DISTRIBUIÇÕES TEÓRICAS DE PROBABILIDADE DE VARIÁVEIS ALEATÓRIAS CONTÍNUAS

INTRODUÇÃO

Nas aulas anteriores tratamos das distribuições de variáveis discretas, isto é, daqueles em que o evento só pode ser expresso em números inteiros. Voltemo-nos agora para as distribuições de variáveis contínuas. Estas, ao contrário das discretas, podem assumir qualquer valor matemático, mesmo fracionário. Podemos citar como exemplos deste tipo de variáveis os salários, as medidas físicas, biológicas, psicológicas, etc. Como, para este tipo de variável, não se pode enumerar todos os valores possíveis e suas respectivas probabilidades, convém elaborar uma função de densidade de probabilidade – ou curva de probabilidade.

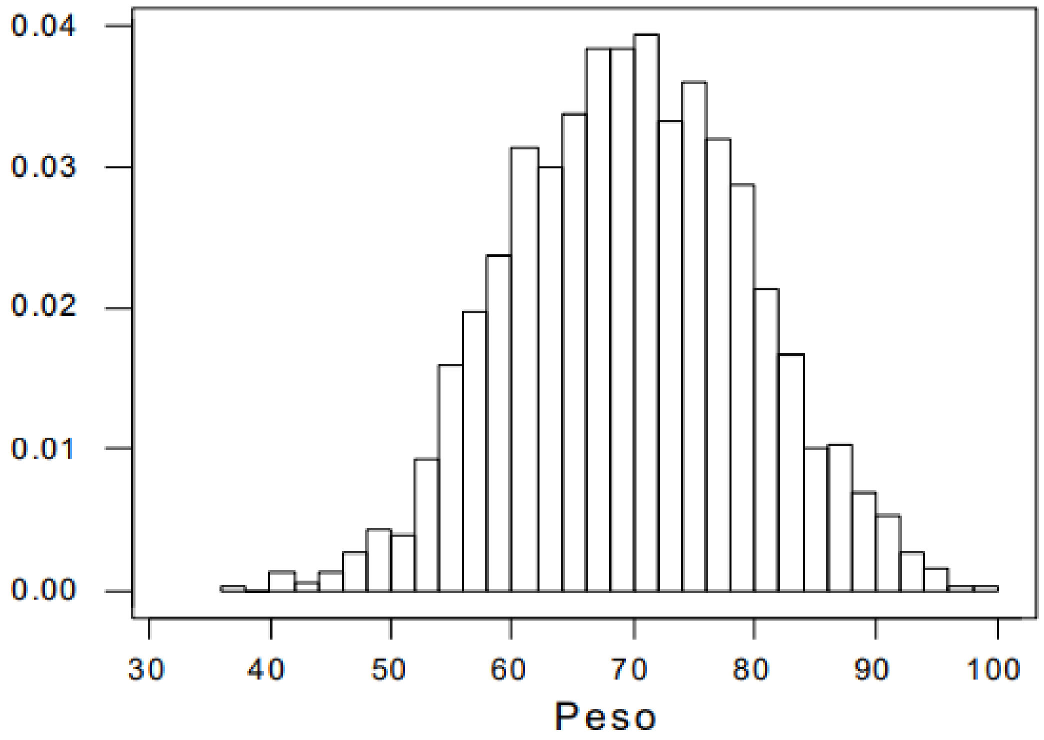
Distribuição Normal, t de Student, Qui-Quadrado e F de Snedecor

DISTRIBUIÇÃO NORMAL

Dentre as distribuições de probabilidades referentes a variáveis contínuas a mais comum é a distribuição normal.

Observamos o peso, em kg, de 1500 pessoas adultas selecionadas ao acaso em uma população.

