Distribuições Contínuas de Probabilidade

Uma variável aleatória contínua é uma função definida sobre o espaço amostral, que associa valores em um intervalo de números reais.

Exemplos:

- Espessura de um item
- Tempo necessário para completar um teste
- Tempo de espera numa fila
- Peso

Para variáveis aleatórias contínuas introduzimos a **função densidade de probabilidade (fdp),** tal que,

$$\text{(a)} \quad f(x) \geq 0$$

$$\text{(b)} \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

A condição (a) implica que a densidade é uma função não negativa e condição (b) corresponde ao fato de que a soma (integral no caso de variáveis contínuas) das probabilidades é igual a um. A integração de ±oo significa que devemos integrar sobre todos os valores de x em que f(x) é definida.

Qualquer função f(x) satisfazendo (a) e (b) é uma fdp.

Note que para variáveis contínuas as somas são substituídas por integrais.

Para variáveis contínuas não podemos obter a probabilidade de x ter o valor num ponto $(x = a)$, ou seja, temos que considerar a probabilidade de x assumir valores num intervalo $a < x < b$.

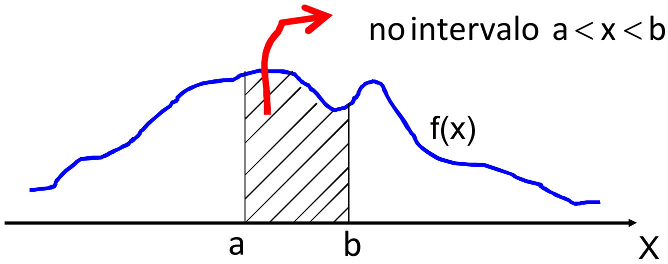
$$P(a < x < b) = \int_a^b f(x)\, dx$$ representa a probabilidade de x assumir valores no intervalo $a < x < b$.

Note que, como a probabilidade da variável x assumir valor num dado ponto é nula, temos que: $P(a < x < b) = P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b)$

Geometricamente

Área entre a fdp f(x) e o eixo x

no intervalo $a < x < b \; = \; P(a < x < b)$



Exemplo: Dada a função

$$f(x) \; = \begin{cases} \frac{1}{3}, & 1 \le x \le 4 \\ 0, & \text{outros valores} \end{cases}$$

(a) verifique se f(x) é uma fdp.

(b) Calcule a probabilidade de x assumir vaores no intervalo $2 < x < 3$

(a) Temos que $f(x) \ge 0$ e $\int_{-\infty}^{+\infty} f(x)\, dx = \int_1^4 \frac{1}{3}\, dx = \frac{1}{3}\, x\Big|_1^4 = \frac{1}{3}\,(4\text{-}1) = 1$

portanto f(x) é uma fdp.

(b) Temos: $P(2 < x < 3) = \int_2^3 \frac{1}{3}\, dx = \frac{1}{3}\, x\Big|_2^3 = \frac{1}{3}\,(3\text{-}2) = \frac{1}{3}$

Do mesmo modo que para variáveis discretas, podemos definir para variáveis contínuas a média, a variância e o desvio padrão. Lembando que introduzimos a fdp associada a variável aleatória considerada e substituímos as somas por integrais.

Assim, para a variável aleatória X, temos:

$$\text{Média}\,(\mu): \qquad \mu = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

$$\text{Variância}\,(\sigma^2): \qquad \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx$$

$$\text{Desvio-padrão:} \qquad \sigma = \sqrt{\sigma^2}$$

Exemplo: Considere a fdp considerada anteriormente

$$f(x) = \begin{cases} \frac{1}{3}, & 1 \le x \le 4 \\ 0, & \text{outros valores} \end{cases}$$

(a) Calcule a média da variável aleatória X.

(b) Calcule a variância e o desvio padrão de X.

(a) Temos: $\mu = \int_{-\infty}^{+\infty} x\, f(x)\, dx = \int_1^4 x\left(\frac{1}{3}\right) dx = \frac{1}{3}\int_1^4 x\, dx = \frac{1}{3}\left(\frac{x^2}{2}\right)\Bigg|_1^4 = \frac{1}{6}(4^2 - 1^2) = \frac{15}{6} = \frac{5}{2}$

(b) Temos: $\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2\, f(x)\, dx = \int_1^4 (x - \frac{5}{2})^2 \left(\frac{1}{3}\right) dx = \frac{1}{3}\int_1^4 (x - \frac{5}{2})^2\, dx = \frac{1}{3}\frac{\left(x - \frac{5}{2}\right)^3}{3}\Bigg|_1^4 = \frac{3}{4}$

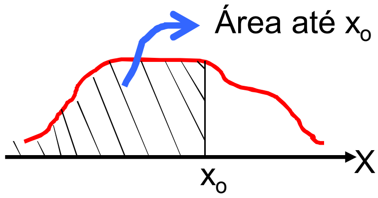
e assim: $\sigma = \sqrt{\sigma^2} = \frac{\sqrt{3}}{2}$

Função Distribuição Acumulada

A função distribuição acumulada (fda) F(x) é a probabilidade de que $X \le x_o$, ou seja,

$$F(x_o) = P(X \le x_o) = \int_{-\infty}^{x_o} f(x)\,dx$$

F(x) é análoga a distribuição de frequências relativas acumuladas (ou % acumuladas) estudadas no início do curso

