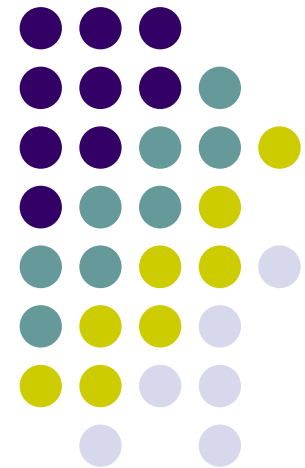


DISTRIBUIÇÕES DE PROBABILIDADE





Distribuições de Probabilidade Discretas

- Distribuição de Bernoulli
- Distribuição Binomial
- Distribuição de Poisson

Distribuições de Probabilidade Contínuas

- Distribuição Uniforme
- Distribuição Normal (Gaussiana)
- Distribuição Exponencial
- Distribuição Qui-Quadrado
- Distribuição t-student
- Distribuição F

Distribuição de Bernoulli



Uma variável aleatória tem distribuição de Bernoulli, se e só se a sua função de probabilidade é dada por:

$$f(x; \theta) = \theta^x (1 - \theta)^{1-x} \quad \text{para } x = 0 \text{ ou } 1$$

“sucesso” \Rightarrow probabilidade igual a θ

“insucesso” \Rightarrow probabilidade igual a $1 - \theta$



Distribuição Binomial

Uma variável aleatória tem distribuição Binomial, se e só se a sua função de probabilidade é dada por:

$$f(x; n, \theta) = C_x^n \theta^x (1 - \theta)^{n-x} \quad \text{para } x = 0, 1, \dots, n$$

Notação: $X \sim \text{Bin}(n, \theta)$ Média: $\mu = n\theta$ Variância: $\sigma^2 = n\theta(1 - \theta)$

“x sucessos em n tentativas”

Exemplo: Encontre a probabilidade de obter 5 caras e 7 coroas em 12 lançamentos de uma moeda equilibrada.

Resolução:

$$n = 12 \quad p = \frac{1}{2} \quad \text{"sucesso" = sair cara}$$

$$X \sim \text{Bin}\left(12, \frac{1}{2}\right)$$

$$P(x = 5) = C_5^{12} \left(\frac{1}{2}\right)^5 \left(1 - \frac{1}{2}\right)^{12-5} \cong 0.19$$



Distribuição de Poisson

Uma variável aleatória tem distribuição de Poisson, se e só se a sua função de probabilidade é dada por:

$$f(x; \lambda) = \frac{\lambda^x \cdot e^{-\lambda}}{x!} \quad \text{para } x = 0, 1, 2, \dots$$

Notação: $X \sim Poi(\lambda)$

Média: $\mu = \lambda$

Variância: $\sigma^2 = \lambda$

 **APROXIMAÇÃO À BINOMIAL:** $\begin{cases} n \geq 20 & \text{e} & p \leq 0.05 \\ n \geq 100 & \text{e} & n \cdot p < 10 \end{cases}$



Exemplo 1: Aproximação à binomial

$$n = 150 \quad p = 0.05 \quad \lambda = 7.5$$

x $P(X = x)$ Binomial $P(X = x)$ Poisson

0 0.0005 0.0006

6 0.1384 0.1367

12 0.0355 0.0366

Exemplo 2: Se 2% dos livros de uma certa impressora têm defeitos, determine a probabilidade de que 5 de entre 400 livros tenham defeito.

$X \rightarrow$ v. a. que designa n° de livros com defeito

$$x = 5 \quad p = 0.02 \quad \lambda = n.p = 400 \times 0.02 = 8$$

$$P(x = 5; \lambda = 8) = \frac{8^5 \cdot e^{-8}}{5!} = 0.0916$$

A distribuição de Poisson pode servir de modelo para o número de sucessos que ocorre durante um dado intervalo de tempo ou uma região específica, quando:

- o número de sucessos ocorrendo em intervalos não sobrepostos são independentes;
- a probabilidade de um único sucesso ocorrendo num certo intervalo é proporcional ao comprimento do intervalo;
- a probabilidade de mais de um sucesso ocorrer num pequeno intervalo é negligível.



Exemplo 3: O número de pessoas que chegam a um determinado Centro de Saúde numa dada cidade num período de 15 minutos é 12. Qual a probabilidade de num qualquer período de 15 minutos cheguem a este C.S. menos de 9 pessoas?

$X \rightarrow$ v. a. que designa nº de pessoas que chegam ao C.S./15 minutos

$$\lambda = 12$$

$$P(x < 9; \lambda = 12) = \sum_{x=0}^8 p(x; \lambda = 12) = p(x=0) + p(x=1) + \dots + p(x=8) = 0.1550$$

Exemplo 4: Encontre a probabilidade de 7 de 10 pessoas recuperarem de uma doença tropical, assumindo independência, e com probabilidade de 0.8 que qualquer um deles recupere da doença.

$X \rightarrow$ v. a. que designa nº de pessoas que recuperam

$n = 10$ $p = 0.8$ "sucesso" = recuperar da doença

$$X \sim \text{Bin}(10, 0.8)$$

$$P(x = 7) = C_7^{10} (0.8)^7 (1 - 0.8)^{10-7} = 0.20$$



Distribuição Uniforme

Uma variável aleatória contínua segue a distribuição uniforme se e só se a sua função densidade é dada por

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha < x < \beta \\ 0 & \text{outros valores} \end{cases}$$

$$\mu = \frac{\beta + \alpha}{2}$$

$$\sigma^2 = \frac{1}{12}(\beta - \alpha)^2$$

Exemplo...



Uma fábrica produz folhas de cartão com uma espessura uniforme entre 0.8 e 1.2 cm. Qual a percentagem de folhas abaixo de 1 cm?

$$P(X < 1) = \int_{0.8}^1 \frac{1}{1.2 - 0.8} dx = \frac{1}{0.4} x \Big|_{0.8}^1 = \frac{0.2}{0.4} = 0.5$$



Distribuição Exponencial Negativa

Uma variável aleatória segue a distribuição Exponencial Negativa, se e só se a sua função densidade de probabilidade é dada por:

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & , x > 0 \text{ e } \theta > 0 \\ 0 & x \leq 0 \end{cases}$$

