PROBABILIDADE E ESTATÍSTICA

Professor: Mauri A. Oliveira

mauriao2@gmail.com

PRINCIPAIS DISTRIBUIÇÕES CONTÍNUAS

Tipos de variáveis:

- VARIÁVEL
  - QUALITATIVA
    - NOMINAL
    - ORDINAL
  - QUANTITATIVA
    - DISCRETA
    - CONTÍNUA
      - UNIFORME
      - NORMAL
      - EXPONENCIAL
      - OUTRAS

O QUE É ESTUDAR UMA VARIÁVEL ALEATÓRIA?

1. f(x) ou P(x)
2. F(x)
3. E(x)
4. Var(x)
5. Utilidade

Distribuição Uniforme [X~U(α,β)]:

➢ **Def: Uma variável aleatória X tem distribuição uniforme se sua função distribuição de probabilidade é dada por:**

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & , se \ \ \alpha \leq x \leq \beta \\ 0 & , caso \text{ contrário} \end{cases}$$

$$F(x) = P(X \leq x) = \begin{cases} 0 & , x < \alpha \\ \dfrac{\text{x} - \alpha}{\beta - \alpha} & , \alpha \leq x \leq \beta \\ 1 & , x > \beta \end{cases}$$

E(X) = $(\alpha+\beta) / 2$

Var(X) = $(\beta - \alpha)^2 / 12$

Distribuição Normal [X~N(μ,σ^2)]:

- **Uma variável aleatória X tem distribuição Normal se sua f.d.p. é:**

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

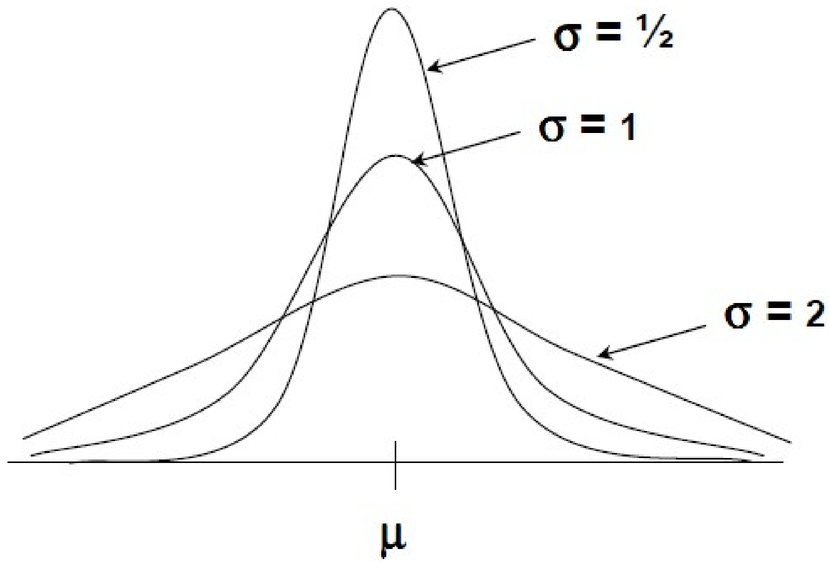
$\rightarrow$ **x $\in \Re$**

$\rightarrow$ **$\mu \in \Re$ ($-\infty < \mu < \infty$)**

$\rightarrow$ **$\sigma \in \Re$+ ($\sigma > 0$)**

$$\rightarrow \begin{cases} E[X] = \mu \\ Var[X] = \sigma^2 \end{cases}$$

- **Propriedades:**



- **Forma de sino centrado em μ**
- **Simétrica**
- **Achatamento depende de σ**
- **Há um único máximo global em x= μ**
- **f(x) é crescente para x < μ**
- **f(x) é decrescente para x > μ**

Distribuição Normal [X~N(μ,σ²)]:

- Distribuição Normal Padronizada: Z~N(0,1): uma variável aleatória Z tem distribuição Normal padronizada (ou reduzida) se sua f.d.p. pode ser escrita como:

$$Z = \frac{X-\mu}{\sigma} \quad \begin{cases} E[Z] = 0 \\ Var[Z] = 1 \end{cases} \qquad f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad , -\infty < z < \infty$$

- Tabulação da Distribuição Normal Padronizada

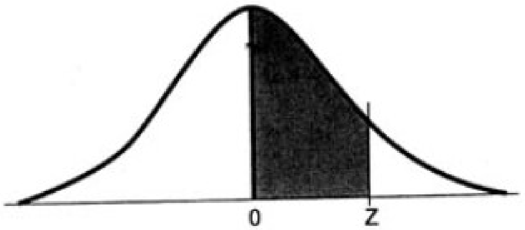
Distribuição Normal [$X \sim N(\mu,\sigma^2)$]:

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

→ **$E[X] = \mu$**

→ **$Var[X] = \sigma^2$**

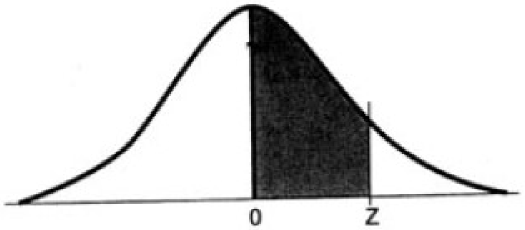
- **Se $X \sim N(\mu,\sigma^2)$ e se $Z=(X-\mu)/\sigma$ então $Z \sim N(0,1)$**



This table presents the area between the mean and the Z score . When Z=1.96, the shaded area is 0.4750.

Areas Under the Standard Normal Curve

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767



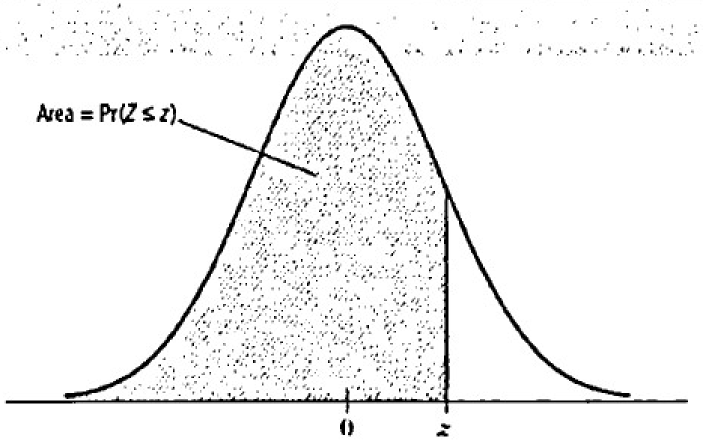
This table presents the area between the mean and the Z score . When Z=1.96, the shaded area is 0.4750.

Areas Under the Standard Normal Curve

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000									

Source: Adapted by permission from *Statistical Methods* by George W. Snedecor and William G. Cochran, sixth edition

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \leq z)$



z	Second Decimal Value of z									
	0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367

NORMAL

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{Exp}\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

In[1]:= $\frac{1}{1*\sqrt{2\pi}} \int_{-\infty}^{\infty} \text{Exp}\left[-\frac{1}{2}\left(\frac{x-0}{1}\right)^2\right] dx$

Out[1]= 1

In[2]:= $\frac{1}{1*\sqrt{2\pi}} \int_{-\infty}^{-2.32} \text{Exp}\left[-\frac{1}{2}\left(\frac{x-0}{1}\right)^2\right] dx$

Out[2]= 0.0101704

➢Exemplo

Uma fábrica de carros sabe que os motores de sua fabricação têm duração normal com média de 150.000 km e desvio padrão de 5.000 km.