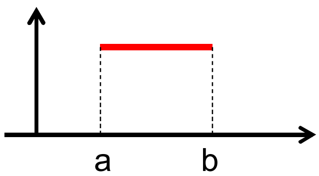


A Distribuição Uniforme

Para a Distribuição Uniforme a fdp é dada por:

$$f(X) = \begin{cases} \dfrac{1}{b-a} & \text{se} \quad a \le X \le b \\ 0 & a < X \ \text{ou} \ X > b \end{cases}$$

gráfico



média: $\mu = \dfrac{a+b}{2}$

desvio padrão: $\sigma = \sqrt{\dfrac{(b-a)^2}{12}}$

A Distribuição Exponencial

- Distribuição contínua que é assimétrica à direita e se estende de zero até o infinito positivo
- Amplamente utilizada na teoria das filas para modelar intervalo do tempo decorrido entre chegadas em processos, tais como:
  - clientes em caixas eletrônicos de bancos
  - pacientes dando entrada em uma emergência de um hospital
  - pesquisas em um endereço da internet

fdp para a distribuição exponencial:

$$f(x) = \begin{cases} \dfrac{\mathbf{e}^{-\frac{x}{\beta}}}{\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

média: $\mu = \beta$

variância: $\sigma^2 = \beta^2$

desvio padrão: $\sigma = \beta$

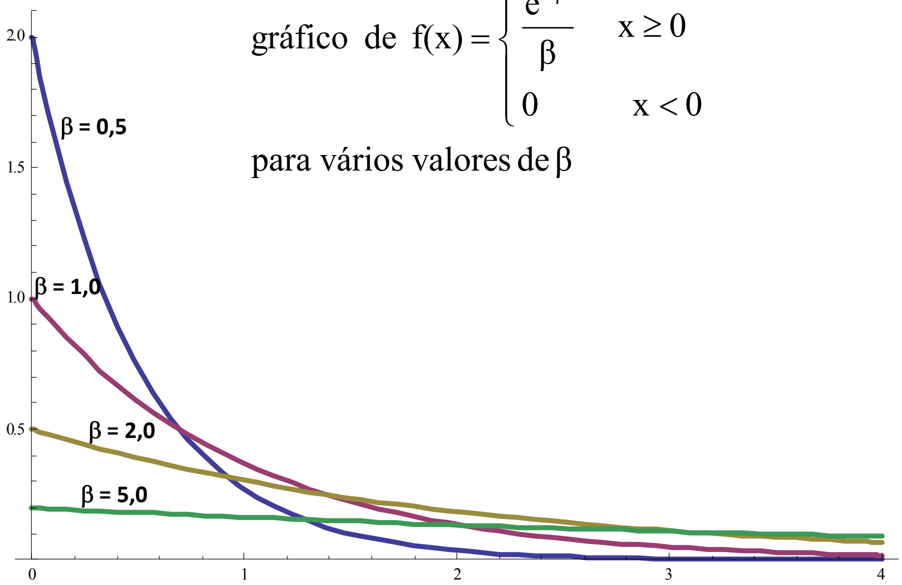
A probabilidade acumulada de zero até $X = a$: $P(x \leq a) = 1 - e^{-\frac{a}{\beta}}$

é área abaixo da curva desde 0 até a.

Note que β representa o valor médio de x e $1/\beta$ representa uma frequência média. Em alguns livros usa-se $\lambda = 1/\beta$.

$$\text{gráfico de } f(x) = \begin{cases} \dfrac{e^{-\frac{x}{\beta}}}{\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

para vários valores de β



Exemplo: Cerca de 15 clientes utilizam o caixa eletrônico por hora. Supondo que a distribuição de tempos de chegada é exponencial, qual é a probabilidade de que o tempo de chegada entre clientes consecutivos seja:

(a) Menor que três minutos?

(b) Maior que 3 minutos?

(c) Entre 2 e 4 minutos?

- A variável aleatória X= tempo

- Note que foi dada a frequência de chegada $\lambda = 1/\beta = 15$ /hora

-3 min = 0,05 horas

-2 min = 0,0333 horas

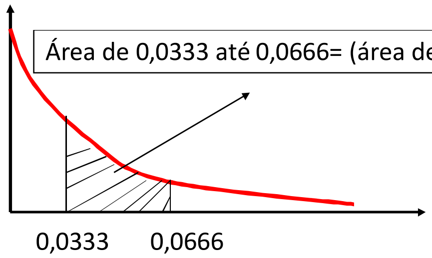
-4 min = 0,0666 horas

(a) $P(t < 0{,}05) = 1 - e^{-\lambda t} = 1 - e^{-(15).(0{,}05)} = 0{,}5276$

(b) $P(t \geq 0{,}05) = 1 - P(t < 0{,}05) = 1 - 0{,}5276 = 0{,}4734$

(c)

Área de 0,0333 até 0,0666= (área de 0 até 0,0666) – (área de 0 até 0,033)



$P(0{,}0333<t < 0{,}0666)= P(t<0{,}0666)- P(t<0{,}0333)= e^{-(15).(0{,}0666))} - e^{-(15).(0{,}0333))}$

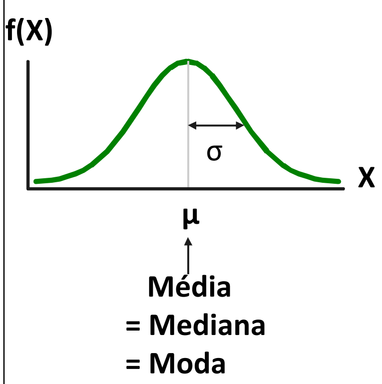
Distribuição Normal

- curva em "**Forma de sino**"
- **Simétrica**
- **Média, Mediana e Moda são iguais**

A localização é determinada pela média, μ

A dispersão é determinado pelo desvio-padrão, σ

A variável aleatória tem uma amplitude teórica infinita:
$+\infty$ **até** $-\infty$