Notação: $X \sim EN\left(\frac{1}{\theta}\right)$

Média: $\mu = \theta$

Variância: $\sigma^2 = \theta^2$

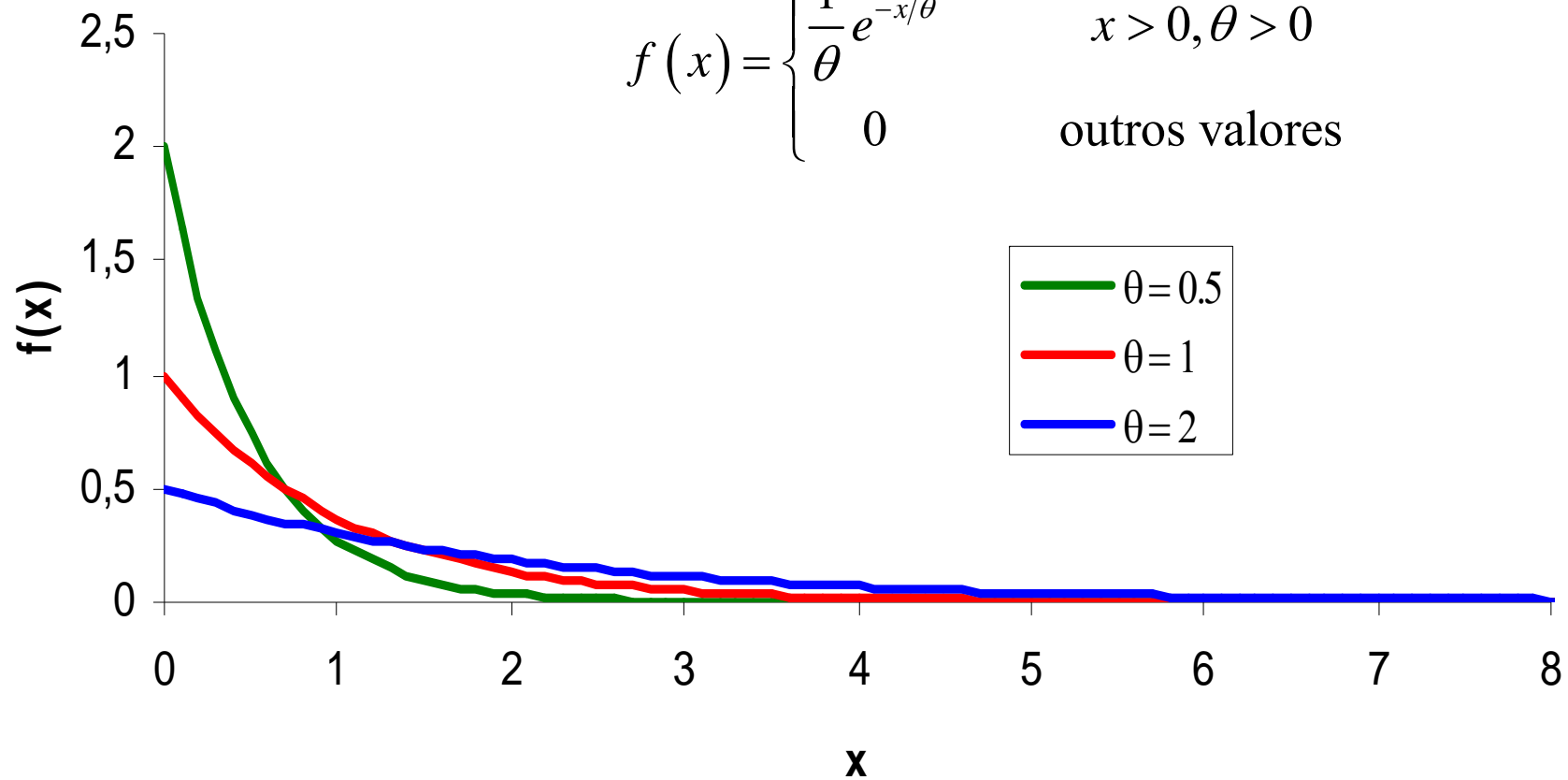
Função distribuição acumulada:

$$F(x; \theta) = \begin{cases} 1 - e^{-\frac{x}{\theta}} & , x > 0 \text{ e } \theta > 0 \\ 0 & x \leq 0 \end{cases}$$

Exponencial



$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & x > 0, \theta > 0 \\ 0 & \text{outros valores} \end{cases}$$





Exemplo...

- Um componente electrónico requer, em média, uma reparação de 2 em 2 anos. Qual a probabilidade de que funcione por pelo menos 3 anos?

$$P(X > 3) = 1 - P(X < 3) = 1 - \int_0^3 \frac{1}{2} e^{-\frac{x}{2}} dx = 1 - \left(-e^{-\frac{x}{2}} \right) \Big|_0^3 = e^{-\frac{3}{2}} = 0.2231$$

- Sabendo que o componente dura há já dois anos, qual a probabilidade de funcionar durante mais um ano?

$$P(X > 3 | X > 2) = \frac{P(X > 3 \cap X > 2)}{P(X > 2)} = \frac{P(X > 3)}{P(X > 2)} = \frac{e^{-\frac{3}{2}}}{e^{-\frac{2}{2}}} = e^{-\frac{1}{2}} = 0.6065$$



Distribuição Normal ou Gaussiana

Uma variável aleatória segue a distribuição Normal, se e só se a sua função densidade de probabilidade é dada por:

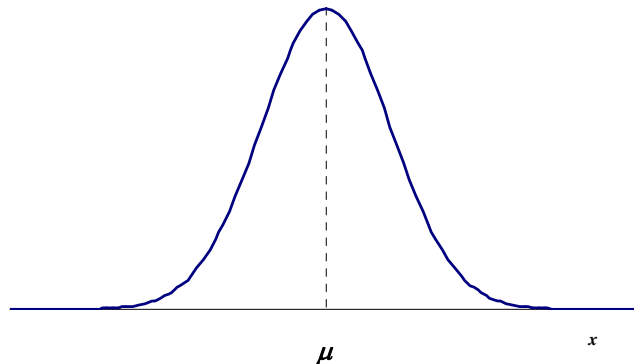
$$f(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \quad -\infty < x < +\infty \text{ com } \sigma > 0$$

Notação: $X \sim N(\mu, \sigma^2)$

Média: $\mu = \mu_X$

Variância: $\sigma^2 = \sigma_X^2$

Gráfico da distribuição Normal





A distribuição Normal em $\mu = 0$ e $\sigma = 1$ é designada por **Normal Padrão** ou **Standard**, e tem como função densidade de probabilidade:

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

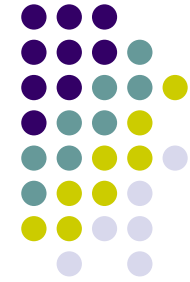
$$Z \sim N(0,1)$$

Se X tem uma distribuição Normal com média μ e desvio padrão σ , então:

$$Z = \frac{X - \mu_X}{\sigma_X} \sim N(0,1)$$

tem uma distribuição Normal Padrão.

NORMAL



Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,5	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010	0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019	0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036	0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048	1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064	1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084	1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110	1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143	1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183	1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233	1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294	1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367	1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455	1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559	2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681	2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823	2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985	2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170	2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379	2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611	2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867	2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148	2,8	0,9974	0,9975	0,9976	0,9977	0,9978	0,9979	0,9980	0,9981	0,9982	0,9983
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451	2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776	3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121	3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483	3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859	3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247	3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641	3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998



Aproximação Normal à distribuição Binomial

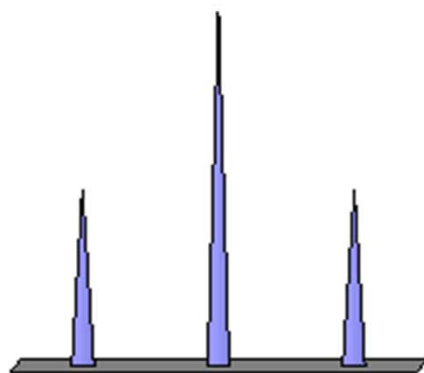
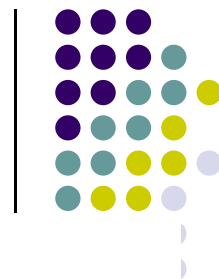
A distribuição Normal fornece uma boa aproximação à distribuição Binomial quando n , número de tentativas é grande, e π , a probabilidade de um sucesso numa tentativa é próxima de 0.5.