Qual a probabilidade de que um carro, escolhido ao acaso, dos fabricados por essa empresa, tenha um motor que dure:

a) menos de 170.000 km?
b) Entre 140.000 km e 165.000 km?

Aproximações da Distribuição Normal:

- **Distribuição Binomial: Se X~Bin(n,p) em que n é relativamente grande ($n \geq 30$) e n.p.(1-p)>5 pode ser aproximada à uma Normal com μ = n.p e σ^2 =n.p.(1-p), ou seja,**

$$\frac{x - n.p}{\sqrt{n.p.(1-p)}} \sim N(0,1)$$

- **Distribuição de Poisson: Uma variável aleatória que segue uma distribuição de Poisson com parâmetro $\lambda \geq 5$ pode ser aproximada à uma distribuição Normal com $\mu = \lambda$ e $\sigma = \lambda^{1/2}$, ou seja,**

$$\frac{x - \lambda}{\sqrt{\lambda}} \sim N(0,1)$$

Distribuição Exponencial [X~Exp(λ)], λ >0:

$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

$\rightarrow$ **E[X] = $\frac{1}{\lambda}$**

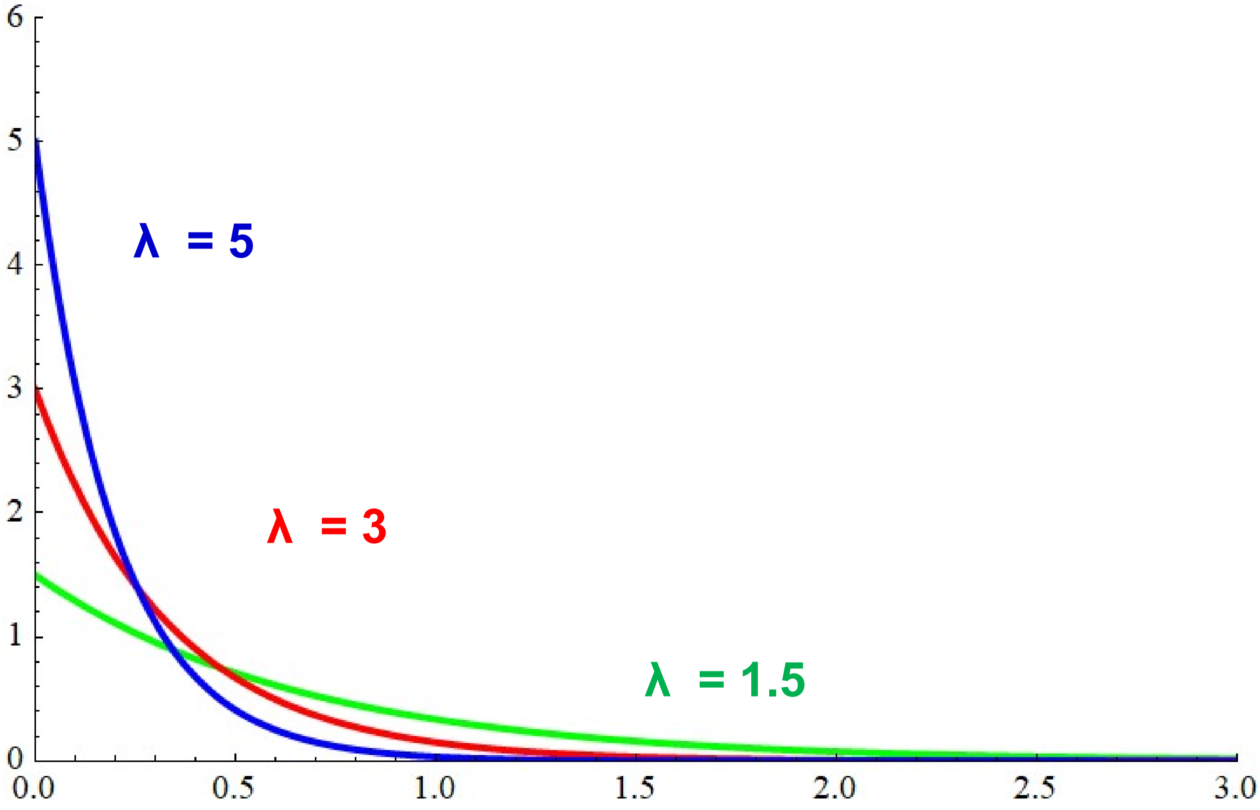
$\rightarrow$ **Var[X] = $\frac{1}{\lambda^2}$**

Propriedade de Falta de Memória

P(X ≥ t + s | X ≥ t) = P(X ≥ s) $\qquad$ **∀s, t ∈ ℝ com s, t ≥ 0**

É a única distribuição de probabilidade contínua em que ocorre falta de memória.

Interpretação: Suponha que queremos determinar a probabilidade de um cliente chegar na próxima meia hora. Essa propriedade nos diz que é o mesmo saber quando o último cliente chegou ou calcular diretamente qual é a probabilidade de chegar nos próximos 30 minutos sem levar em consideração o passado.



A distribuição exponencial é uma distribuição também muito utilizada na prática para modelar tempo de falha de objetos.

Distribuição Gamma [X~Gamma(α,β)]:

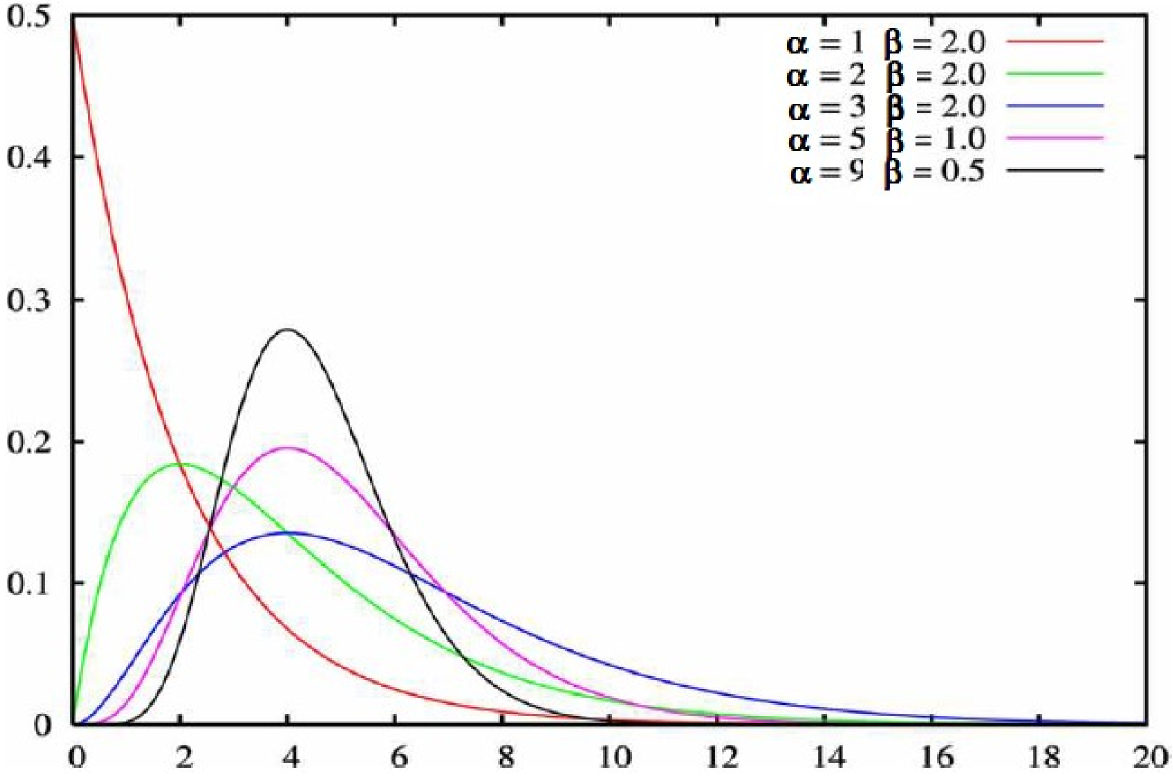
- **Def - Função Gamma:** $\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}e^{-x}dx$