O histograma por densidade é o seguinte:



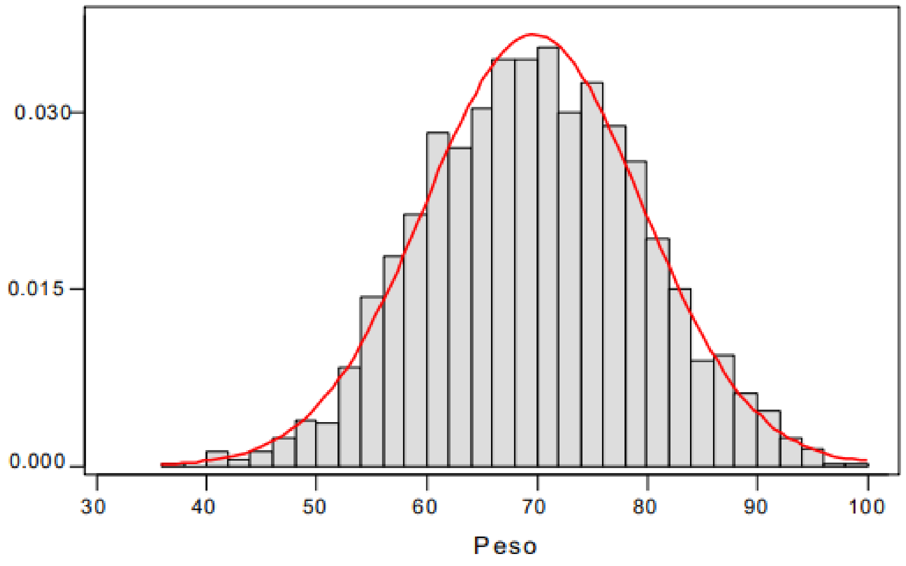
DISTRIBUIÇÃO NORMAL

Vamos definir a variável aleatória

X: peso, em kg, de uma pessoa adulta escolhida ao acaso da população

Como se distribuem os valores da variável aleatória X, isto é, qual a distribuição de probabilidades de X?

A curva contínua da figura denomina-se curva Normal.



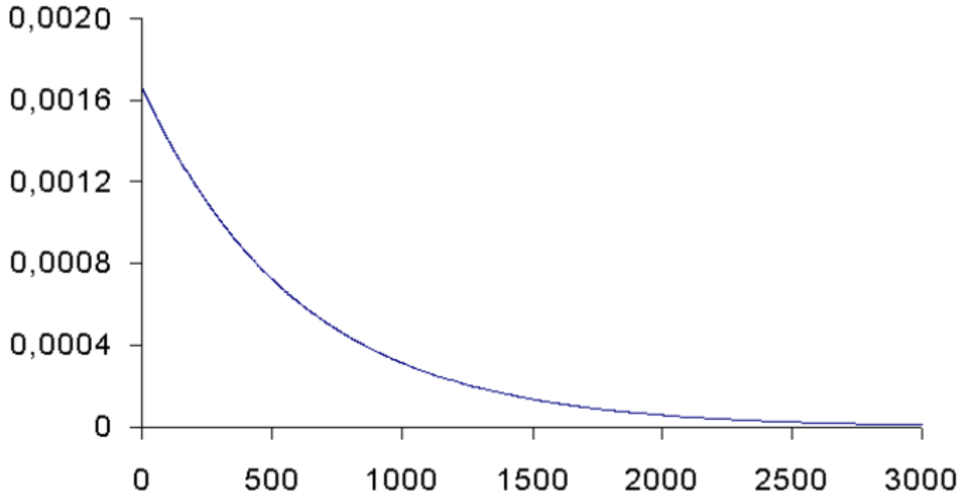
DISTRIBUIÇÃO NORMAL

Nem todos os fenômenos se ajustam à distribuição Normal.

Exemplo

Y: Duração, em horas, de uma lâmpada de certa marca.

A experiência sugere que esta distribuição deve ser assimétrica – grande proporção de valores entre 0 e 500 horas e pequena proporção de valores acima de 1500 horas.



DISTRIBUIÇÃO NORMAL

Características da curva normal

1º) A variável aleatória X pode assumir qualquer valor real;

2º) A apresentação gráfica da distribuição normal corresponde a uma curva em forma de sino;

3º) A distribuição normal é simétrica em torno da média; assim, se considerarmos as ordenadas referentes a dois valores de X com afastamento idênticos em relação à média, positiva e negativamente, obteremos a mesma ordenada e o mesmo valor de probabilidade;

4º) A curva normal admite uma única ordenada máxima (pico), situada na média, e assim as medidas de tendência central - $\bar{X}$, Md e Mo – são idênticas;

DISTRIBUIÇÃO NORMAL

Características da curva normal

5º) Os valores concentram-se em torno da média e, à medida que se afastam desta, qualquer que seja o seu sentido, tendem a se tornar mais raros;

