vel gerar probabilidades e
es discutidos neste capítu-
rar probabilidades e CDF

= x) e $P(X \leqslant x)$ para uma

ial 14 0.3.
Distribution Function
n = 14 and p = 0.300000

x)	x	P(X <= x)
3	6	0.9067
5	7	0.9685
3	8	0.9917
2	9	0.9983
2	10	0.9998
5	11	1.0000

abilidades e probabilidades
parâmetro $\lambda = 5,2$.

Quadro 6.2 Probabilidades de Poisson geradas pelo Minitab.

MTB > PDF;
SUBC> Poisson 5.2.
Probability Density Function
Poisson with mu = 5.20000

x	P(X = x)	x	P(X = x)
0	0.0055	9	0.0423
1	0.0287	10	0.0220
2	0.0746	11	0.0104
3	0.1293	12	0.0045
4	0.1681	13	0.0018
5	0.1748	14	0.0007
6	0.1515	15	0.0002
7	0.1125	16	0.0001
8	0.0731	17	0.0000

MTB > CDF;
SUBC> Poisson 5.2.
Cumulative Distribution Function
Poisson with mu = 5.20000

x	P(X <= x)	x	P(X <= x)
0	0.0055	9	0.9603
1	0.0342	10	0.9823
2	0.1088	11	0.9927
3	0.2381	12	0.9972
4	0.4061	13	0.9990
5	0.5809	14	0.9997
6	0.7324	15	0.9999
7	0.8449	16	1.0000
8	0.9181		

Na planilha Excel podem ser usadas funções específicas dentro da categoria Estatística. Por exemplo, para cálculos com a distribuição binomial, usar a função DISTRBINOM; para a distribuição de Poisson, usar a função POISSON.

6.10 Problemas e Complementos

29. Um florista faz estoque de uma flor de curta duração que lhe custa \$0,50 e que ele vende a \$1,50 no primeiro dia em que a flor está na loja. Toda flor que não é vendida nesse primeiro dia não serve mais e é jogada fora. Seja X a variável aleatória que denota o número de flores que os fregueses compram em um dia casualmente escolhido. O florista descobriu que a função de probabilidade de X é dada pela tabela abaixo.

x	0	1	2	3
$p(x)$	0,1	0,4	0,3	0,2

Quantas flores deveria o florista ter em estoque a fim de maximizar a média (valor esperado) do seu lucro?

30. As cinco primeiras repetições de um experimento custam \$10,00 cada. Todas as repetições subseqüentes custam \$5,00 cada. Suponha que o experimento seja repetido até que o primeiro sucesso ocorra. Se a probabilidade de sucesso de uma repetição é igual a 0,9, e se as repetições são independentes, qual é o custo esperado da operação?

31. Na manufatura de certo artigo, é sabido que um entre dez dos artigos é defeituoso. Qual a probabilidade de que uma amostra casual de tamanho quatro contenha:
(a) nenhum defeituoso?
(b) exatamente um defeituoso?
(c) exatamente dois defeituosos?
(d) não mais do que dois defeituosos?

32. Um fabricante de peças de automóveis garante que uma caixa de suas peças conterá, no máximo, duas defeituosas. Se a caixa contém 18 peças, e a experiência tem demonstrado que esse processo de fabricação produz 5% das peças defeituosas, qual a probabilidade de que uma caixa satisfaça a garantia?

33. Um curso de treinamento aumenta a produtividade de uma certa população de funcionários em 80% dos casos. Se dez funcionários quaisquer participam desse curso, encontre a probabilidade de:
   (*a*) exatamente sete funcionários aumentarem a produtividade;
   (*b*) não mais do que oito funcionários aumentarem a produtividade; e
   (*c*) pelo menos três funcionários não aumentarem a produtividade.

34. O número de petroleiros que chegam a uma refinaria em cada dia ocorre segundo uma distribuição de Poisson, com $\lambda = 2$. As atuais instalações podem atender, no máximo, a três petroleiros por dia. Se mais de três aportarem num dia, o excesso é enviado a outro porto.
   (*a*) Em um dia, qual a probabilidade de se enviar petroleiros para outro porto?
   (*b*) De quanto deverão ser aumentadas as instalações para permitir atender a todos os navios que chegarem pelo menos em 95% dos dias?
   (*c*) Qual o número médio de petroleiros que chegam por dia?

35. Na tabela abaixo, X significa número de filhos homens em famílias com 12 filhos. Calcule para cada valor da variável o número de famílias que você deveria esperar se $X \sim b(12; 0{,}5)$.