  - $\alpha>1$, $\Gamma(\alpha)=(\alpha-1).\ \Gamma(\alpha-1)$
  - Para n inteiro e positivo, $\Gamma(n)=(n-1)!$
  - $\Gamma(½) = \pi^{½}$

- **Def: Uma variável aleatória X tem distribuição gamma se sua função distribuição de probabilidade é dada por:**

$$f(x)=\begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}\mathrm{x}^{\alpha\text{-}1}\ \mathrm{e}^{\text{-x}/\beta} & ,x\geq 0 \\ 0 & ,caso\ \text{contrário} \end{cases} \qquad ,\alpha>0,\beta>0$$

Distribuição Gamma [X~Gamma(α,β)]:



$$f(x)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}\mathrm{x}^{\alpha\text{-}1}\,\mathrm{e}^{-\mathrm{x}/\beta} & ,x\geq 0\\ 0 & ,caso\ \text{contrário}\end{cases} ,\alpha>0,\beta>0$$

$\longrightarrow$ **E[X] =** $\dfrac{\alpha}{\beta}$

$\longrightarrow$ **Var[X] =** $\dfrac{\alpha}{\beta^2}$

Família da Distribuição Gamma:

- **Distribuição Exponencial: α= 1 e β=1/λ ,** $f(x) = \begin{cases} \lambda \, \mathrm{e}^{-\lambda \mathrm{x}} & , x \geq 0 \\ 0, caso \text{ contrário} \end{cases}$

- **Distribuição Chi-quadrado: α= ν/2 e β=2**

$$f(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \mathrm{x}^{(\nu/2)\text{-}1} \, \mathrm{e}^{-\mathrm{x}/2} & , x \geq 0 \\ 0 & , caso \text{ contrário} \end{cases}$$

em que ν é o número de graus de liberdade de X

Algumas outras Distribuições

- Distribuição Log-normal
- Distribuição Weibull
- Distribuição Beta
- Distribuição Erlang
- Distribuição Wald
- Distribuição Cauchy
- Distribuição Pareto
- Distribuição Rayleigh

DISTRIBUIÇÕES AMOSTRAIS

Estatísticas e suas distribuições

- Quando estudamos estatística descritiva, observamos uma amostra única (x_1, x_2, …,x_n) de tamanho n da população para obter estimativas pontuais dos momentos amostrais ($\bar{x}$ e s^2).

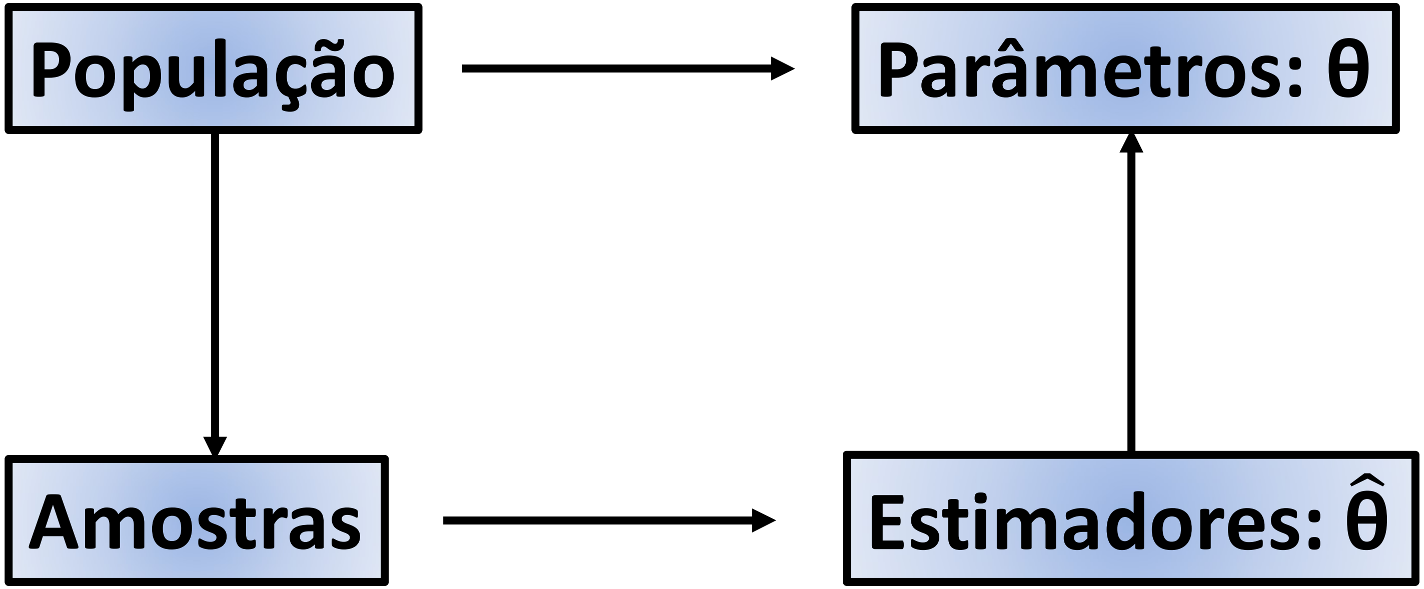
- Se selecionarmos 2 amostras (cada uma de tamanho n) é de se esperar que as observações da segunda amostra serão diferentes das observações da primeira amostra. Por exemplo:

  Amostra 1 (n=3 carros) $\rightarrow x_1=30{,}7\ x_2=29{,}4\ x_3=31{,}1$

  Amostra 2 (n=3 carros) $\rightarrow x_1=28{,}8\ x_2=30{,}0\ x_3=31{,}3$

- Essas estimativas pontuais são utilizadas para fazer inferências em relação ao valor esperado e ao desvio-padrão / variância da população, os quais são desconhecidos.

Estatísticas e suas distribuições



- Def: ESTATÍSTICA é qualquer quantidade cujo valor pode ser calculado a partir de uma amostra. Como há incerteza em relação a esses valores toda estatística é uma variável aleatória (iid).

- Toda estatística, por ser uma variável aleatória, tem uma distribuição de probabilidade. A distribuição de probabilidade de uma estatística é chamada distribuição amostral.

- Def: Distribuição amostral é a distribuição de probabilidade de uma estatística e descreve como o valor da estatística varia em função das amostras selecionadas.

- A distribuição amostral depende da distribuição de probabilidade da população, do tamanho da amostra e do método de amostragem (neste curso vamos estudar apenas a aleatória)

Distribuição da média amostral ($\bar{x}$)

- Sejam $x_1, x_2, \ldots, x_n$ n elementos de uma amostra aleatória de uma população normal com parâmetros $E[x] = \mu$ e $Var[x] = \sigma^2$. Então:

  i. $E[\bar{x}] = \mu$

  ii. $Var[\bar{x}] = \sigma^2/n$

  iii. $To = x_1 + \ldots + x_n$, $E[To] = n\mu$, $Var[To] = n\sigma^2$

- De acordo com i a distribuição amostral de $\bar{x}$ é centrada no valor esperado da população da qual a amostra foi selecionada, com ii que a distribuição se torna mais concentrada em torno de μ com o crescimento do tamanho da amostra (n).