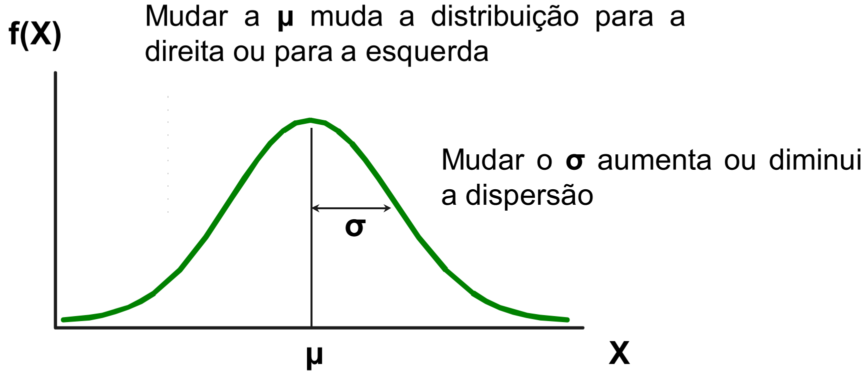
A fórmula para a fdp da distribuição normal é:

$$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

Notação:

$N(\mu,\sigma^2)$ distribuição normal com média μ e variância σ^2

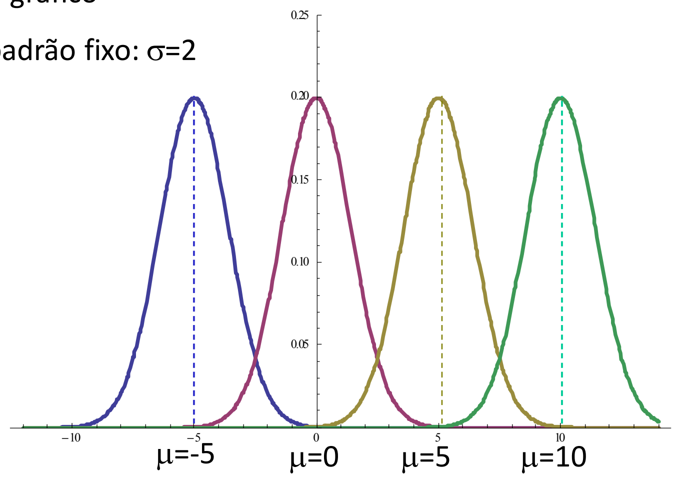
Variando os parâmetros μ e σ, obtém-se diferentes distribuições normais



Mudar a média **μ** desloca a distribuição para a direita ou para a esquerda

Exemplo gráfico

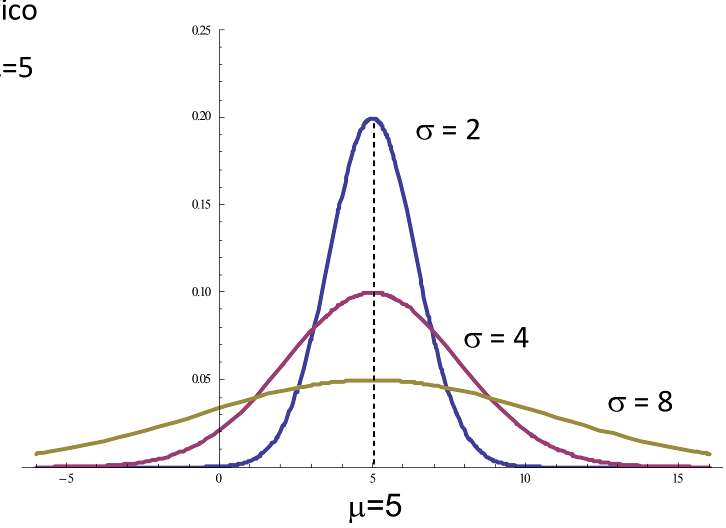
Desvio padrão fixo: $\sigma=2$



Mudar o desvio padrão σ aumenta ou diminui a dispersão

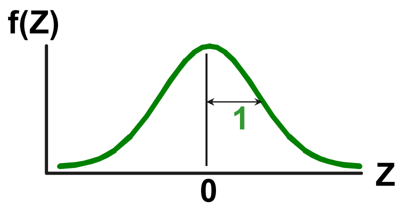
Exemplo gráfico

Média fixa: $\mu=5$



Distribuição Normal Padronizada(N(0,1)): distribuição normal com média zero e variância um.

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



Valores acima da média possuem valor Z positivo

Valores abaixo da média possuem valor Z negativo

Transformação para a Distribuição Normal Padronizada

Transformação de X em normal padronizada (distribuição "Z") mudando a variável

$$Z = \frac{X - \mu}{\sigma}$$

A distribuição Z sempre tem média zero e desvio-padrão 1

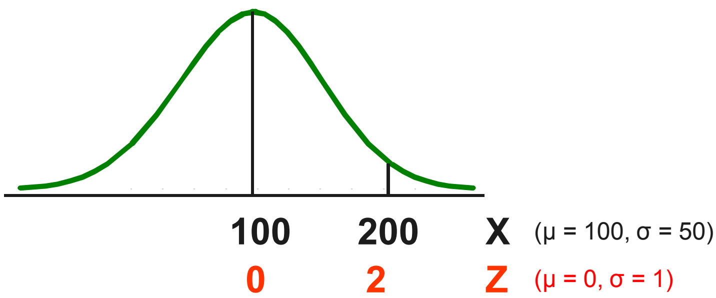
Exemplo

- Se X é normalmente distribuído com média 100 e desvio padrão 50, o valor Z para X = 200 é