6º) Ao se afastarem os valores de X em relação à média, quer positiva ou negativamente, a curva não só modifica sua tendência em direção ao eixo das abcissas, como não chega a tocá-lo; por isso, dizemos que a curva normal varia de $-\infty$ a $+\infty$. Em outras palavras, podemos afirmar que a distribuição normal em sua representação gráfica – curva normal – apresenta sempre dois pontos de inflexão e é assintótica em relação ao eixo da variável X;

7º) A área total sob a curva corresponde à proporção 1 ou à porcentagem 100%;

8º) A probabilidade de ocorrer valor maior ou menor que a média é equivalente, sendo igual a 0,50 ou 50%.

DISTRIBUIÇÃO NORMAL

A variável aleatória X tem distribuição com parâmetros μ e σ^2 se sua função densidade de probabilidade é dada por

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

Pode ser mostrado que

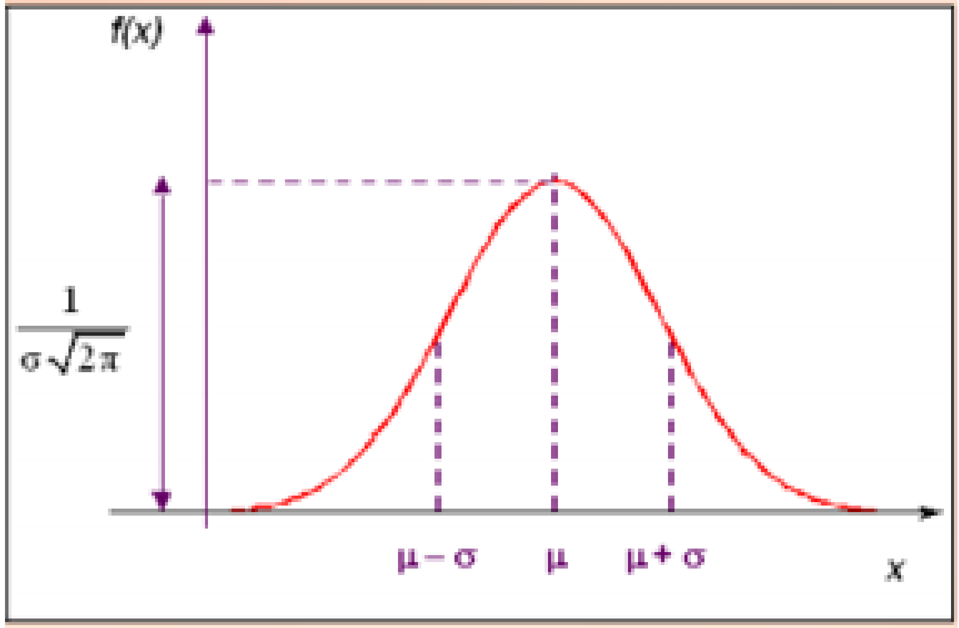
1. μ é o valor esperado (média) de X $(-\infty < \mu < \infty)$;

2. σ^2 é a variância de X $(\sigma^2 > 0)$.

Notação: X~N(μ; σ^2)