Distribuição da média amostral ($\bar{x}$)

$$E[\bar{x}] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i]$$

$$= \frac{1}{n}\sum_{i=1}^{n} E[x] = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n}.n\mu$$

$$E[\bar{x}] = \mu$$

Quando E[$\hat{\theta}$] = θ, o estimador é não viciado, não viesado ou não tendencioso.
Logo, $\bar{x}$ é um estimador não tendencioso de μ.

Distribuição da média amostral ($\overline{x}$)

$$Var[\overline{x}] = Var\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n^2} Var[x_1 + x_2 + \ldots + x_n]$$

$$= \frac{1}{n^2}\left(Var[x_1] + Var[x_2] + \ldots + Var[x_n]\right) = \frac{1}{n^2} .n\sigma^2$$

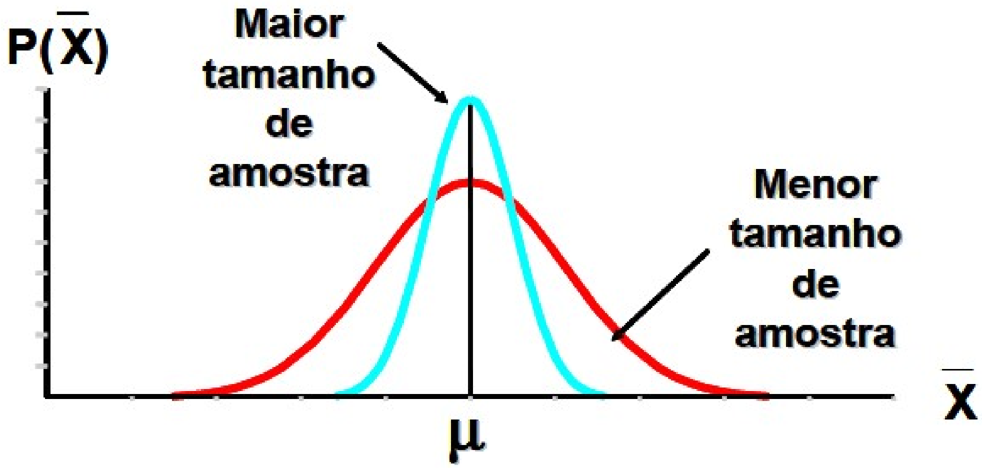
$$Var[\overline{x}] = \frac{\sigma^2}{n}$$

Distribuição da média amostral ($\bar{x}$)

➢ Proposição: Seja $x_1, x_2, \ldots, x_n$ uma amostra aleatória de uma população normal com $E[x] = \mu$ e $Var[x] = \sigma^2$, então, para qualquer n:

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

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



$$To \sim N\left(n\mu, n\sigma^2\right) \Rightarrow \frac{To - n\mu}{\sigma\sqrt{n}} \sim Z$$

Distribuição de uma combinação linear de v.a.'s:

- Sendo $Y = a_1X_1 + \ldots + a_nX_n$

$$\Rightarrow E[Y] = a_1E[X_1] + \ldots + a_nE[X_n] = a_1\mu_1 + \ldots + a_n\mu_n$$

Se as variáveis aleatórias são independentes,

$$Var[a_1X_1 + \ldots + a_nX_n] = a_1^2Var[X_1] + \ldots + a_n^2Var[X_n]$$

- Distribuição da diferença entre médias de populações normais:

Sejam A e B duas populações normais com parâmetros μ_A, σ_A, μ_B e σ_B.

Assim,

$$\overline{X}_A \sim N\left(\mu_A, \sigma_A^2/n_A\right) \quad \text{e} \quad \overline{X}_B \sim N\left(\mu_B, \sigma_B^2/n_B\right)$$

$$\overline{X}_A - \overline{X}_B \sim N\left(\mu_A - \mu_B, \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right) \Rightarrow \frac{\left(\overline{X}_A - \overline{X}_B\right) - \left(\mu_A - \mu_B\right)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim Z$$

Distribuição amostral de estatísticas

- Existem dois métodos para obter a distribuição amostral de uma estatística:

  - Método 1: Utilizando regras da probabilidade

Ex:

Serviço	A	B	C
Preço: x	40	45	50
p(x)	0,2	0,3	0,5

$$\mu = E[x] = 40 * 0,2 + 45 * 0,3 + 50 * 0,5 = 46,5$$

$$\sigma^2 = Var[x] = 0,2 * (40-46,5)^2 + 0,3 * (45-46,5)^2 + 0,5 * (50-46,5)^2 = 15,25$$

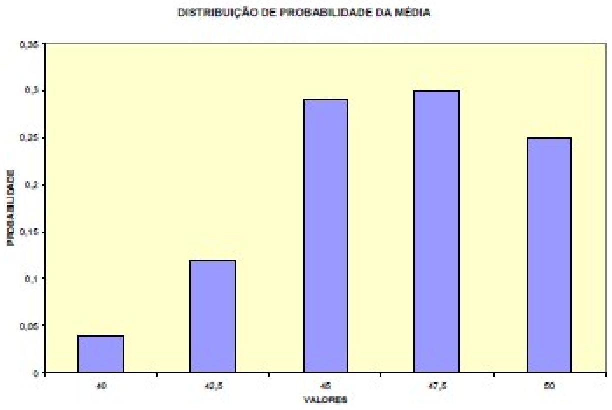
  - Método 2: Realização de um experimento de simulação

Usa-se um computador para obter k amostras aleatórias de tamanho n da distribuição designada. Para cada amostra calcula-se as estatísticas de interesse e cria-se um histograma para avaliar a distribuição amostral aproximada da estatística.

Distribuição amostral de estatísticas

- Método 1: n=2

x1	x2	p(x1)	p(x2)	p(x1,x2)	$\overline{x}$	s^2
40	40	0,2	0,2	0,04	40	0
40	45	0,2	0,3	0,06	42,5	12,5
40	50	0,2	0,5	0,1	45	50
45	40	0,3	0,2	0,06	42,5	12,5
45	45	0,3	0,3	0,09	45	0
45	50	0,3	0,5	0,15	47,5	12,5
50	40	0,5	0,2	0,1	45	50
50	45	0,5	0,3	0,15	47,5	12,5
50	50	0,5	0,5	0,25	50	0



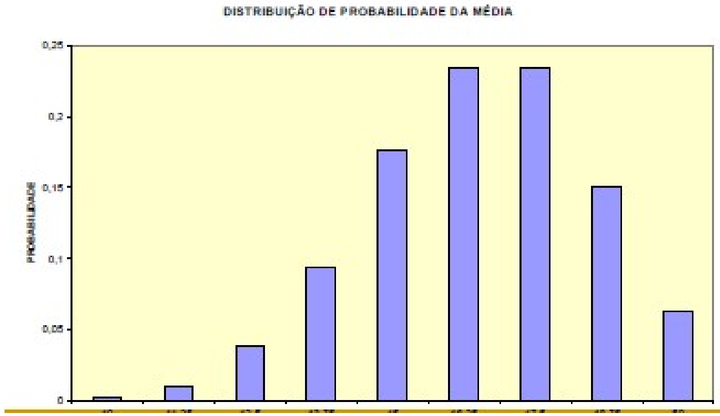
$$E[\bar{x}] = 40 * 0,04 + 42,5 * 0,12 + \ldots + 50 * 0,25 = 46,5$$