Teorema: Se X é uma variável aleatória seguindo uma distribuição binomial com parâmetros n e π , então:

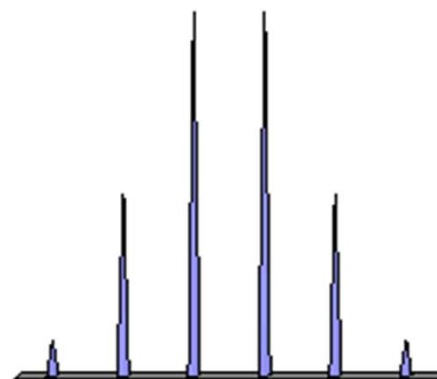
$$Z = \frac{X - n.\pi}{\sqrt{n.\pi.(1 - \pi)}}$$

aproxima-se da distribuição Normal quando $n \rightarrow \infty$.

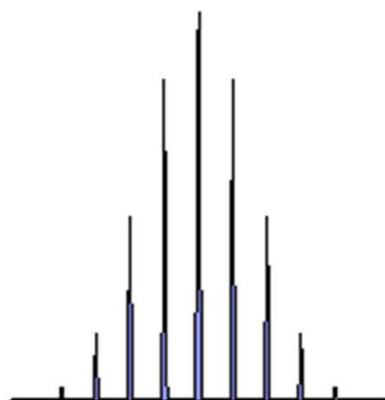
Condições: $n.\pi > 5$ e $n.(1 - \pi) > 5$



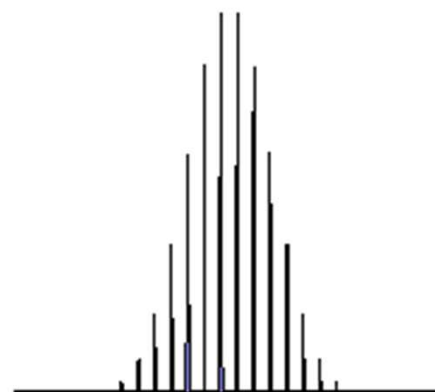
$n = 2, \pi = 0.5$



$n = 5, \pi = 0.5$



$n = 10, \pi = 0.5$



$n = 25, \pi = 0.5$

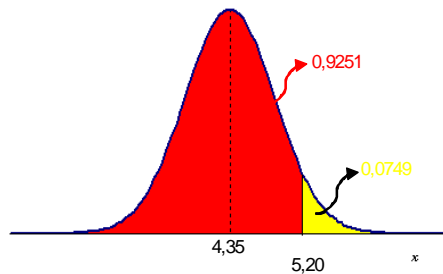


Exemplo: Suponha que a quantidade de radiação cósmica que uma pessoa é exposta ao viajar de avião é uma variável aleatória Normal com média $\mu = 4.35 \text{ mrcm}$ e o desvio padrão é $\sigma = 0.59 \text{ mrcm}$. Qual é a probabilidade de que uma pessoa seja exposta a mais de 5.20 mrcm de radiação cósmica?

$$X \sim N(4.35, 0.59^2)$$

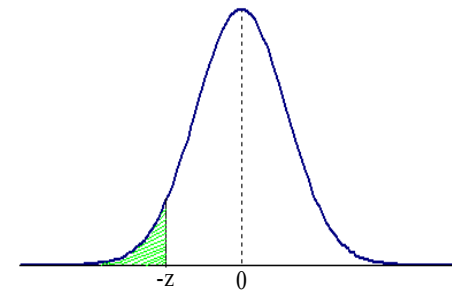
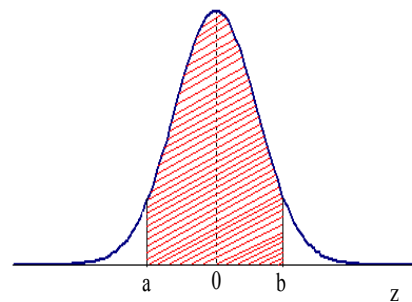
$$z = \frac{5.20 - 4.35}{0.59} = 1.44$$

$$\begin{aligned} P(X > 5.20) &= P(Z > 1.44) = \\ &= 1 - P(Z < 1.44) = 1 - 0.9251 = 0.0749 \end{aligned}$$



Algumas propriedades:

- $P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$
- $\Phi(-z) = 1 - \Phi(z)$





Exemplo: Use a aproximação Normal à distribuição binomial para determinar a probabilidade de obter 6 caras e 10 coroas em 16 lançamentos de uma moeda equilibrada.

$X \rightarrow$ v. a. que designa nº de caras

$n=16$ $\pi=0.5$ "sucesso" = sair cara

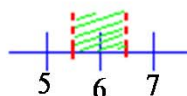
$X \sim \text{Bin}(16, 0.5)$

$$P(x=6) = C_6^{16} (0.5)^6 (1-0.5)^{16-6} = 0.1222$$

Utilizando a aproximação:

$$\mu = n.\pi = 16 * 0.5 = 8$$

$$\sigma = \sqrt{n.\pi.(1-\pi)} = 2$$



$$\begin{aligned} P(x=6) &= P(5.5 < x < 6.5) = P(-1.25 < z < 0.75) = \\ &= \Phi(0.75) - \Phi(-1.25) = 0.2266 - 0.1056 = 0.1210 \end{aligned}$$