$$Z = \frac{X - \mu}{\sigma} = \frac{200 - 100}{50} = 2$$

- X = 200 está dois desvios-padrão acima da média

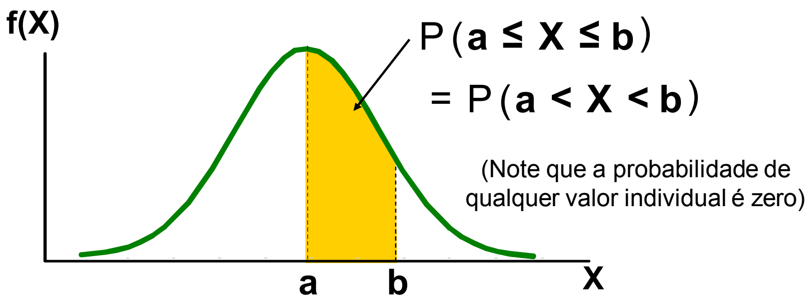
Comparando X e Z



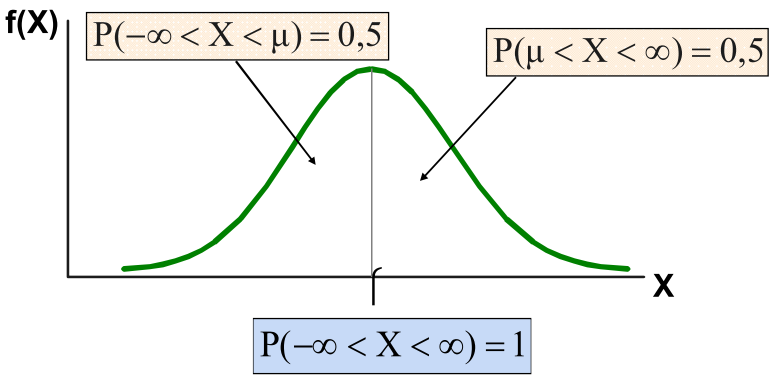
Note que a distribuição é a mesma, apenas a escala mudou. O problema pode ser expresso em unidades originais (X) ou em unidades padronizadas (Z)

Encontrando Probabilidades com a distribuição normal

A **probabilidade** é dada pela área abaixo da curva

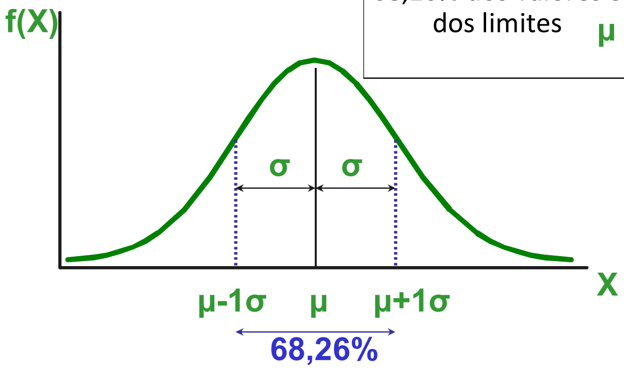


A **área total abaixo da curva é igual a 1** e a curva é simétrica, então metade da curva está acima da média e a outra metade está abaixo



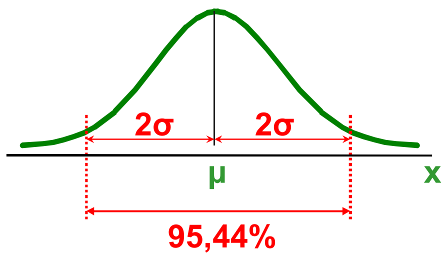
Regras Gerais:

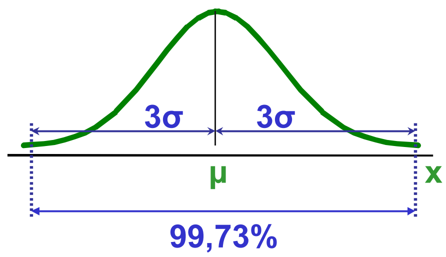
68,26% dos valores estão dentro dos limites $\mu \pm 1\sigma$



95,44 % dos valores estão dentro dos limites $\mu \pm 2\sigma$

99,73% dos valores estão dentro dos limites $\mu \pm 3\sigma$





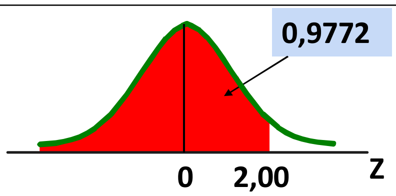
A Tabela da Normal Padronizada

- A Tabela da Normal Padronizada fornece a probabilidade menor do que um valor desejado para Z (isto é, do infinito negativo até Z)

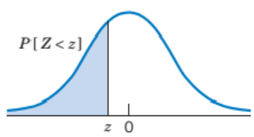
Exemplo:

P(Z < 2,00) = 0,9772

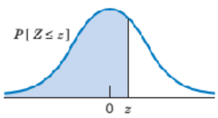
Esse valor é encontrado na tabela



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998



Encontrando o valor na tabela

A **coluna** fornece o valor da segunda casa decimal de Z

Z	0,00	0,01	0,02 ...
0,0			
0,1			
,			
2,0	0,9772		

A **linha** fornece o valor da primeira casa decimal de Z

O valor dentro da tabela representa a probabilidade

P(Z < 2,00) = 0,9772

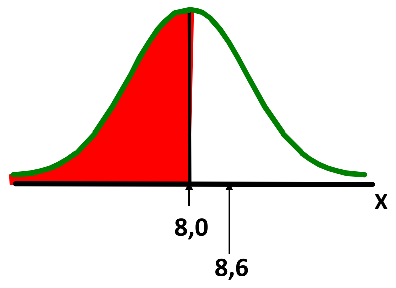
Procedimentos Gerais para Encontrar Probabilidades