DISTRIBUIÇÃO NORMAL

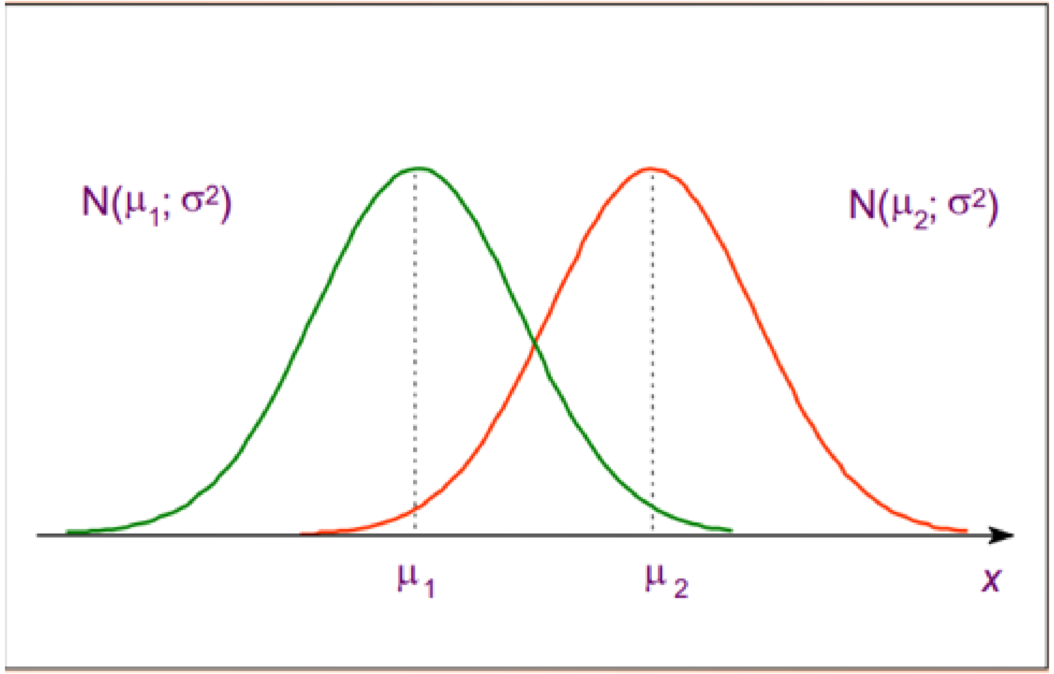
Propriedades de X~N(μ; σ^2)



- E(X) = μ (média ou valor esperado);
- Var(X) = σ^2 (e portanto, DP(X) = σ);
- $f(x) \to 0$ quando $x \to \pm\infty$;
- $x = \mu$ é o ponto máximo de $f(x)$;
- $\mu - \sigma$ e $\mu + \sigma$ são pontos de inflexão de $f(x)$;
- a curva Normal é simétrica em torno da média μ.

DISTRIBUIÇÃO NORMAL

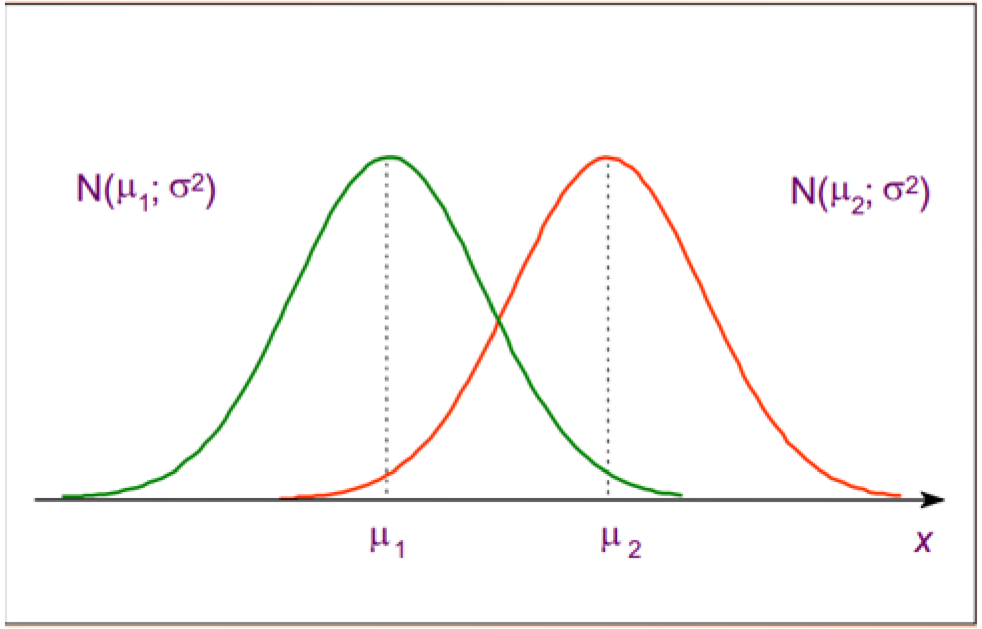
A distribuição normal depende dos parâmetros μ e σ^2