Distribuição do Qui-quadrado

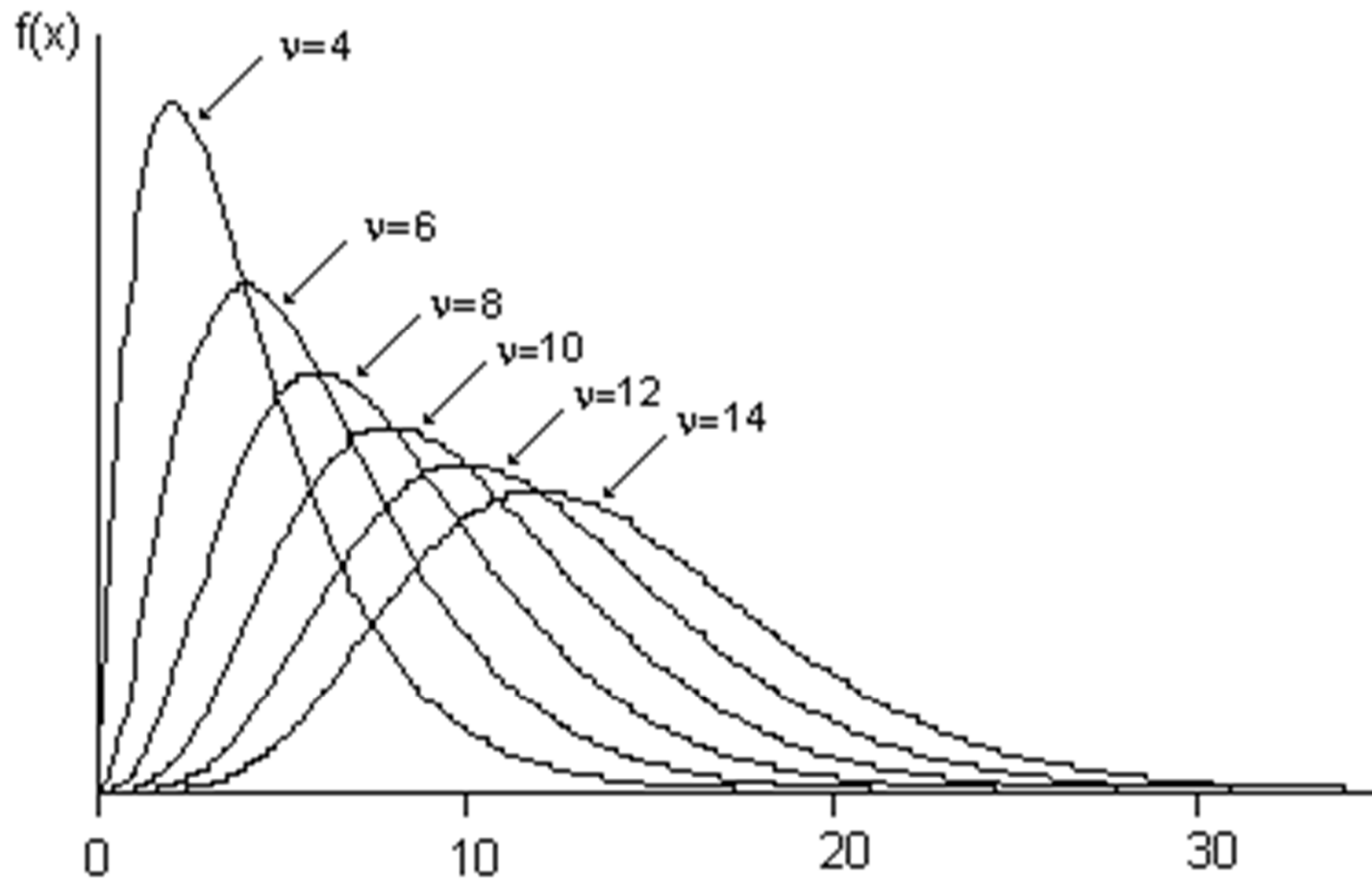


Uma variável aleatória segue a distribuição de Qui-quadrado com ν graus de liberdade, se a sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{\frac{\nu-2}{2}} e^{-x/2} & x > 0 \\ 0 & \text{outros valores} \end{cases}$$

$$\mu = \nu \qquad \sigma^2 = 2\nu$$

Qui-quadrado



QUI-QUADRADO



g.l.	0,995	0,990	0,975	0,950	0,050	0,025	0,010	0,005
1	0,000	0,000	0,001	0,004	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	43,773	46,979	50,892	53,672
40	20,707	22,164	24,433	26,509	55,758	59,342	63,691	66,766
50	27,991	29,707	32,357	34,764	67,505	71,420	76,154	79,490
60	35,534	37,485	40,482	43,188	79,082	83,298	88,379	91,952
70	43,275	45,442	48,758	51,739	90,531	95,023	100,425	104,215
80	51,172	53,540	57,153	60,391	101,879	106,629	112,329	116,321
90	59,196	61,754	65,647	69,126	113,145	118,136	124,116	128,299
100	67,328	70,065	74,222	77,929	124,342	129,561	135,807	140,169



Qui-quadrado

- Se X segue uma distribuição normal padrão, então X^2 segue a distribuição de Qui-Quadrado com 1 grau de liberdade.

- Se X_1, X_2, \dots, X_n são variáveis aleatórias independentes que seguem uma distribuição de Qui-Quadrado com v_1, v_2, \dots, v_n , graus de liberdade, então

$$Y = \sum X_i$$

segue a distribuição de Qui-Quadrado com $v_1 + v_2 + \dots + v_n$, graus de liberdade.



QUI-QUADRADO

Se \bar{x} e s^2 são a média e a variância de uma amostra aleatória de tamanho n de uma população normal com média μ e desvio padrão σ , então

- \bar{x} e s^2 são independentes,
- a variável aleatória $(n-1)s^2 / \sigma^2$ segue uma distribuição de Qui-Quadrado com $n-1$ graus de liberdade.



Distribuição t-Student

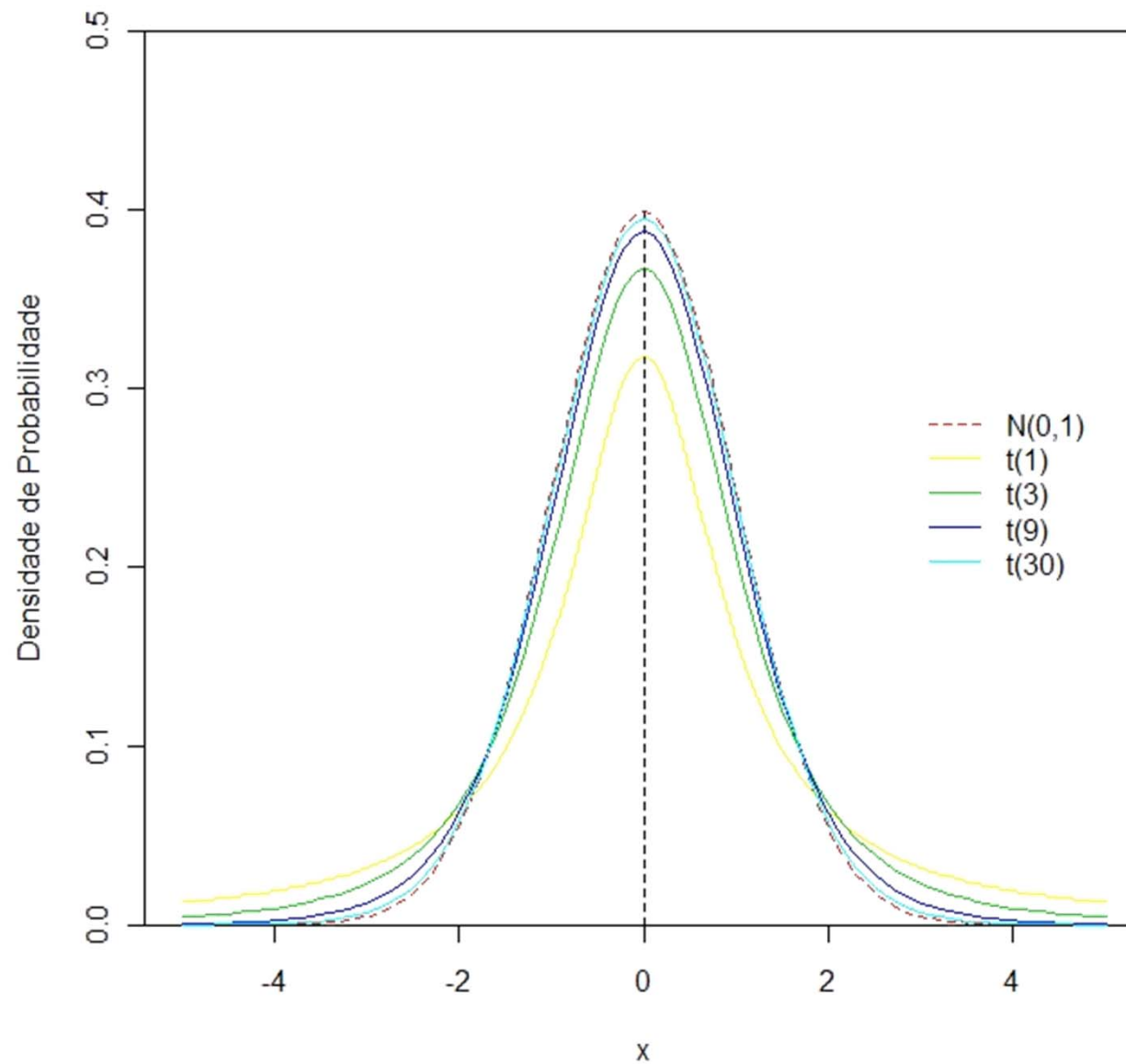
Se y e z são variáveis aleatórias independentes, y com uma distribuição de Qui-quadrado com ν graus de liberdade e z uma distribuição normal padrão, então a distribuição de