Para encontrar P(a < X < b) quando X é normalmente distribuído:

- Desenhe a curva normal para o problema em termos de X
- Transforme os valores X em valores Z
- Use a Tabela da Normal Padronizada

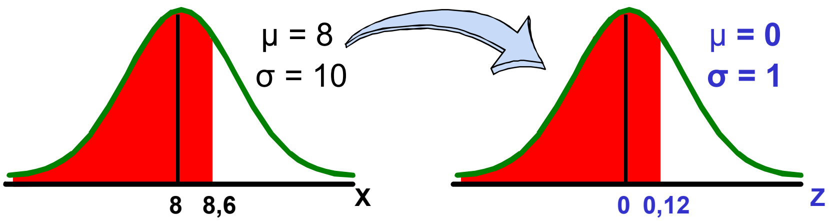
Encontrando Probabilidades Normais

- Suponha que X é normal com média igual a 8 e desvio-padrão igual a 5
- Encontre P(X < 8,6)



Encontrando o valor padronizado

$$Z = \frac{X - \mu}{\sigma} = \frac{8{,}6 - 8}{5} = 0{,}12$$



P(X < 8,6)

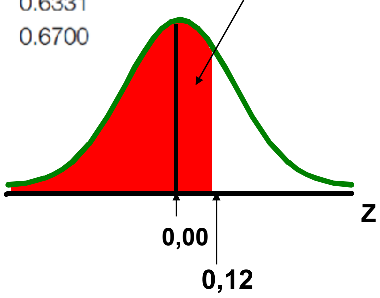
P(Z < 0,12)

Procurando P(Z < 0,12) na tabela

Standard Normal Probabilities

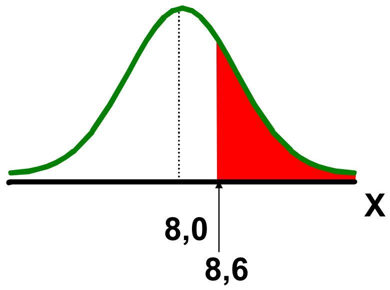
x	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700

P(X < 8,6)= P(Z < 0,12)=
0,5478



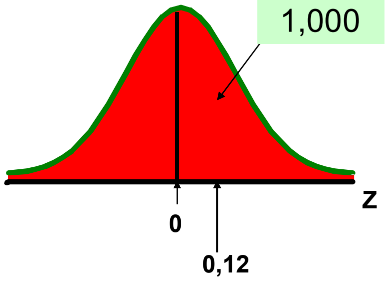
Probabilidades Acima da Média

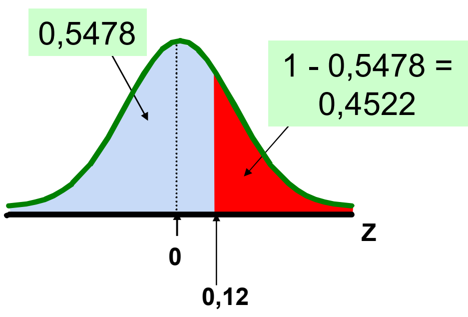
- Suponha que X é normal com média igual a 8 e desvio-padrão igual a 5
- Encontre P(X > 8,6)



P(X > 8,6) = P(Z > 0,12) = 1 - P(Z ≤ 0,12)

= 1 - 0,5478 = 0,4522





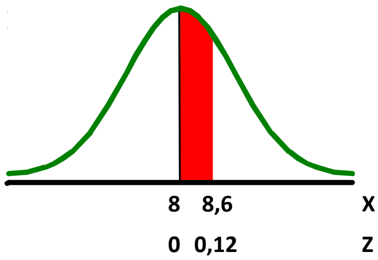
Probabilidades entre Dois Valores

- Suponha que X é normal com média igual a 8 e desvio-padrão igual a 5. Encontre P(8 < X < 8,6).

Cálculo do valor Z:

$$Z = \frac{X - \mu}{\sigma} = \frac{8 - 8}{5} = 0$$

$$Z = \frac{X - \mu}{\sigma} = \frac{8,6 - 8}{5} = 0,12$$



P(8 < X < 8,6)

= P(0 < Z < 0,12)

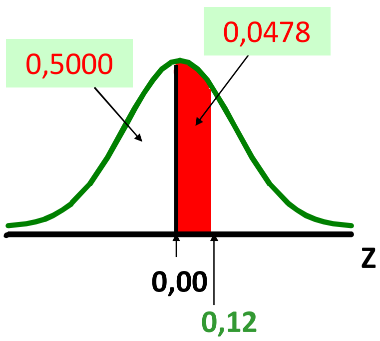
P(8 < X < 8,6) = P(0 < Z < 0,12)

= P(Z < 0,12) – P(Z ≤ 0)

= 0,5478 - 0,5000 = 0,0478

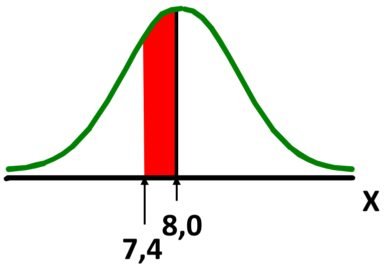
Tabela da Probabilidade Normal Padronizada

Z	0.00	0.01	0.02	0.03
0.0	0.5000	0.5040	0.5080	0.5120
0.1	0.5398	0.5438	0.5478	0.5517
0.2	0.5793	0.5832	0.5871	0.5910
0.3	0.6179	0.6217	0.6255	0.6293



