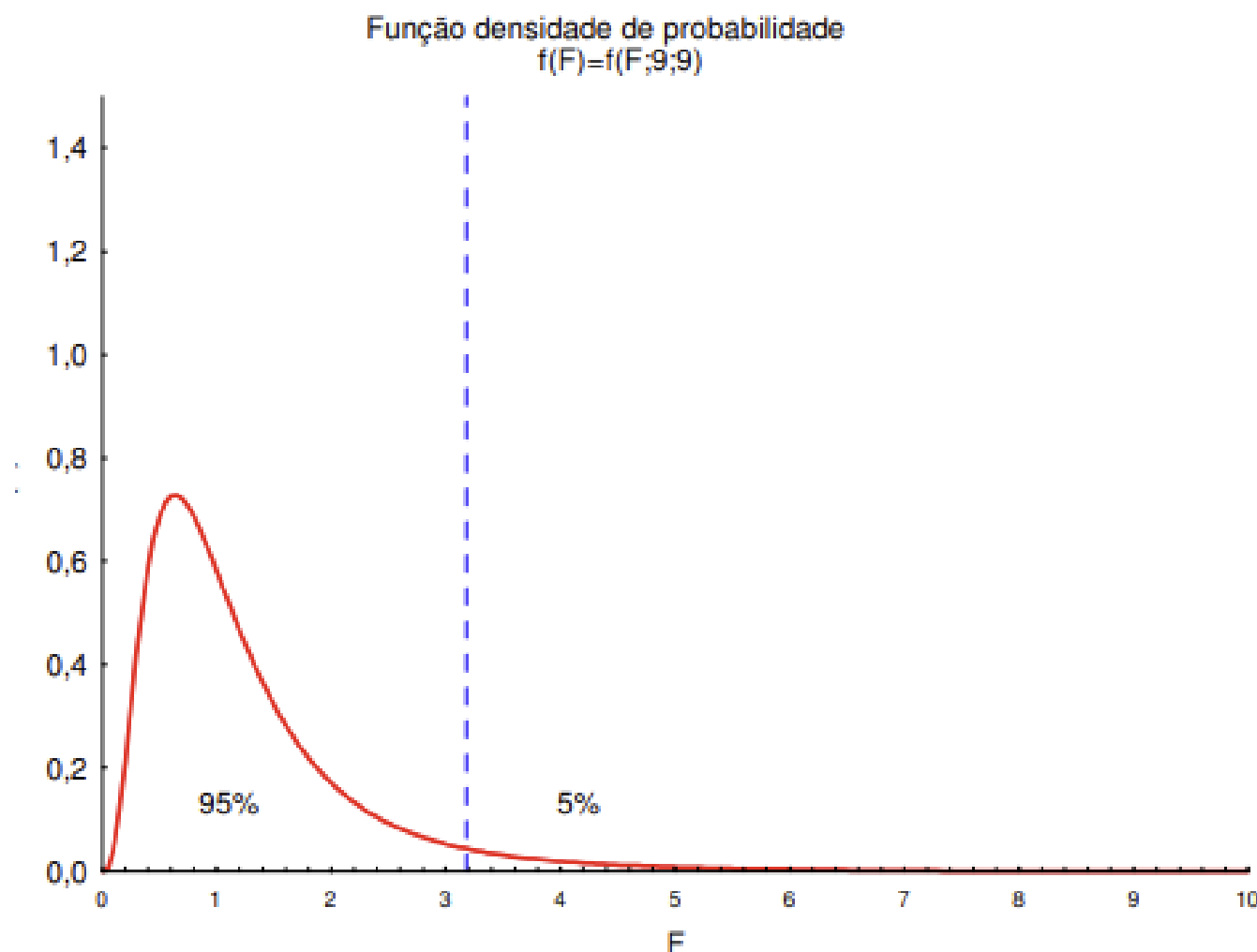
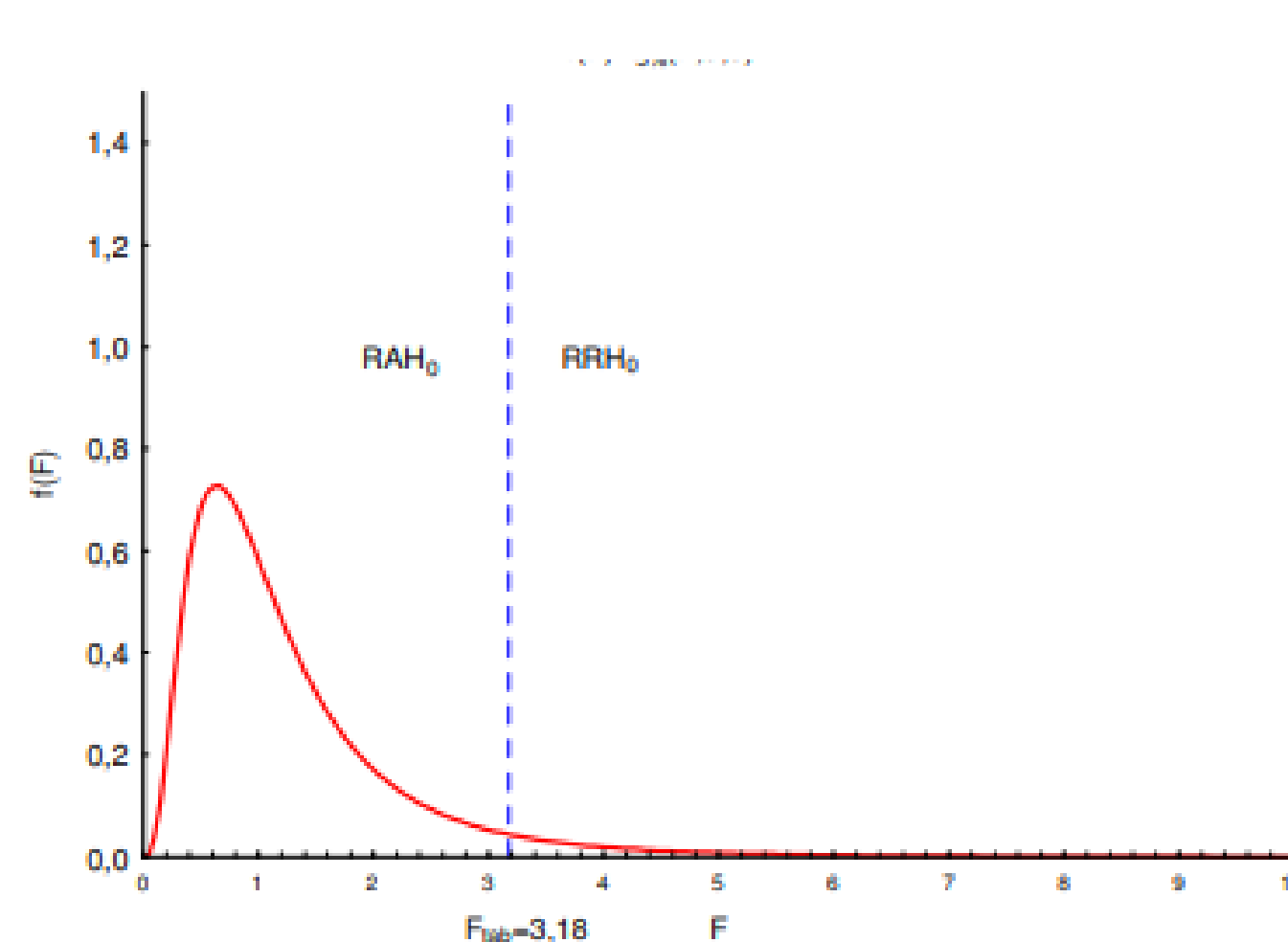
$$\int_0^{3,18} f(F) dF = 0,95 = 95\%$$