$$t = \frac{z}{\sqrt{y/\nu}}$$

é dada por

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

t-Student





t-Student

Se \bar{x} e s^2 são a média e a variância de uma amostra aleatória de tamanho n de uma população normal com média μ e desvio padrão σ , então

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

segue uma distribuição t-Student com $n-1$ graus de liberdade.

t-Student



g.l.	0,25	0,15	0,10	0,05	0,025	0,010	0,005
1	1,000	1,963	3,078	6,314	12,706	31,821	63,657
2	0,816	1,386	1,886	2,920	4,303	6,965	9,925
3	0,765	1,250	1,638	2,353	3,182	4,541	5,841
4	0,741	1,190	1,533	2,132	2,776	3,747	4,604
5	0,727	1,156	1,476	2,015	2,571	3,365	4,032
6	0,718	1,134	1,440	1,943	2,447	3,143	3,707
7	0,711	1,119	1,415	1,895	2,365	2,998	3,499
8	0,706	1,108	1,397	1,860	2,306	2,896	3,355
9	0,703	1,100	1,383	1,833	2,262	2,821	3,250
10	0,700	1,093	1,372	1,812	2,228	2,764	3,169
11	0,697	1,088	1,363	1,796	2,201	2,718	3,106
12	0,695	1,083	1,356	1,782	2,179	2,681	3,055
13	0,694	1,079	1,350	1,771	2,160	2,650	3,012
14	0,692	1,076	1,345	1,761	2,145	2,624	2,977
15	0,691	1,074	1,341	1,753	2,131	2,602	2,947
16	0,690	1,071	1,337	1,746	2,120	2,583	2,921
17	0,689	1,069	1,333	1,740	2,110	2,567	2,898
18	0,688	1,067	1,330	1,734	2,101	2,552	2,878
19	0,688	1,066	1,328	1,729	2,093	2,539	2,861
20	0,687	1,064	1,325	1,725	2,086	2,528	2,845
21	0,686	1,063	1,323	1,721	2,080	2,518	2,831
22	0,686	1,061	1,321	1,717	2,074	2,508	2,819
23	0,685	1,060	1,319	1,714	2,069	2,500	2,807
24	0,685	1,059	1,318	1,711	2,064	2,492	2,797
25	0,684	1,058	1,316	1,708	2,060	2,485	2,787
26	0,684	1,058	1,315	1,706	2,056	2,479	2,779
27	0,684	1,057	1,314	1,703	2,052	2,473	2,771
28	0,683	1,056	1,313	1,701	2,048	2,467	2,763
29	0,683	1,055	1,311	1,699	2,045	2,462	2,756
30	0,683	1,055	1,310	1,697	2,042	2,457	2,750
40	0,681	1,050	1,303	1,684	2,021	2,423	2,704
60	0,679	1,045	1,296	1,671	2,000	2,390	2,660
120	0,677	1,041	1,289	1,658	1,980	2,358	2,617
∞	0,674	1,036	1,282	1,645	1,960	2,326	2,576



Distribuição F

Se U e V são variáveis aleatórias independentes seguindo distribuições de Qui-Quadrado com ν_1 e ν_2 graus de liberdade, então

$$x = \frac{U/\nu_1}{V/\nu_2}$$

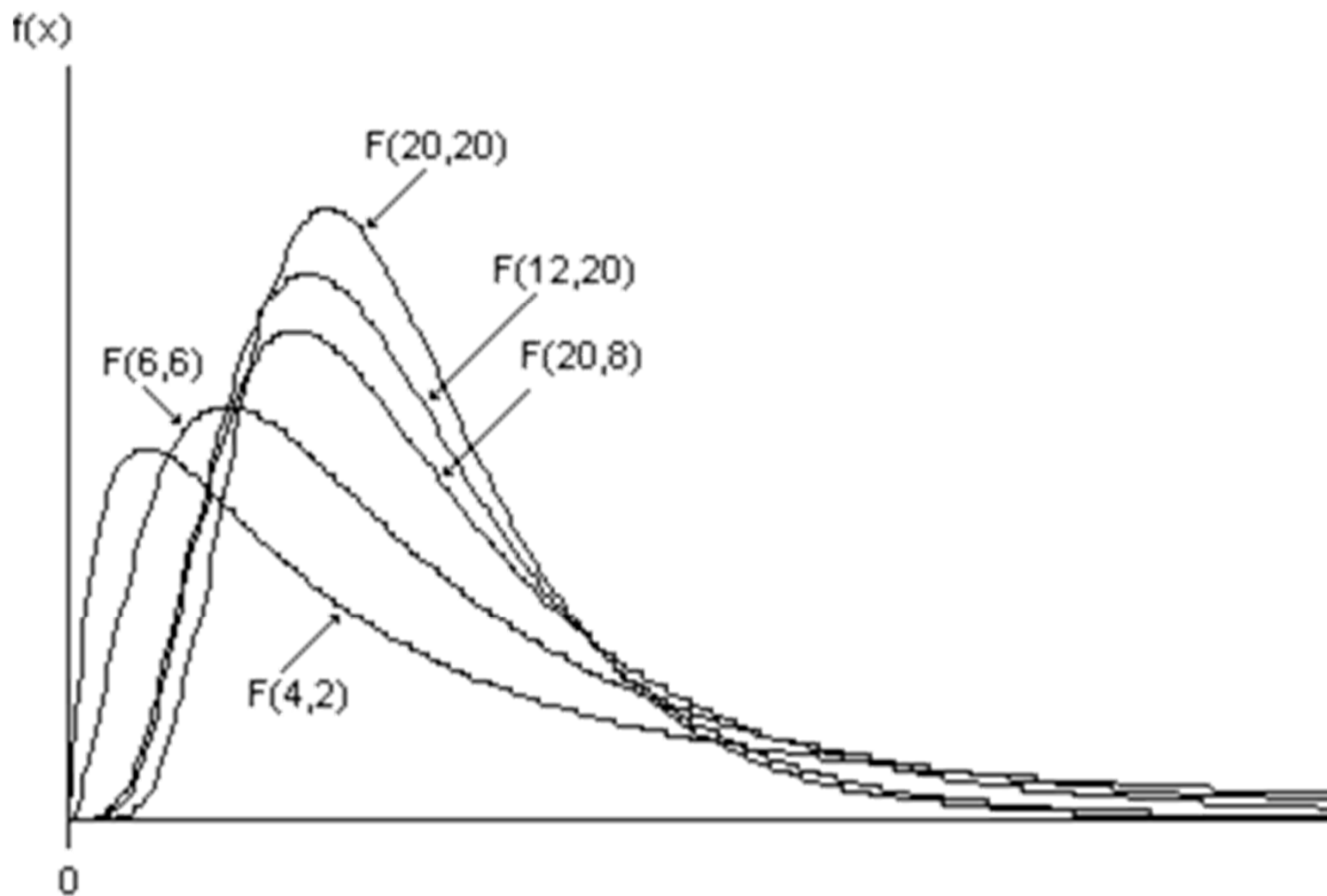
é uma variável aleatória seguindo a distribuição F com ν_1 e ν_2 graus de liberdade.