Curvas normais com mesma variância σ^2, mas médias diferentes $(\mu_2 > \mu_1)$

DISTRIBUIÇÃO NORMAL

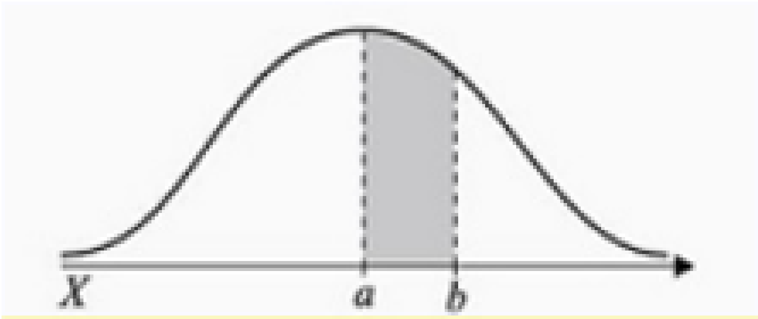
Influência de σ^2 na curva normal



Curvas normais com mesma média μ, mas com variâncias diferentes ${\sigma_2}^2 > {\sigma_1}^2$.

DISTRIBUIÇÃO NORMAL

Agora observe a figura:



Se quisermos calcular a probabilidade indicada na figura, devemos fazer:

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

que apresenta um grau relativo de dificuldade.

DISTRIBUIÇÃO NORMAL

Seja X: (μ, σ^2), definimos:

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

Z também tem distribuição normal. Z é chamada de variável normal reduzida, normal padronizada ou variável normalizada e E(X) = 0 e Var(X) = 1.

Logo, se:

$$X: (\mu, \sigma^2) \Rightarrow Z: N(0, 1)$$

A função densidade de probabilidade de Z é $f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$ para $-\infty < Z < +\infty$.

Essa curva é também simétrica com relação a μ_Z.

DISTRIBUIÇÃO NORMAL

Agora vamos verificar a correspondência entre X e Z, por meio do seguinte exemplo:

Seja X: N(20, 4). Encontrar os valores reduzidos correspondentes a:

$X_1 = 14$

$X_2 = 16$

$X_3 = 18$

$X_4 = 20$

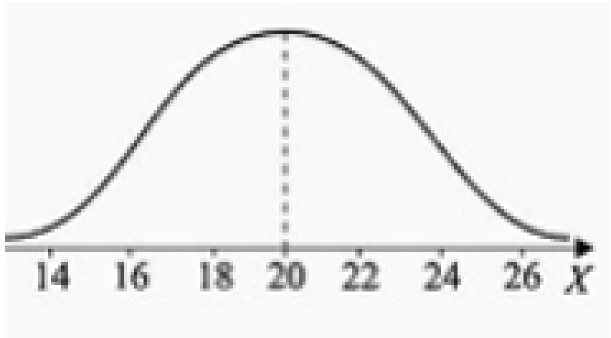
$X_5 = 22$

$X_6 = 24$

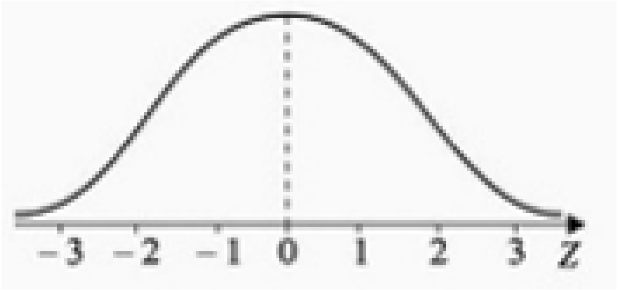
$X_7 = 26$

DISTRIBUIÇÃO NORMAL

Graficamente:



$$\begin{cases} 14 = \mu_x - 3\sigma \\ 16 = \mu_x - 2\sigma \\ 18 = \mu_x - \sigma \\ 20 = \mu_x \\ 22 = \mu_x + \sigma \\ 24 = \mu_x + 2\sigma \\ 26 = \mu_x + 3\sigma \end{cases}$$

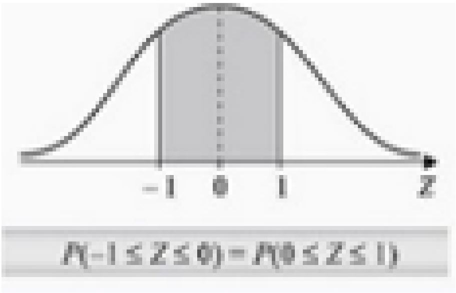


$$\begin{cases} -3 = \mu_z - 3\sigma \\ -2 = \mu_z - 2\sigma \\ -1 = \mu_z - \sigma \\ 0 = \mu_z \\ 1 = \mu_z + \sigma \\ 2 = \mu_z + 2\sigma \\ 3 = \mu_z + 3\sigma \end{cases}$$

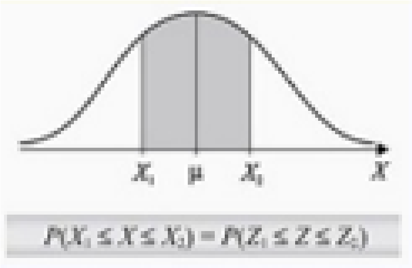
Concluímos que a variável Z indica quantos desvios padrões a variável X está afastada da média.