Probabilidades Abaixo da Média

- Suponha que X é normal com média igual a 8 e desvio-padrão igual a 5.
- Encontre P(7,4 < X < 8)



Calculando os valores Z:

$$Z = \frac{X - \mu}{\sigma} = \frac{8 - 8}{5} = 0$$

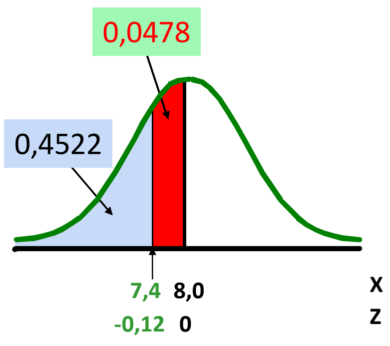
$$Z = \frac{X - \mu}{\sigma} = \frac{7{,}4 - 8}{5} = -0{,}12$$

Temos:

P(7,4 < X < 8) = P(-0,12 < Z < 0)

= P(Z < 0) – P(Z ≤ -0,12)

= 0,5000 - 0,4522 = **0,0478**



Lembre-se: A distribuição normal é simétrica e a probabilidade é igual a P(0 < Z < 0,12)

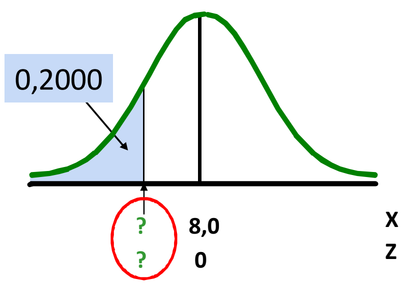
Encontrando o valor X para uma probabilidade conhecida

- Passos para encontrar o valor X para uma probabilidade conhecida:
  1. Encontre o valor Z para a probabilidade conhecida
  2. Converta o valor Z em X pela fórmula:

$$\boxed{X = \mu + Z\sigma}$$

Exemplo:

- Suponha que X tem distribuição normal com média igual a 8 e desvio-padrão igual a 5
- Encontre o valor X, considerando que apenas 20% dos valores são menores que esse valor X

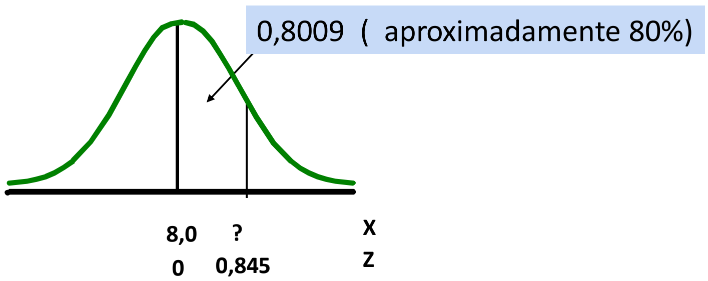


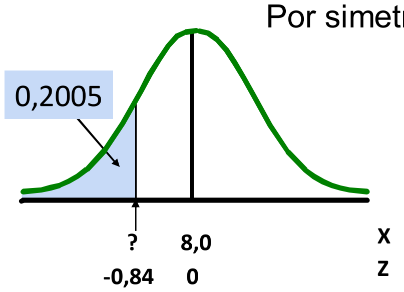
Lembre-se a distribuição é simétrica

Tabela da Probabilidade Normal Padronizada: valores positivos

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365

Como o valor 80% está entre dois valores, utilizamos o ponto médio: Temos aproximadamente P = 0,8009 dos valores ocorrem para x<0,845.





Por simetria, valores abaixo da média:

Z = -0,845

2. Converta o valor Z em X pela fórmula:

$$X = \mu + Z\sigma$$
$$= 8 + (-0,845).5$$
$$= 3,775$$

Portanto: 20% dos valores da distribuição com média igual a 8 e desvio-padrão igual a 5 são menores que 3,775

Exemplo:
A média das ações de empresas que compoem as S&P500 é de US$ 30 e o desvio padrão é de US$8,20. Supoha que os preços das ações se distribuam normalmente.

(a) qual a probabilidade de uma empresa ter um preço de, no mínimo, US$ 40 para suas ações?

(b) Qual a probabilidade de uma empresa ter um preço não superior a US$ 20 para suas ações?

(c) qual o preço das ações para que a empresa seja incluída nas 10% maiores?

(a) para $x = 40: \ z = 1{,}22$

$P(x < 40) = P(z < 1{,}22) = 0{,}8888$

portanto $P(x \geq 40) = 1 - P(x < 40) = 0{,}1112$

(b) para $x = 20: \ z = -1{,}22$

$P(x > 20) = P(z > -1{,}22) = P(z < 1{,}22) = 0{,}8888$

portanto $P(x \leq 20) = 1 - 0{,}8888 = 0{,}1112$

(**c**) área de 10% na cauda superior corresponde $z = 1{,}285$

(tabela e utilizando o ponto médio)

então: $x = 30 + 8{,}2.(1{,}285) = 40{,}50$

a empresa estara entre as 10% maiores se tiver um preço $\geq 40{,}50$

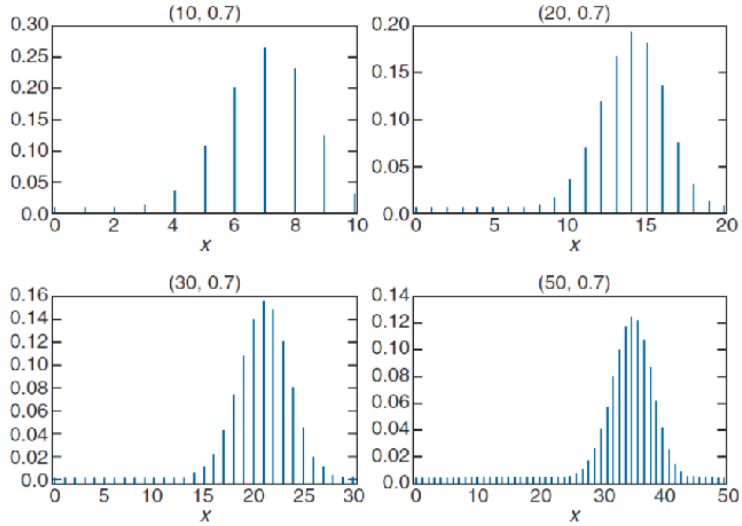
Aproximação da Normal para a Distribuição Binomial