Distribuição F

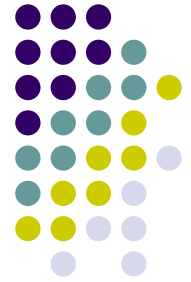


$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right) x^{\frac{\nu_1}{2}-1} \left(1 + \frac{\nu_1}{\nu_2}x\right)^{-\frac{1}{2}(\nu_1 + \nu_2)} & x > 0 \\ 0 & \text{outros valores} \end{cases}$$

Distribuição F



DISTRIBUIÇÃO F



Se s_1^2 e s_2^2 são as variâncias de variáveis aleatórias independentes de dimensão n_1 e n_2 de populações normais com variâncias σ_1^2 e σ_2^2 , então

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

é uma variável aleatória seguindo a distribuição F com n_1-1 e n_2-1 graus de liberdade.



Teorema do Limite Central:

Se X_1, X_2, \dots, X_n formam uma sequência de n variáveis independentes, com médias e variâncias respectivamente iguais a μ_{X_i} e $\sigma_{X_i}^2$, $i = 1, 2, \dots, n$ e se construirmos uma v.a. U ,

$U = X_1 + X_2 + \dots + X_n$ então a estatística

$$Z = \frac{U - \sum_{i=1}^n \mu_{X_i}}{\sqrt{\sum_{i=1}^n \sigma_{X_i}^2}}$$

tem uma distribuição assintótica $N(0,1)$.

Deste teorema resulta que:

– Se \bar{X} é a média da amostra aleatória de tamanho n , retirada de uma população

Normal com média μ_X e variância σ_X^2 , então:

$$Z = \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{n}}} \underset{n \rightarrow \infty}{\sim} N(0,1)$$

ESTIMAÇÃO



- ESTIMAÇÃO PONTUAL
- ESTIMAÇÃO POR INTERVALOS

Objectivo da estimação pontual

Consiste em tentar encontrar a “estatística”, cujo valor numérico, obtido através dos dados da amostra, esteja próximo do parâmetro da população, que é constante mas desconhecido.

$\theta \rightarrow$ parâmetro da população

$\hat{\theta} \rightarrow$ estimador pontual para θ

PROPRIEDADES DE UM ESTIMADOR



- TENDÊNCIA NULA (NÃO TENDENCIOSO, CENTRADO, NÃO ENVIESADO)
- MÉDIA QUADRÁTICA DO ERRO MÍNIMA
- EFICIENTE
- CONSISTENTE
- SUFICIENTE
- ROBUSTO

Tendência



$$t_T(\theta)$$

$$t_T(\theta) = E[T] - \theta$$

Diz-se que uma estatística T é um estimador não tendencioso (ou centrado) em relação ao parâmetro θ , se e só se:

$$t_T(\theta) = 0 \Leftrightarrow E[T] = \theta$$



Exemplo:

$X \sim \text{bin}(n, \theta)$. Mostrar que $\frac{X}{n}$ é um estimador não tendencioso de θ .

Resolução:

$$E[X] = n\theta$$

$$E\left[\frac{X}{n}\right] = \frac{1}{n}E[X] = \frac{1}{n}n\theta = \theta$$

$\therefore T = \frac{X}{n}$ é um estimador não tendencioso para θ

Exemplo:

Se X_1, X_2, \dots, X_n constituem uma amostra aleatória duma população dada por

$$f(x) = \begin{cases} e^{-(x-\theta)} & x > \theta \\ 0 & \text{outros valores} \end{cases}$$

Mostre que $T = \bar{X}$ é um estimador tendencioso de θ .

Resolução:

$$T = \frac{\sum_{i=1}^n X_i}{n} \Rightarrow E[T] = E\left[\frac{\sum_{i=1}^n X_i}{n}\right] = \frac{1}{n} \cdot E\left[\sum_{i=1}^n X_i\right] = \frac{1}{n} \cdot \sum_{i=1}^n E[X_i] = E[X_i]$$

$$\mu = E[X_i] = \int_{\theta}^{+\infty} x \cdot f(x) dx = \int_{\theta}^{+\infty} x \cdot e^{-(x-\theta)} dx = \left[-x \cdot e^{-(x-\theta)} \right]_{\theta}^{+\infty} - \int_{\theta}^{+\infty} -e^{-(x-\theta)} dx = 1 + \theta$$

$$E[T] = E[X_i] = 1 + \theta$$

$$t_T(\theta) = E[T] - \theta \Leftrightarrow t_T(\theta) = 1 \neq \theta \Rightarrow T \text{ é um estimador tendencioso para } \theta.$$

Se se considerar $T' = \bar{X} - 1$ então T' é não tendencioso, pois $E[T'] = E[\bar{X}] - 1 = \theta$.



MÉDIA QUADRÁTICA DO ERRO (MQE)



A medida, do desempenho de um estimador, mais utilizada é a média quadrática do erro, definida por:

$$MQE = E\left[(T - \theta)^2\right]$$

$$E\left[(T - \theta)^2\right] = \text{var}[T] + \underbrace{\left(E[T] - \theta\right)^2}_{t_T(\theta)}$$

QUANDO O ESTIMADOR É NÃO TENDENCIOSO A MQE RESUME-SE À VARIÂNCIA DO ESTIMADOR.