$$F_{\text{tab}} \Rightarrow F_{5\%}(9, 9) = 3,179$$

$$F_{\text{cal}} \Rightarrow \frac{>S^2}{<S^2} = \frac{1,76}{0,46} = 3,82$$

Denominando a linha pontilhada de  $F_{\text{tab}}$ :

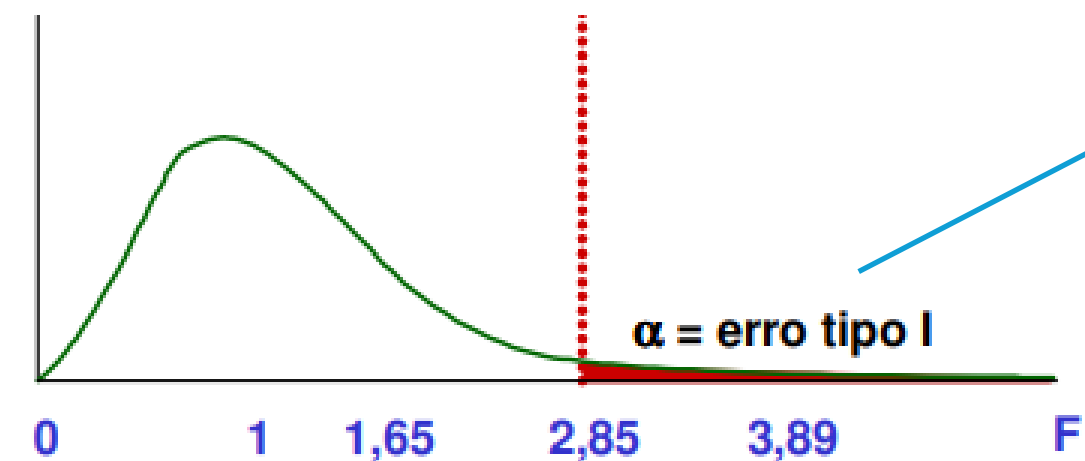
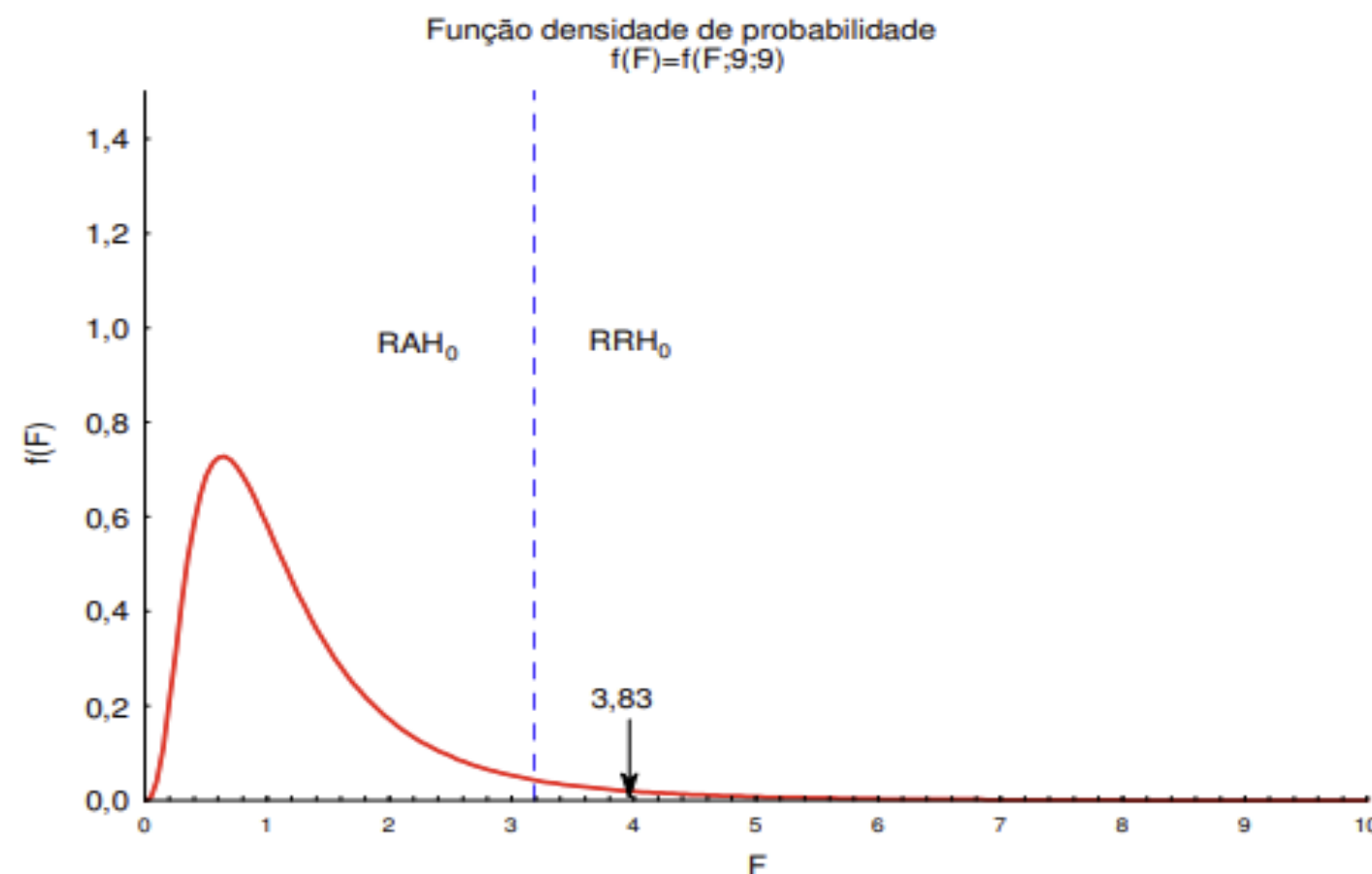
- $F_{\text{cal}} < F_{\text{tab}}$ : aceita-se a igualdade
- $F_{\text{cal}} \geq F_{\text{tab}}$ : rejeita-se a igualdade