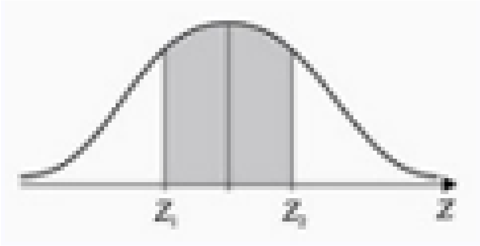
DISTRIBUIÇÃO NORMAL

Como as curvas são simétricas em relação às médias:



Concluímos então que se X: N(μ, σ^2):





DISTRIBUIÇÃO NORMAL

Pois

$$P(X_1 \leq X \leq X_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$P(Z_1 \leq Z \leq Z_2) = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz$$

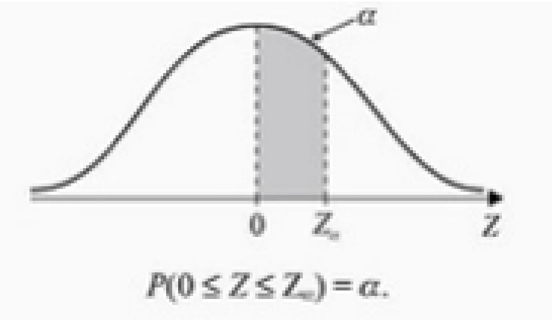
onde

$Z_1 = \frac{X_1-\mu}{\sigma}$ e $Z_2 = \frac{X_2-\mu}{\sigma}$

DISTRIBUIÇÃO NORMAL

Uso da tabela

A vantagem de se usar a variável $Z = \frac{X-\mu}{\sigma}$ é que podemos tabelar os valores da área, ou as probabilidades, pois para cada X dado, a área depende de μ e σ^2. Como $\mu_Z = 0$ e $\sigma^2 = 1$, uma tabela de Z é suficiente.



DISTRIBUIÇÃO NORMAL

Exemplos

1) Seja $X: N(100, 25)$, calcular:

$a) P(100 \leq X \leq 106)$

$b) P(89 \leq X \leq 107)$

$c) P(112 \leq X \leq 116)$

$d) P(X \geq 108)$

DISTRIBUIÇÃO NORMAL

2) Os 60 alunos de uma turma de faculdade responderam ao Teste de Inteligência Cia, sendo o conjunto de notas resultantes aproximadamente normal, com média 100 e desvio padrão 15. Pede-se:

a) Que proporção de QIs supera 130?

b) Qual a porcentagem de alunos com QIs entre 85 e 115?

c) Na amostra em questão (60 alunos), quantos teriam QI entre 70 e 120?

d) Qual o QI corresponde aos 10% dos alunos com resultados mais baixos?

e) Qual o QI corresponde aos 5% dos extremos?

DISTRIBUIÇÃO NORMAL

3) Suponhamos que o tempo necessário para terminar uma prova escolar tenha distribuição aproximadamente normal, com média de 90 minutos e desvio padrão de 15 minutos.

a) Que proporção de estudantes terminará a tarefa em, no máximo, uma hora (60 minutos)?

b) Qual o tempo necessário para permitir que 80% dos estudantes terminem a prova?

DISTRIBUIÇÃO NORMAL

4) A taxa normal de glicose no sangue humano é uma variável aleatória com distribuição aproximadamente normal, média = 100mg/100ml de sangue e desvio padrão = 20mg/100ml de sangue. Pergunta-se:

a) Qual a probabilidade de um indivíduo apresentar taxa de glicose superior a 135mg/100ml de sangue?

b) Qual a porcentagem de indivíduos com taxas de glicose entre 90 e 110mg/100 ml de sangue?