UM “BOM” ESTIMADOR CORRESPONDE ÀQUELE QUE POSSUIR MENOR MQE.

Eficiência



Se T_1 e T_2 são dois estimadores não tendenciosos do parâmetro θ duma população e se $\text{var}[T_1] < \text{var}[T_2]$, diz-se que T_1 é relativamente mais eficiente que T_2 .

$$ef(T_1, T_2) = \frac{\text{var}[T_1]}{\text{var}[T_2]}$$

Se T é um estimador tendencioso dum dado parâmetro θ , as comparações devem ser feitas com base na média quadrática do erro.

$$ef(T_1, T_2) = \frac{MQE[T_1]}{MQE[T_2]}$$

Consistência



A estatística T é um estimador consistente do parâmetro θ se e só se para cada $c > 0$,

$$\lim_{n \rightarrow \infty} P(|T - \theta| < c) = 1$$

De notar que a consistência é uma propriedade assintótica.

Se T é um estimador não tendencioso do parâmetro θ e $\text{var}[T] \rightarrow 0$ à medida que $n \rightarrow \infty$, então T é um estimador consistente de θ .

Suficiência



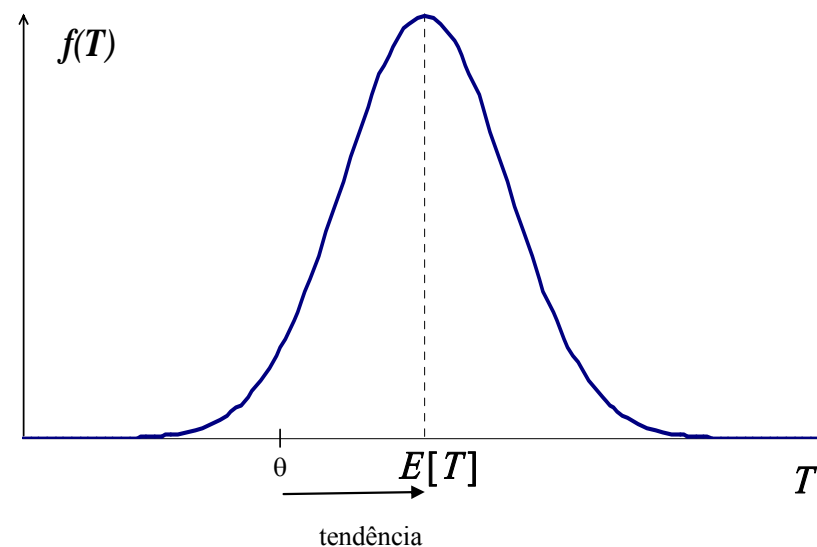
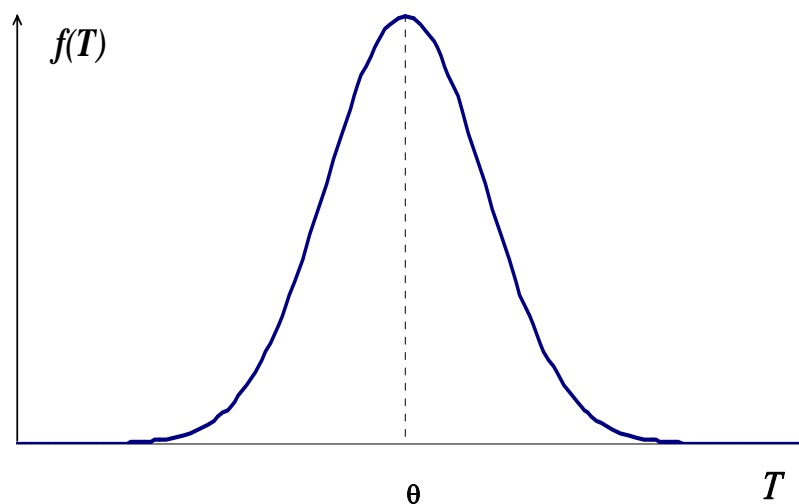
Um estimador é suficiente se toda a informação na amostra relevante para a estimação de θ , isto é, se todo o conhecimento acerca de θ que pode ser ganho a partir dos valores individuais e da sua ordem, pode também ser ganho pelo valor de T por si só.

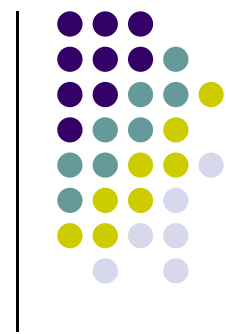
A estatística T é um estimador suficiente do parâmetro θ se e só se para cada valor de T a probabilidade condicional da amostra aleatória X_1, X_2, \dots, X_n dado $T = t$ é independente de θ .



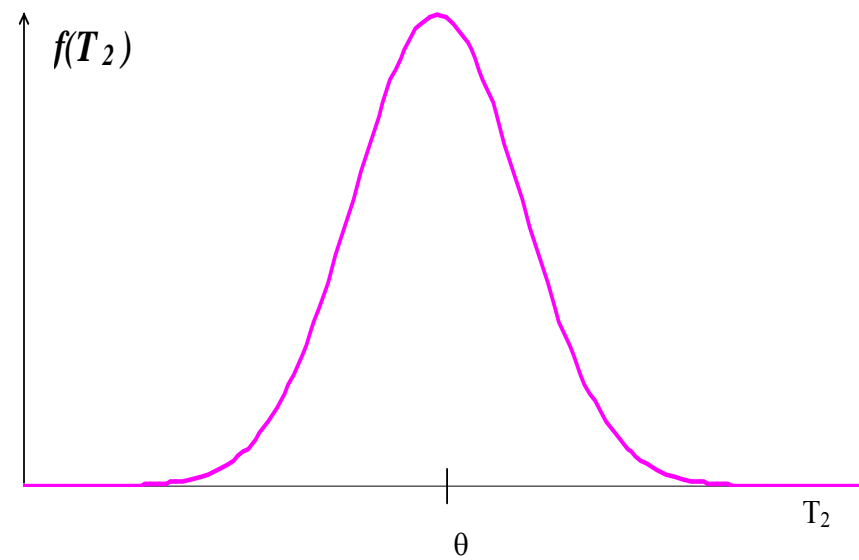
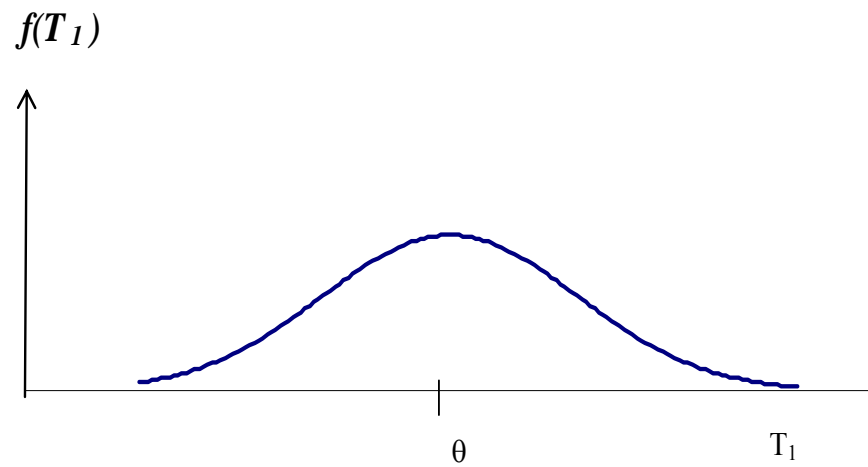
Estimadores Pontuais: Propriedades