$$Var[\bar{x}] = 0,04 * (40-46,5)^2 + \ldots + 0,25 * (50-46,5)^2 = 7,625$$

$$E[s^2] = 0 * 0,38 + 12,5 * 0,42 + 50 * 0,20 = 15,25$$

➢Método 1: n=4

x1	x2	x3	x4	p(x1)	p(x2)	p(x3)	p(x4)	p(x1,x2,x3,x4)	$\bar{x}$	s^2
40	40	40	40	0,2	0,2	0,2	0,2	0,0016	40	0
40	40	40	45	0,2	0,2	0,2	0,3	0,0024	41,25	6,25
40	40	40	50	0,2	0,2	0,2	0,5	0,004	42,5	25
40	40	45	40	0,2	0,2	0,3	0,2	0,0024	41,25	6,25
40	40	45	45	0,2	0,2	0,3	0,3	0,0036	42,5	8,333
40	40	45	50	0,2	0,2	0,3	0,5	0,006	43,75	22,92
40	40	50	40	0,2	0,2	0,5	0,2	0,004	42,5	25
40	40	50	45	0,2	0,2	0,5	0,3	0,006	43,75	22,92
40	40	50	50	0,2	0,2	0,5	0,5	0,01	45	33,33
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
50	50	50	40	0,5	0,5	0,5	0,2	0,025	47,5	25
50	50	50	45	0,5	0,5	0,5	0,3	0,0375	48,75	6,25
50	50	50	50	0,5	0,5	0,5	0,5	0,0625	50	0



$$E[\,\bar{x}\,]=40*0{,}016+\ldots+50*0{,}0625=46{,}5$$

$$Var[\,\bar{x}\,] = 3{,}8125\ (s^2\ /\ 4)$$

➢ Distribuição de $\bar{x}$ para uma população não normal:

Para exemplificar, vamos considerar uma população finita X: 1, 2, 3, 4, 5. Ou seja, N = 5.

Sabemos que:

$$E[x] = \mu_x = \sum_{i=1}^{N} x_i \, . \, p(x_i)$$

e

$$Var[x] = \sigma_x^2 = \sum_{i=1}^{N} (x_i - \mu_x)^2 \, . \, p(x_i)$$

Dessa forma,

$$E[x] = \mu_x = \frac{1}{5}(1+2+3+4+5) = \frac{15}{5} = 3$$

Para a variância,

x	$P(x)$	$x - \mu_x$	$(x-\mu_x)^2$	$(x-\mu_x)^2 . P(x)$
1	1/5	-2	4	4/5
2	1/5	-1	1	1/5
3	1/5	0	0	0
4	1/5	1	1	1/5
5	1/5	2	4	4/5
Σ	1			2

$$Var[x] = \sigma_x^2 = 2$$

Vamos retirar dessa população de tamanho N=5 todas as amostras com reposição de tamanho n=2. Podemos retirar $N^n = 5^2 = 25$ amostras. Calculamos as médias de cada amostra.

Amostras		$\bar{x}_i$
1	(1,1)	1,0
2	(1,2)	1,5
3	(1,3)	2,0
4	(1,4)	2,5
5	(1,5)	3,0
6	(2,1)	1,5
7	(2,2)	2,0
8	(2,3)	2,5
9	(2,4)	3,0
10	(2,5)	3,5

Amostras		$\bar{x}_i$
11	(3,1)	2,0
12	(3,2)	2,5
13	(3,3)	2,0
14	(3,4)	2,5
15	(3,5)	4,0
16	(4,1)	2,5
17	(4,2)	3,0
18	(4,3)	3,5
19	(4,4)	4,0
20	(4,5)	4,5

Amostras		$\bar{x}_i$
21	(5,1)	3,0
22	(5,2)	3,5
23	(5,3)	4,0
24	(5,4)	4,5
25	(5,5)	5,0

Como $\overline{x}$ varia de amostra para amostra, $\overline{x}$ é uma variável aleatória, nesse caso, discreta.
Determinamos a distribuição da variável $\overline{x}$ e calculamos E[$\overline{x}$] e Var[$\overline{x}$].

$\overline{x}$	$P(\overline{x})$	$\overline{x}.P(\overline{x})$	$\overline{x}^2.P(\overline{x})$
1,0	1/25	1/25	1/25
1,5	2/25	3/25	4,5/25
2,0	3/25	6/25	12/25
2,5	4/25	10/25	25/25
3,0	5/25	15/25	45/25
3,5	4/25	14/25	49/25
4,0	3/25	12/25	48/25
4,5	2/25	9/25	40,5/25
5,0	1/25	5/25	25/25
Σ	1	3	10

$$E\left[\overline{x}\right] = \mu_{\overline{x}} = \sum_{i=1}^{N} \overline{x}_i . p\left(\overline{x}_i\right)$$

$$E\left[\overline{x}\right] = \mu_{\overline{x}} = 3$$

Como

$$E\left[\overline{x}^2\right] = \sum_{i=1}^{N} \overline{x}_i^2 . p\left(\overline{x}_i\right),$$