- Quanto mais perto p estiver de 0,5, melhor será a aproximação da normal para a distribuição binomial
- Quanto maior o tamanho da amostra, n, melhor será a aproximação da normal para a distribuição binomial

- Regra Geral:
  - A distribuição normal pode ser utilizada para aproximar a distribuição binomial se

$$np \geq 5 \text{ e } n(1 - p) \geq 5$$

Gráficos da distribuição binomial para p=0,7: note que a medida que n aumenta o gráfico fica mais próximo ao de uma distribuição normal.



- A média e o desvio-padrão da distribuição binomial são

$$\mu = n.p$$

- Transforme a binomial em normal utilizando a fórmula:

$$\sigma = \sqrt{np(1-p)}$$

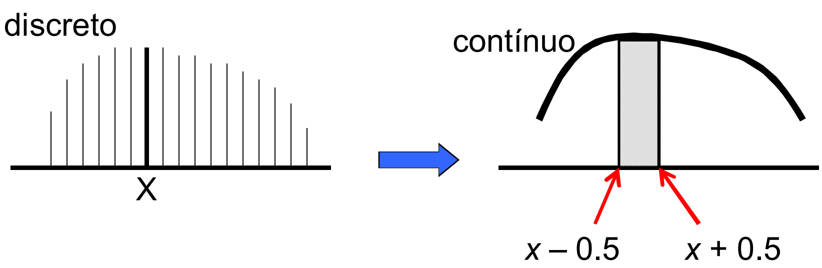
$$\boxed{Z = \frac{X-\mu}{\sigma} = \frac{X-np}{\sqrt{np(1-p)}}}$$

Correção de continuidade

Note que:

A distribuição Normal é uma distribuição contínua, enquanto que a distribuição Binomial é uma distribuição discreta.

Assim, para aproximar a distribuição Binomial (que é discreta) pela normal (que é contínua) fazemos uma **correção de continuidade** ao valor discreto x na distribuição binomial representando o valor x pelo intervalo de $x - 0.5$ a $x + 0.5$.

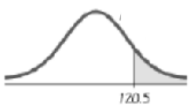


Por exemplo:

X = <u>**pelo menos**</u> **120**
= 120, 121, 122, . . .

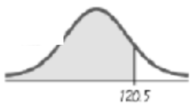
$X \geq 120 \rightarrow X \geq 119{,}5$

X = <u>**mais do que**</u> **120**
= 121, 122, 123, . . .



$X > 120 \rightarrow X > 120{,}5$

X = <u>**no máximo**</u> **120**
= 0, 1, . . . 118, 119, 120

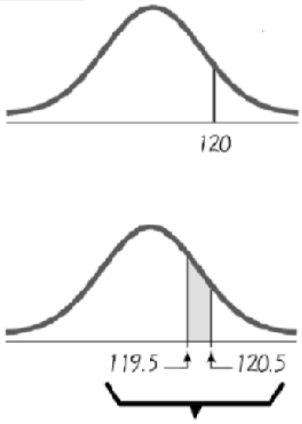


$X \leq 120 \rightarrow X \leq 120{,}5$

X = <u>**menos do que**</u> **120**
= 0, 1, . . . 118, 119

$X < 120 \rightarrow X < 119{,}5$

***x* = exactamente 120**



Intervalo que representa o valor discreto 120

Exemplo: Considere a distribuição binomial com n=1000 e p=0,02. Qual é $P(X \leq 180)$?

Temos: np=1000.0,02=20>5 e n(1-p)=1000.0,98=980>5

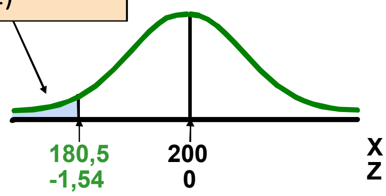
Portanto, podemos aproximar a distribuição binomial pela normal.

Correção de continuidade: Aproximar $P(X \leq 180)$ para $P(X \leq 180{,}5)$

Transformar a binomial em normal padronizada:

$$Z = \frac{X - n.p}{\sqrt{n.p.(1-p)}} = \frac{180{,}5 - (1000).(0{,}2)}{\sqrt{(1000).(0{,}2).(1-0{,}2)}} = -1{,}54$$

Consultando a tabela:
$P(Z \leq -1{,}54) = 0{,}0618$



Algumas propriedades

Sejam X e Y variáveis aleatórias independentes com distribuições normais de probabilidade, com médias μ_x e μ_y e desvios-padrão σ_x e σ_y, respectivamente. Então a variável aleatória X+Y tem distribuição normal com média

$$\mu_{x+y} = \mu_x + \mu_y$$

e desvio-padrão

$$\sigma_{x+y} = \sqrt{\sigma_x^2 + \sigma_y^2}$$

Verificando normalidade

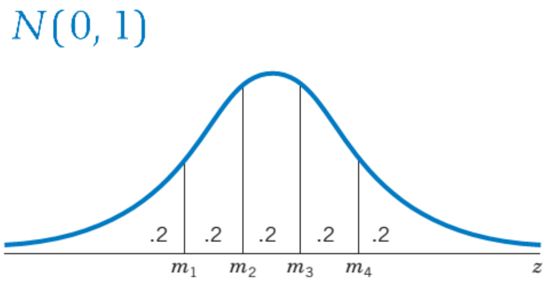
Muitos procedimentos estatísticos requerem que a população tenha uma distribuição próxima da distribuição normal.