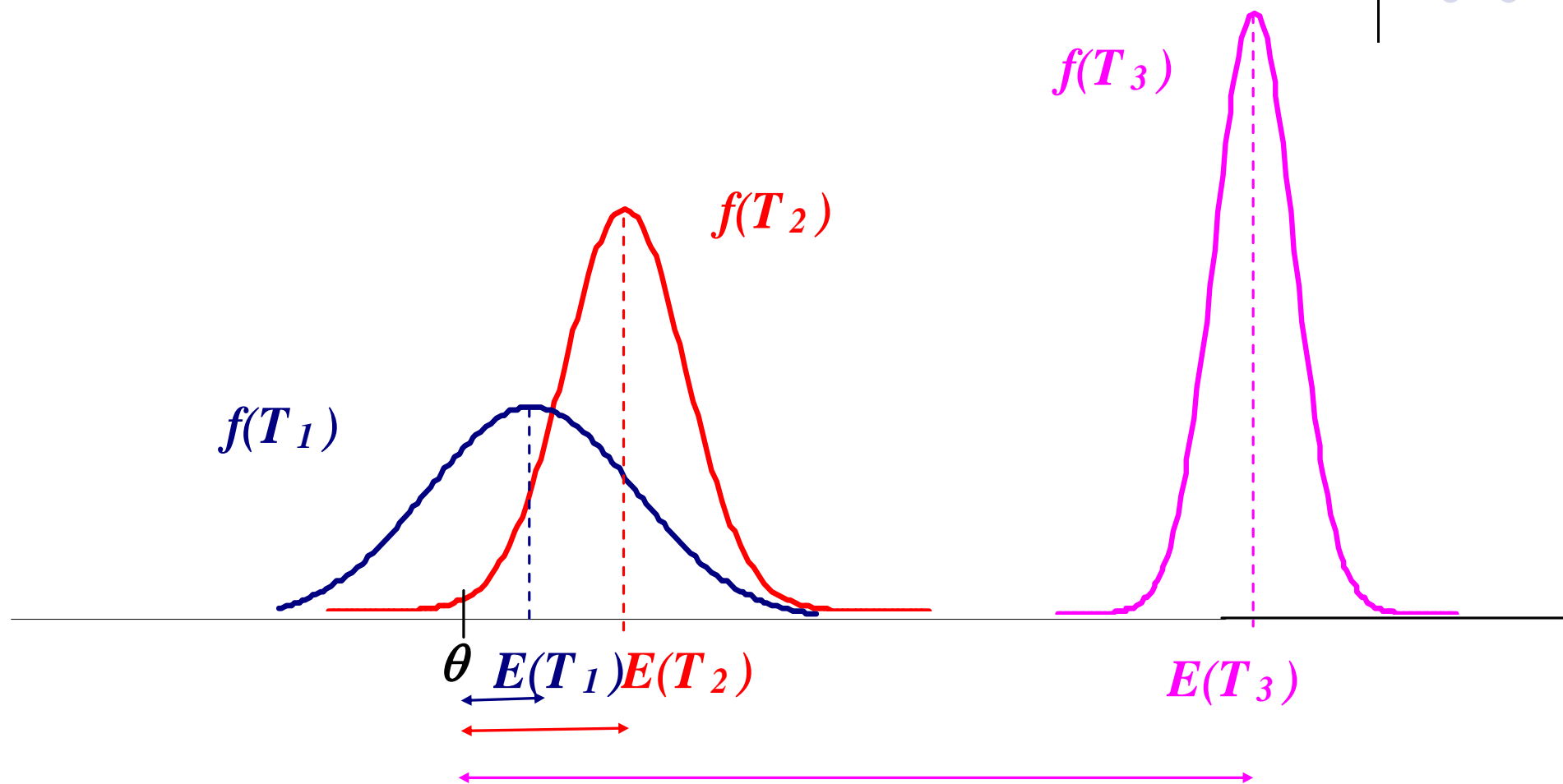
Não Tendenciosos $t_T(\theta) = 0 \Leftrightarrow E[T] = \theta$





VARIÂNCIA MÍNIMA





MÉTODOS DE ESTIMAÇÃO



- Método da Máxima Verosimilhança
- Método dos Mínimos quadrados

Definição: Se x_1, x_2, \dots, x_n são os valores de uma a.a. de uma população com parâmetro θ , a função de verosimilhança, $L(\theta)$, é dada por:

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta)$$

para valores de θ no domínio dado, $f(x_1, x_2, \dots, x_n; \theta)$ é o valor da função (densidade) de probabilidade conjunta das variáveis aleatórias x_1, x_2, \dots, x_n observadas.



ESTIMADOR DE MV

ALGORITMO DE CÁLCULO:

1º - Determinação da função de verosimilhança

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta) = f(x_1; \theta) \cdot f(x_2; \theta) \cdot \dots \cdot f(x_n; \theta)$$

2º - O máximo de $L(\theta)$ é também o máximo de $\ln L(\theta)$

$$\ln L(\theta) = \ln \left(\prod_{i=1}^n f(x_i; \theta) \right)$$

3º - Derivar em ordem ao parâmetro(s) para determinar o máximo

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \Rightarrow \theta = \dots$$



Exemplo: Sejam x_1, x_2, \dots, x_n os valores de uma amostra de uma distribuição exponencial negativa com f.d.p dada por:

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & , x > 0 \text{ e } \theta > 0 \\ 0 & x \leq 0 \end{cases}$$

determine o estimador de máxima verosimilhança para θ .

Resolução:

$$L(\theta) = \prod_{i=1}^n \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \frac{1}{\theta} e^{-\frac{x_1}{\theta}} \times \frac{1}{\theta} e^{-\frac{x_2}{\theta}} \times \dots \times \frac{1}{\theta} e^{-\frac{x_n}{\theta}} = \left(\frac{1}{\theta}\right)^n e^{-\frac{1}{\theta} \sum_{i=1}^n x_i}$$

$$\ln L(\theta) = -n \ln \theta - \frac{1}{\theta} \sum_{i=1}^n x_i \quad \frac{d \ln L(\theta)}{d\theta} = 0 \Leftrightarrow -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i = 0 \Leftrightarrow \frac{1}{\theta} \left(-n + \frac{1}{\theta} \sum_{i=1}^n x_i \right) = 0 \Leftrightarrow \theta = \frac{\sum_{i=1}^n x_i}{n}$$

$$\therefore \hat{\theta} = \bar{X}$$