temos

$$E\left[\overline{x}^2\right] = 10.$$

Sendo

$$Var[\overline{x}] = E\left[\overline{x}^2\right] - \{E[\overline{x}]\}^2,$$

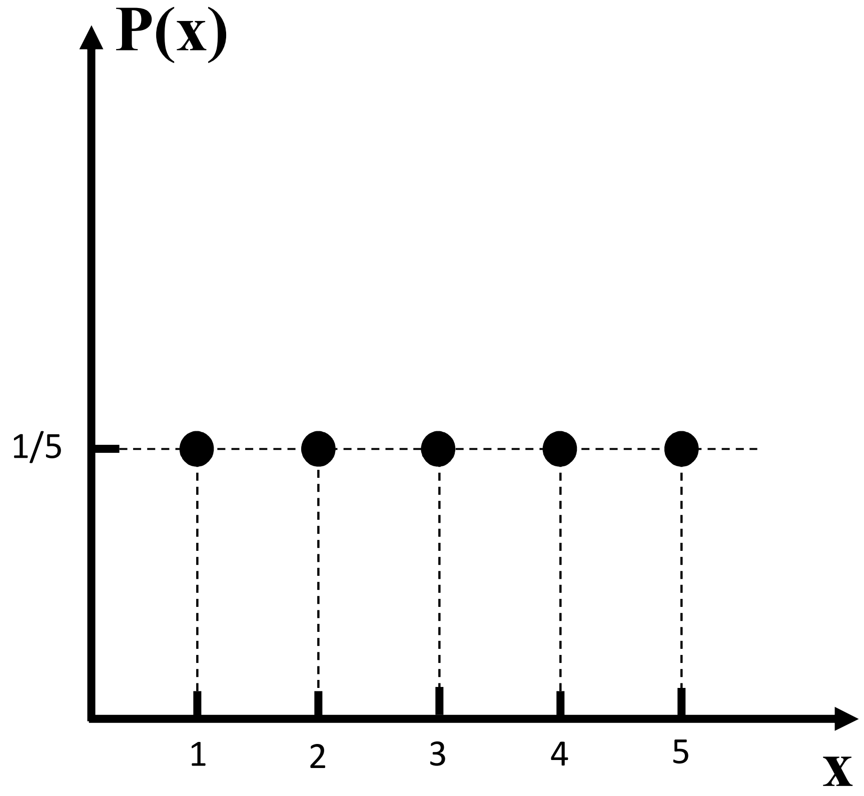
então,

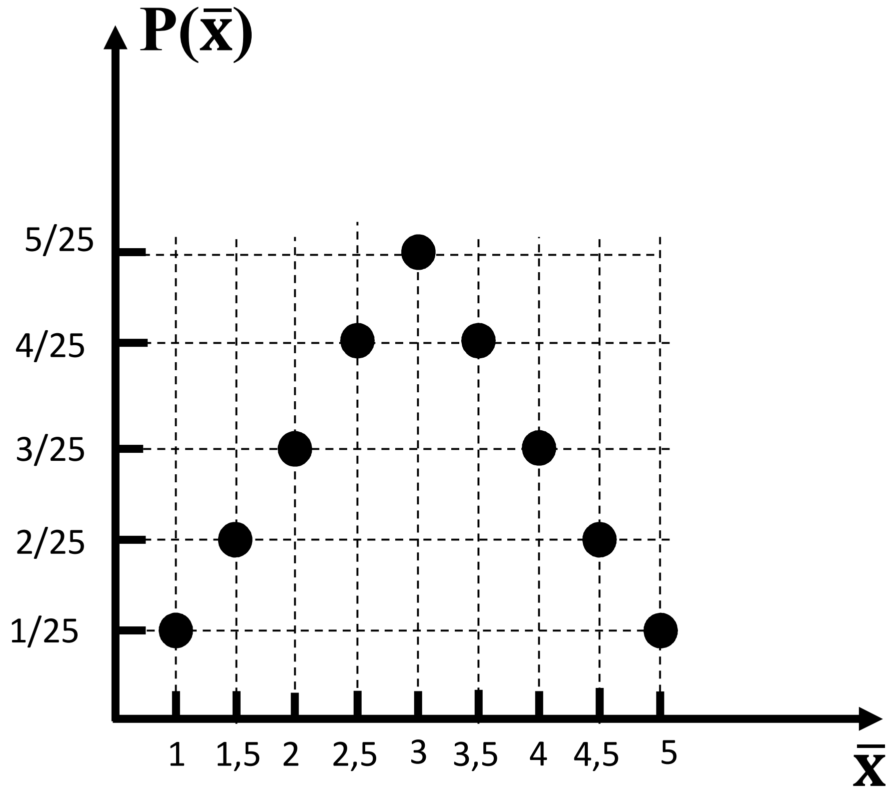
$$Var[\overline{x}] = 10 - 3^2.$$

Logo:

$$\sigma_{\overline{x}}^2 = Var[\overline{x}] = 1.$$

Graficamente





Concluímos

Proposição – 1

A média das médias amostrais, ou E[$\overline{x}$], é igual a média populacional, ou E[$\overline{x}$] = $\mu_{\overline{x}}$.

Proposição – 2

A variância da média amostral é igual a variância populacional dividida pelo tamanho da amostra.

LEI DOS GRANDES NÚMEROS

Lei dos Grandes Números

A média aritmética dos resultados da realização da mesma experiência repetidas vezes tende a se aproximar do valor esperado à medida que mais tentativas se sucederem.

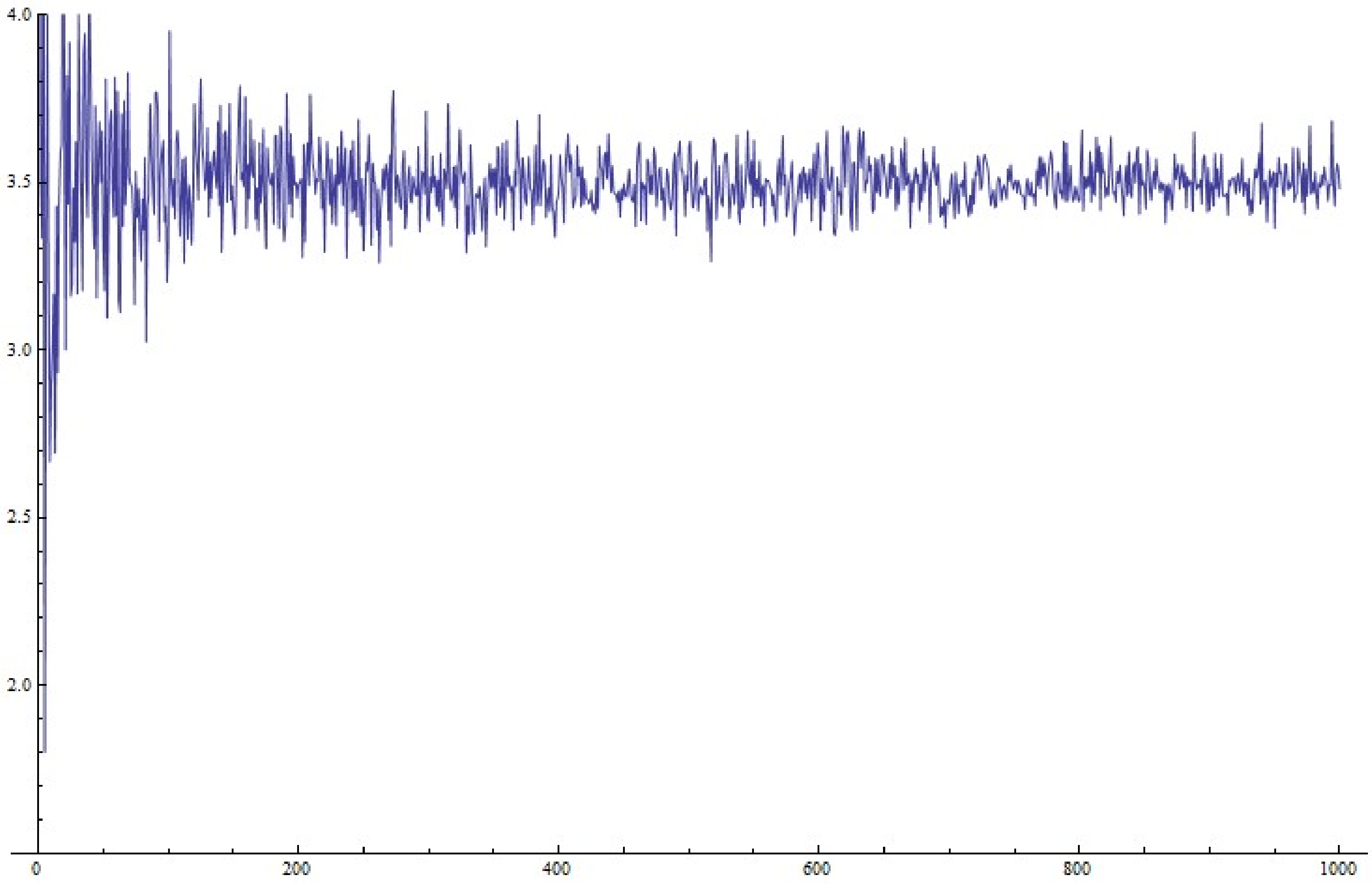
Lei dos Grandes Números

Seja $X_1, X_2, \ldots$ uma sequência de variáveis aleatórias i.i.d. com média finita μ.

As somas parciais $S_n = X_1 + X_2 + \ldots + X_n$ satisfazem

$$\frac{S_n}{n} \to \mu.$$

Lei dos Grandes Números



TEOREMA CENTRAL DO LIMITE

Teorema Central do Limite

The term "central limit theorem" most likely traces back to Georg Pólya. As he recapitulated at the beginning of a paper published in 1920, it was "generally known that the appearance of the Gaussian probability density[1] e^{-x^2}" in a great many situations "can be explained by one and the same limit theorem," which plays "a central role in probability theory" [Pólya 1920, 171]. Laplace had discovered the essentials of this fundamental theorem in 1810, and with the designation "central limit theorem of probability theory," which was even emphasized in the paper's title, Pólya gave it the name that has been in general use ever since.