Espera-se que a amostra reflita a distribuição da população: histograma parecido, distribuição acumulada...

Outra maneira de verificar se normalidade é razoável:
normal-score plot

Exemplo: considere os dados: 68, 42, 44, 75
Para n=4, temos 5 faixas com probabilidade 1/5=0.2.
Cada valor dos dados ordenados acumula 1/5=0.2

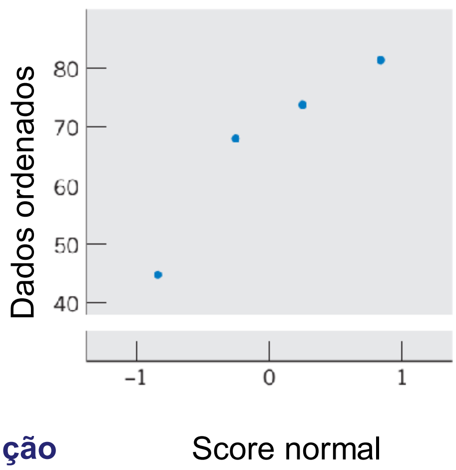
E corresponderiam, se fossem de uma distribuição normal, aos pontos:



$$m_1 = -.84$$
$$m_2 = -.25$$
$$m_3 = .25$$
$$m_4 = .84$$

normal-score plot

Score normal	Dados ordenados
$m_1 = -.84$	44
$m_2 = -.25$	68
$m_3 = .25$	75
$m_4 = .84$	82

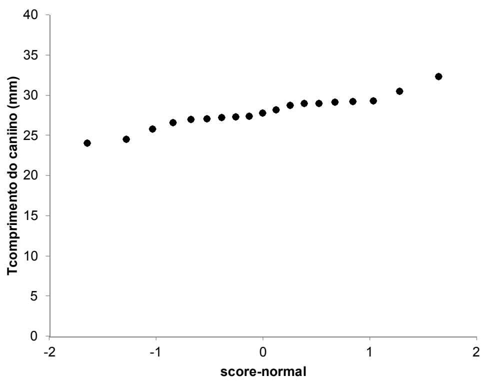


Um padrão retilineo indica proximidade com a distribuição normal.

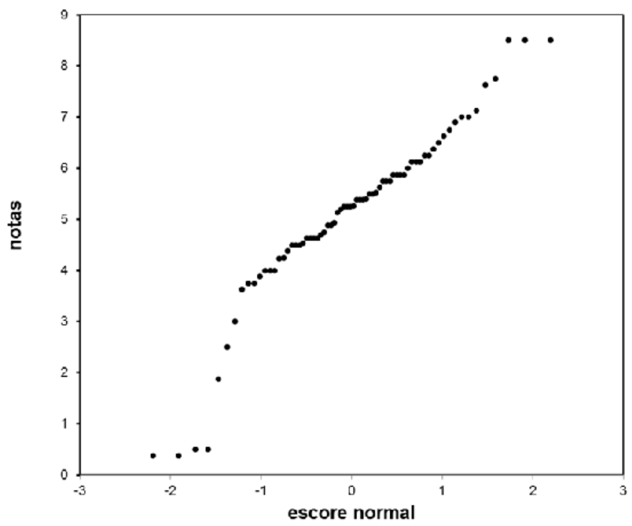
Dados tabela D9 apendice do livro texto

	A	B	C
1		score normal	comprimento do canino de lobos
2	1	=INV.NORMP.N(A2/20)	24
3	2	INV.NORMP.N(probabilidade)	24.5
4	3	-1.036433389	25.8
5	4	-0.841621234	26.6
6	5	-0.67448975	27
7	6	-0.524400513	27.1
8	7	-0.385320466	27.2
9	8	-0.253347103	27.3
10	9	-0.125661347	27.4
11	10	0	27.8
12	11	0.125661347	28.2
13	12	0.253347103	28.7
14	13	0.385320466	29
15	14	0.524400513	29
16	15	0.67448975	29.1
17	16	0.841621234	29.2
18	17	1.036433389	29.3
19	18	1.281551566	30.5
20	19	1.644853627	32.3

	A	B	C
1		score normal	canino (mm)
2	1	-1.644853627	24
3	2	-1.281551566	24.5
4	3	-1.036433389	25.8
5	4	-0.841621234	26.6
6	5	-0.67448975	27
7	6	-0.524400513	27.1
8	7	-0.385320466	27.2
9	8	-0.253347103	27.3
10	9	-0.125661347	27.4
11	10	0	27.8
12	11	0.125661347	28.2
13	12	0.253347103	28.7
14	13	0.385320466	29
15	14	0.524400513	29
16	15	0.67448975	29.1
17	16	0.841621234	29.2
18	17	1.036433389	29.3
19	18	1.281551566	30.5
20	19	1.644853627	32.3



Normal-score plot para notas de uma disciplina



Transformando dados para obter normalidade:

Some Useful Transformations

Make large values larger:	Make large values smaller:
x^3, x^2	$\sqrt{x}$, $\sqrt[4]{x}$, $\log_e x$, $\frac{1}{x}$