X	Nº observado de famílias
0	6
1	29
2	160
3	521
4	1.198
5	1.921
6	2.360
7	2.033
8	1.398
9	799
10	298
11	60
12	7
Total	10.690

Você acha que o modelo binomial é razoável para explicar o fenômeno?

36. Houve uma denúncia por parte dos operários de uma indústria de que, toda vez que ocorria um acidente em uma seção da indústria, ocorriam outros em outras seções mais ou menos no mesmo horário. Em outras palavras, os acidentes não estavam ocorrendo ao acaso. Para verificar essa hipótese, foi feita uma contagem do número de acidentes por hora durante um certo número de dias (24 horas por dia). Os resultados da pesquisa foram apresentados no quadro a seguir.

ixa de suas peças conterá, no
experiência tem demonstrado
ituosas, qual a probabilidade

a certa população de funcio-
rticipam desse curso, encontre

ade;
lutividade; e
tividade.

ada dia ocorre segundo uma
em atender, no máximo, a três
cesso é enviado a outro porto.
os para outro porto?
a permitir atender a todos os

dia?

amílias com 12 filhos. Calcule
veria esperar se $X \sim b(12; 0{,}5)$.

fenômeno?

ústria de que, toda vez que
utros em outras seções mais
tes não estavam ocorrendo
m do número de acidentes
. Os resultados da pesquisa

Nº de acidentes por hora	Nº de horas
0	200
1	152
2	60
3	30
4	13
5	9
6	7
7	5
8	4

(a) Calcule o número médio de acidentes por hora nessa amostra.

(b) Se o número de acidentes por hora seguisse uma distribuição de Poisson, com média igual à que você calculou, qual seria o número esperado de dias com $0, 1, 2, \ldots$ etc. acidentes?

(c) Os dados revelam que a suspeita dos operários é verdadeira?

37. Determinado tipo de parafuso é vendido em caixas com 1.000 peças. É uma característica da fabricação produzir 10% com defeito. Normalmente, cada caixa é vendida por \$13,50. Um comprador faz a seguinte proposta: de cada caixa, ele escolhe uma amostra de 20 peças; se a caixa não tiver parafusos defeituosos, ele paga \$20,00; um ou dois defeituosos, ele paga \$10,00; três ou mais defeituosos, ele paga \$8,00. Qual alternativa é a mais vantajosa para o fabricante? Justifique.

38. Uma certa região florestal foi dividida em 109 quadrados para estudar a distribuição de *Primula Simenses Selvagem*. *A priori*, supomos que esse tipo distribua-se aleatoriamente na região. O quadro abaixo indica o número de quadrados com X *Primula Simenses*; o número médio de plantas por quadrado foi de 2,2.

X plantas por quadrado	Nº de quadrados com X plantas
0	26
1	21
2	23
3	14
4	11
5	4
6	5
7	4
8	1
acima de 8	0

(a) Se as plantas realmente se distribuem aleatoriamente na região, qual a probabilidade de encontrarmos pelo menos duas *Primulas*?

(b) Dê as freqüências esperadas para os valores de $X = 0$, $X = 1$ e $X = 2$.

(c) Apenas comparando os resultados de *(b)* com as freqüências observadas, qual a conclusão a que você chegaria?

(d) Quais as causas que você daria para a conclusão?

39. Uma fábrica produz válvulas, das quais 20% são defeituosas. As válvulas são vendidas em caixas com dez peças. Se uma caixa não tiver nenhuma defeituosa, seu preço de venda é \$10,00; tendo uma, o preço é \$8,00; duas ou três, o preço é \$6,00; mais do que três, o preço é \$2,00. Qual o preço médio de uma caixa?

40. Um industrial fabrica peças, das quais 1/5 são defeituosas. Dois compradores A e B, classificaram as partidas adquiridas em categorias I e II, pagando \$1,20 e \$0,80 respectivamente do seguinte modo:

    Comprador A: retira uma amostra de cinco peças; se encontrar mais que uma defeituosa, classifica como II.

    Comprador B: retira amostra de dez peças; se encontrar mais que duas defeituosas, classifica como II.

    Em média, qual comprador oferece maior lucro?

41. Se $X \sim b(n, p)$, prove que $E(X) = np$ e $\text{Var}(X) = npq$.
    (Sugestão: calcule $E(X)$ e $\text{Var}(X)$ para $n = 1, 2, \ldots$ etc.)