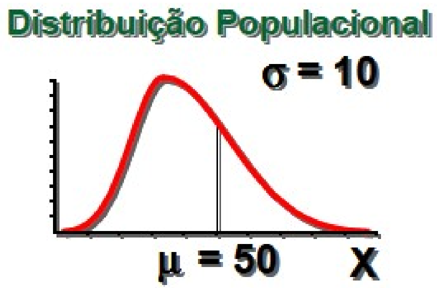
A History of the Central Limit Theorem: From Classical to Modern Probability Theory [Hans Fisher, 2011]

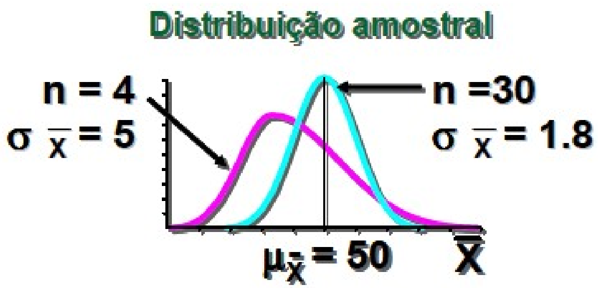
Teorema Central do Limite

TCL : Seja $x_1, x_2, \ldots, x_n$ uma amostra aleatória de uma distribuição qualquer com parâmetros $E[X] = \mu$ e $Var[X] = \sigma^2$. Se n é suficientemente grande, a distribuição de $\bar{x}$ aproxima-se a uma normal com $E[\bar{x}] = \mu$ e $Var[\bar{x}] = \sigma^2/n$ e a distribuição do Total $(x_1+\ldots+x_n)$ aproxima-se a uma distribuição normal com $E[To] = n\mu$ e $Var[To] = n\sigma^2$. Assim,

$$\bar{x} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \dot{\sim} Z$$

$$To \dot{\sim} N\left(n\mu, n\sigma^2\right) \Rightarrow \frac{To - n\mu}{\sigma\sqrt{n}} \dot{\sim} Z$$





Distribuição amostral da proporção (e frequência)

- O TCL pode ser utilizado para justificar a aproximação da Binomial pela Normal pois se X é uma variável aleatória que "conta" o número de sucessos em n tentativas independentes, como a probabilidade de sucesso é constante os Xs são iid.

- Portanto, se n é suficientemente grande

$$n^o\, sucessos = X \dot{\sim} N\left(np, np(1-p)\right) \Rightarrow \frac{X - n.p}{\sqrt{n.p.(1-p)}} \dot{\sim} Z$$

$$\frac{X}{n} = \hat{p} \dot{\sim} N\left(p, p(1-p)/n\right) \Rightarrow \frac{\hat{p} - p}{\sqrt{p.(1-p)/n}} \dot{\sim} Z$$

Distribuição amostral de s^2

- Quando há incerteza em relação ao valor de σ estima-se o seu valor por:

$$\hat{\sigma}^2 = s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 \Big/ n-1$$

- Teorema: Se s^2 é a variância de uma amostra aleatória de tamanho n, retirada de uma população normal com parâmetros μ e σ^2, então a estatística

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

tem distribuição χ^2 com n-1 graus de liberdade.

- Na utilização é necessário obter $\chi^2_{calc} = (n-1)s^2/\sigma^2$

Distribuição Qui-quadrado

A distribuição qui-quadrado é a distribuição da soma de *m* variáveis aleatórias normais padronizadas independentes ao quadrado. Essa distribuição depende de *m*, que é chamado de graus de liberdade da distribuição qui-quadrado.

Distribuição Qui-quadrado

Por exemplo, seja Z_1, Z_2 e Z_3 variáveis aleatórias normais padrão independentes. Então, $Z_1^2 + Z_2^2 + Z_3^2$ possui uma distribuição qui-quadrado com três graus de liberdade.
O nome dessa distribuição deriva da letra grega usada para representa-la: uma distribuição qui-quadrado com *m* graus de liberdade é representada por χ_m^2.
Por exemplo, o 95º percentil da distribuição χ_3^2 é 7,81, de modo que $P\left(Z_1^2 + Z_2^2 + Z_3^2 \le 7,81\right) = 0,95$.

Função Densidade de Probabilidade

$$f\left(x\right)=\frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}x^{\frac{m}{2}-1}e^{-\frac{x}{2}} \qquad x>0$$

TABLE 3 Critical Values for the χ^2 Distribution

Degrees of Freedom	Significance Level		
	10%	5%	1%
1	2.71	3.84	6.63
2	4.61	5.99	9.21
3	6.25	7.81	11.34
4	7.78	9.49	13.28
5	9.24	11.07	15.09
6	10.64	12.59	16.81
7	12.02	14.07	18.48
8	13.36	15.51	20.09
9	14.68	16.92	21.67
10	15.99	18.31	23.21
11	17.28	19.68	24.72
12	18.55	21.03	26.22

QUI - QUADRADO

$$f(x) = \frac{1}{2^{\frac{m}{2}}\,\Gamma\left(\frac{m}{2}\right)}\, x^{\frac{m}{2}-1}\,\mathrm{Exp}\left[-\frac{x}{2}\right]$$

In[8]:= $\int_0^{\infty} \frac{1}{2^{\frac{3}{2}}\,\mathrm{Gamma}\left[\frac{3}{2}\right]}\, x^{\frac{3}{2}-1}\,\mathrm{Exp}\left[-\frac{x}{2}\right]\,dx$

Out[8]= 1

In[9]:= $\int_0^{7.81} \frac{1}{2^{\frac{3}{2}}\,\mathrm{Gamma}\left[\frac{3}{2}\right]}\, x^{\frac{3}{2}-1}\,\mathrm{Exp}\left[-\frac{x}{2}\right]\,dx$

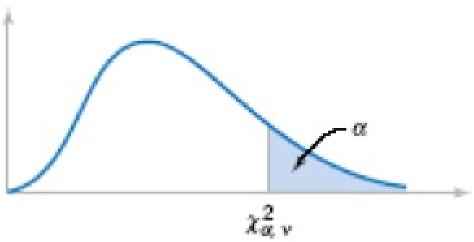
Out[9]= 0.949894

Distribuição amostral de s^2

Table III Percentage Points $\chi^2_{\alpha,\nu}$ of the Chi-Squared Distribution

ν \ α	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

ν = degrees of freedom.



Graus de liberdade

- Os graus de liberdade podem ser vistos como uma medida da informação amostral.
- Sabendo que

$$\underbrace{\sum_{i=1}^{n}(x_i-\mu)^2 \Big/ \sigma^2}_{\sim\chi^2_n} \rightarrow \underbrace{\sum_{i=1}^{n}(x_i-\bar{x})^2 \Big/ \sigma^2}_{\sim\chi^2_{n-1}}$$

quando μ não é conhecido há 1 grau de liberdade a menos ou considera-se que um grau de liberdade é perdido na estimação de μ.

- Generalização: Há n graus de liberdade, ou partes de informação independentes, em uma amostra aleatória de uma população normal e cada vez que se utiliza uma estatística em substituição a um parâmetro perde-se um grau de liberdade.

$\bar{x}$, s e a distribuição t de Student

- A dificuldade prática em relação ao TCL é que ele demanda o conhecimento de σ.