# DISTRIBUIÇÃO F DE SNEDECOR

## Conclusão:

Portanto, como o valor de prova ( $F_{\text{cal}} = 3,83$ ), e admitindo uma probabilidade de 5% de erro, deve-se decidir que os resultados produzidos pelos dois métodos não podem ser considerados como provenientes de uma mesma população. A precisão dos métodos não pode ser considerada similar, significando que um método é mais preciso que o outro. Implica dizer que o método (A:  $s^2 = 1,76$ ) é menos preciso que o método (B:  $s^2 = 0,46$ ), e que, para tomar esta decisão, admitiu-se um erro de 5%.



Estamos cometendo o erro tipo I

# DISTRIBUIÇÕES

## Diferenças:

Qui-quadrado: Usada para dados categóricos e testa a independência entre variáveis. Não assume uma distribuição normal e é baseada em frequências.

t de Student: Usada para dados contínuos quando o tamanho da amostra é pequeno e o desvio padrão da população é desconhecido. Baseia-se na estimativa do desvio padrão da amostra e é mais “larga” que a distribuição normal, proporcionando mais flexibilidade para estimativas com menos dados.

F de Snedecor: Usada para comparar duas ou mais variâncias de amostras e é uma razão de variâncias. É aplicada em testes ANOVA para comparar médias de três ou mais grupos.