- A ferramenta utilizada para lidar com essa dificuldade é a distribuição t (William Gosset, 1908) que é formada pela divisão

$$\frac{Z}{\sqrt{\dfrac{\chi^2_{n-1}}{n-1}}} = \frac{\bar{x}-\mu / \sigma / \sqrt{n}}{\sqrt{\dfrac{(n-1)s^2/\sigma^2}{n-1}}} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} . \frac{\sigma}{s} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

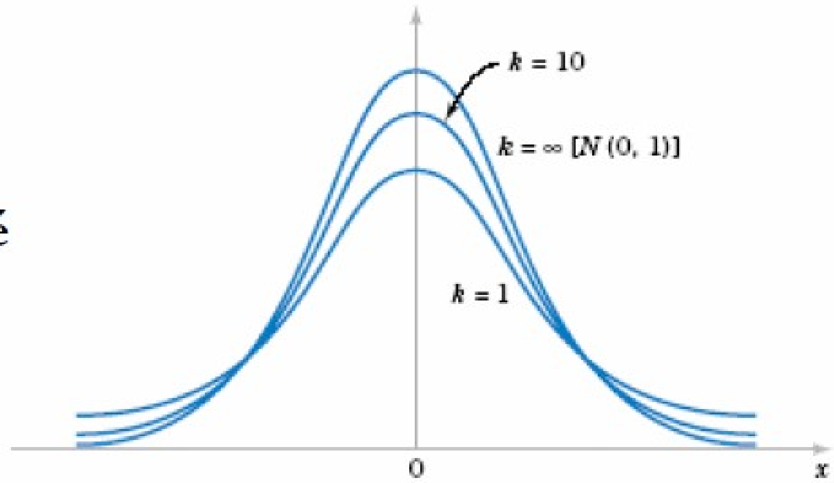
Assim, é possível fazer inferências em relação a μ usando uma estimativa de σ fornecida pela mesma amostra que gerou $\bar{x}$.

- Def: Seja $x_1, \ldots, x_n$ uma amostra aleatória de uma população normal com $E[X] = \mu$ e $Var[X] = \sigma^2$, então $T = (\bar{x}-\mu)/(s/\sqrt{n})$ tem distribuição t com n-1 graus de liberdade.

$\overline{x}$, s e a distribuição t de Student

- Propriedades das distribuições t-student: Seja t_ν a curva da função densidade dos gl de ν.

1. Cada curva t_ν possui um formato de sino e está centrada em 0.
2. Toda curva t_ν é mais dispersa que a curva normal padronizada (Z).
3. À medida que ν aumenta, a dispersão da curva t_ν correspondente diminui.
4. À medida que $\nu \rightarrow \infty$ a seqüência das curvas t_ν se aproxima da curva normal padronizada (de forma que a curva Z é chamada geralmente de t com gl= ∞).

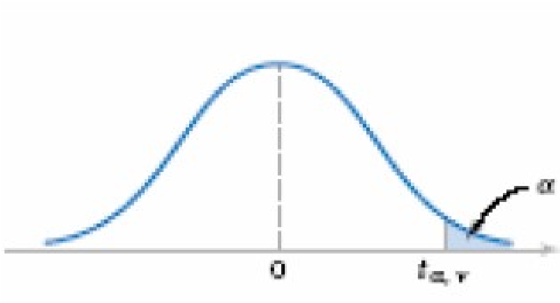


$\bar{x}, s$ e a distribuição t de Student

Table IV Percentage Points $t_{\alpha,\nu}$ of the t-Distribution

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.



Função Densidade de Probabilidade da distribuição t de Student

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

TABLE 2 Critical Values for Two-Sided and One-Sided Tests Using the Student *t* Distribution

Degrees of Freedom	Significance Level: 20% (2-Sided) 10% (1-Sided)	10% (2-Sided) 5% (1-Sided)	5% (2-Sided) 2.5% (1-Sided)	2% (2-Sided) 1% (1-Sided)	1% (2-Sided) 0.5% (1-Sided)
1	3.08	6.31	12.71	31.82	63.66
2	1.89	2.92	4.30	6.96	9.92
3	1.64	2.35	3.18	4.54	5.84
4	1.53	2.13	2.78	3.75	4.60
5	1.48	2.02	2.57	3.36	4.03
6	1.44	1.94	2.45	3.14	3.71
7	1.41	1.89	2.36	3.00	3.50
8	1.40	1.86	2.31	2.90	3.36
9	1.38	1.83	2.26	2.82	3.25
10	1.37	1.81	2.23	2.76	3.17
11	1.36	1.80	2.20	2.72	3.11
12	1.36	1.78	2.18	2.68	3.05

t de Student

$$f(x) = \frac{\Gamma\left[\frac{m+1}{2}\right]}{\sqrt{m\pi}\ \Gamma\left[\frac{m}{2}\right]} * \frac{1}{\left(1+\frac{x^2}{m}\right)^{\frac{m+1}{2}}}$$

In[30]:= $\int_{-\infty}^{\infty} \frac{\text{Gamma}\left[\frac{7+1}{2}\right]}{\sqrt{7\pi}\ \text{Gamma}\left[\frac{7}{2}\right]} * \frac{1}{\left(1+\frac{x^2}{7}\right)^{\frac{7+1}{2}}}\, dx$

Out[30]=

1

In[32]:= $\int_{-\infty}^{1.89} \frac{\text{Gamma}\left[\frac{7+1}{2}\right]}{\sqrt{7\pi}\ \text{Gamma}\left[\frac{7}{2}\right]} * \frac{1}{\left(1+\frac{x^2}{7}\right)^{\frac{7+1}{2}}}\, dx$

Out[32]=

0.949662

Distribuição F

A distribuição F tem enorme aplicação na comparação de variâncias amostrais.

As aplicações da distribuição F são encontradas em problemas envolvendo duas ou mais amostras.

A estatística F é definida como sendo a razão de duas variáveis aleatórias qui-quadrado independentes, cada uma dividida pelo seu número de graus de liberdade.

Consequentemente, nós podemos escrever

$$F = \frac{W/m}{V/n}$$

onde, W e V são variáveis aleatórias independentes tendo distribuições qui-quadrado com m e n graus de liberdade, respectivamente.

Função Densidade de Probabilidade

$$f\left(x;m,n\right)=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}m^{\frac{m}{2}}n^{\frac{n}{2}}x^{\frac{m}{2}-1}\left(mx+n\right)^{-\left(\frac{m+n}{2}\right)} \qquad x>0$$

TABLE 5B Critical Values for the F_{n_1,n_2} Distribution—5% Significance Level

Denominator Degrees of Freedom (n_2)	Numerator Degrees of Freedom (n_1)									
	1	2	3	4	5	6	7	8	9	10
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.39	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

F de Snedecor

$$f(x) = \frac{\Gamma\left[\frac{m+n}{2}\right]}{\Gamma\left[\frac{m}{2}\right]\Gamma\left[\frac{n}{2}\right]} m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1} (mx+n)^{\frac{-(m+n)}{2}}$$

In[14]:= $\int_0^{\infty} \frac{\text{Gamma}\left[\frac{3+30}{2}\right]}{\text{Gamma}\left[\frac{3}{2}\right]\text{Gamma}\left[\frac{30}{2}\right]} * 3^{\frac{3}{2}} * 30^{\frac{30}{2}} * x^{\frac{3}{2}-1} * (3*x+30)^{\frac{-(3+30)}{2}}\, dx$

Out[14]=

1

In[15]:= $\int_0^{2.92} \frac{\text{Gamma}\left[\frac{3+30}{2}\right]}{\text{Gamma}\left[\frac{3}{2}\right]\text{Gamma}\left[\frac{30}{2}\right]} * 3^{\frac{3}{2}} * 30^{\frac{30}{2}} * x^{\frac{3}{2}-1} * (3*x+30)^{\frac{-(3+30)}{2}}\, dx$

Out[15]=

0.94988

Exercício TCL

Seja X_i denotando o tempo de espera (em minutos) para o i-ésimo cliente. Um assistente afirma que μ, o tempo médio de espera de toda a população de clientes, é de 2 minutos. O gerente não acredita na afirmação de seu assistente, então ele observa uma amostra aleatória de 36 clientes. O tempo médio de espera para os 36 clientes é de 3,2 minutos. O gerente deve rejeitar a afirmação de seu assistente?