# Referências e material de apoio

*Slide de referência para construção do atual.*

*Oliveira, B. et.al (2023.2). Variáveis Aleatórias Contínuas.*

*Faria, J. C. (2009). Notas de aula expandida.*

*[https://lec.pro.br/download/faria/apostilas/CET018\\_10ed\\_1pf.pdf](https://lec.pro.br/download/faria/apostilas/CET018_10ed_1pf.pdf)*

*Bonat, W. H. Funções de probabilidade, densidade e distribuição.*

*[http://www.leg.ufpr.br/~paulojus/estbas/slides/302\\_variaveis\\_aleatorias.pdf](http://www.leg.ufpr.br/~paulojus/estbas/slides/302_variaveis_aleatorias.pdf)*

*Maioli, D. (2021). Probabilidade Aula 21 - Distribuição Normal.*

*[https://www.youtube.com/watch?v=AB4wWER\\_wAI](https://www.youtube.com/watch?v=AB4wWER_wAI)*

*UNIVESP. (2018). Estatística e Probabilidade - Aula 12 - Teste Qui-quadrado*

*<https://www.youtube.com/watch?v=4QfHVbpAoSg>*

# Auxílio visual para os exemplos em R

Dada a função:

$$f(y) = \begin{cases} 2y, & \text{para } 0 < y < 1 \\ 0 & \text{para quaisquer outros valores} \end{cases}$$

Verificar se  $f(y)$  é uma FDP

Calcule a probabilidade de  $0 < y < 0,5$

Calcule a probabilidade de  $0,5 < y < 1$

# Auxílio visual para os exemplos em R

Dada a função:

$$f(x) = \begin{cases} 2e^{-2x}, & \text{se } x \geq 0 \\ 0 & , \text{se } x < 0 \end{cases}$$

Mostre que essa função é uma FDP

Calcule a probabilidade de  $X > 1$

Calcule a probabilidade de que  $0.2 < X < 0.8$



# Distribuição Normal Padrão

No R, a distribuição normal pode ser utilizada por meio das funções abaixo, em todas elas pode ser definir a média (mean) e o desvio padrão (sd):

`dnorm(x, mean, sd)` - densidade de probabilidade no ponto x

`pnorm(x, mean, sd)` - função de probabilidade acumulada no ponto x

`qnorm(p, mean, sd)` - quantil correspondente a uma dada probabilidade p

`rnorm(n, mean, sd)` - amostra da distribuição normal de tamanho n

# Distribuição Qui-Quadrado ( $\chi^2$ )

No R, a distribuição  $\chi^2$  pode ser utilizada por meio das funções abaixo, em todas elas pode ser definir o grau de liberdade (df):

`dchisq(x, df)` - densidade de probabilidade no ponto  $x$

`pchisq(x, df)` - função de probabilidade acumulada no ponto  $x$

`qchisq(p, df)` - quantil correspondente a uma dada probabilidade  $p$

`rchisq(n, df)` - amostra da distribuição  $\chi^2$  de tamanho  $n$

# Distribuição t de Student

No R, a distribuição t pode ser utilizada por meio das funções abaixo, em todas elas pode ser definir o grau de liberdade (df):

`dt(x, df)` - densidade de probabilidade no ponto x

`pt(x, df)` - função de probabilidade acumulada no ponto x

`qt(p, df)` - quantil correspondente a uma dada probabilidade p

`rt(n, df)` - amostra da distribuição t de tamanho n

# Distribuição F de Snedecor

No R, a distribuição F pode ser utilizada por meio das funções abaixo, em todas elas pode se definir os graus de liberdade do numerador (df1) e denominador (df2):

`df(x, df1, df2)` - densidade de probabilidade no ponto  $x$

`pf(x, df1, df2)` - função de probabilidade acumulada no ponto  $x$

`qf(p, df1, df2)` - quantil correspondente a uma dada probabilidade  $p$

`rf(n, df1, df2)` - amostra da distribuição F de tamanho  $n$

# Obrigado!

Contato : [wssilva.cic@uesc.br](mailto:wssilva.cic@uesc.br)