42. Aceitação de um lote. Suponha que um comprador queira decidir se vai aceitar ou não um lote de itens. Para isso, ele retira uma amostra de tamanho n do lote e conta o número x de defeituosos. Se $x \leq a$, o lote é aceito, e se $x > a$, o lote é rejeitado; o número a é fixado pelo comprador. Suponha que $n = 19$ e $a = 2$. Use a Tabela I a fim de encontrar a probabilidade de aceitar o lote, ou seja, $P(X \leq 2)$ para as seguintes proporções de defeituosos no lote:

    (a) $p = 0{,}10$ (b) $p = 0{,}20$ (c) $p = 0{,}05$

43. Prove que, quando $n \to \infty$ e $p \to 0$, mas de tal sorte que $np \to \lambda$, temos

$$\binom{n}{k} p^k \cdot (1-p)^{n-k} \to \frac{e^{-\lambda} \cdot \lambda^k}{k!}.$$

    Sugerimos que você use o fato: $\left(1 - \dfrac{\lambda}{n}\right)^n \to e^{-\lambda}$ quando $n \to \infty$.

44. Suponha que X seja uma v.a. discreta, com f.p. $p(x) = 2^{-x}$, $x = 1, 2, \ldots$ Calcule:

    (a) $P(X \text{ ser par})$ (b) $P(X \leq 3)$ (c) $P(X > 10)$

45. Prove (6.4), (6.5) e (6.6).

46. Prove que $E(X) = \text{Var}(X) = \lambda$, se a $P(X = k)$ for dada por (6.24).

47. Prove a relação (6.19).

48. Num teste tipo certo/errado, com 50 questões, qual é a probabilidade de que um aluno acerte 80% das questões, supondo que ele as responda ao acaso?

49. Repita o Problema 48, considerando cinco alternativas para cada questão.

50. Em um experimento binomial com três provas, a probabilidade de exatamente dois sucessos é 12 vezes a probabilidade de três sucessos. Encontre p.

51. No sistem
pendênci

(a) o sist
(b) o sist
(c) exata
(d) pelo

52. Prove que

53. Encontre

54. Encontre

55. Distribuiç
probabili
o experim
Seja X =

pois se X

(a) Prov

(b) Mos
[Sug

(c) Se s

s válvulas são vendidas
efeituosa, seu preço de
ço é $6,00; mais do que

ois compradores A e B,
lo $1,20 e $0,80 respec-

ais que uma defeituosa,

que duas defeituosas,

dir se vai aceitar ou não
no n do lote e conta o
e é rejeitado; o número
se a Tabela I a fim de
ra as seguintes propor-

)5

temos

2,... Calcule:
10)

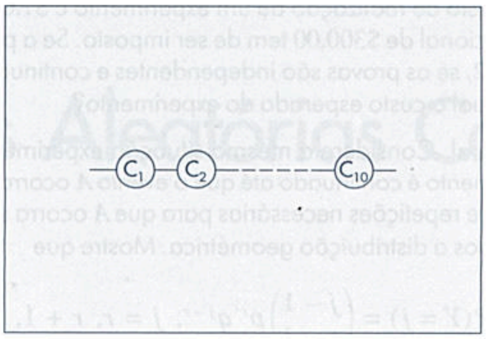
dade de que um aluno
o?

a questão.

exatamente dois suces-

51. No sistema abaixo, cada componente tem probabilidade p de funcionar. Supondo independência de funcionamento dos componentes, qual a probabilidade de:

X



(a) o sistema funcionar?
(b) o sistema não funcionar?
(c) exatamente dois componentes funcionarem?
(d) pelo menos cinco componentes funcionarem?

52. Prove que

$$b(k+1; n, p) = \frac{(n-k)p}{(k+1)(1-p)} \cdot b(k; n, p).$$

53. Encontre a mediana da v.a. Z com distribuição

Z	0	1	2	3
$p(Z)$	1/4	1/4	1/4	1/4

54. Encontre os quantis de ordens $p = 0{,}25,\ 0{,}60,\ 0{,}80$ da v.a. Z do exercício 53.

55. Distribuição Geométrica. Suponha que, ao realizar um experimento, ocorra o evento A com probabilidade p ou não ocorra A (ou seja, ocorre A^c com probabilidade $1-p$). Repetimos o experimento de forma independente até que o evento A ocorra pela primeira vez.

Seja X = número de repetição do experimento até que se obtenha A pela primeira vez. Então,

$$P(X=j) = (1-p)^{j-1}.\ p, \quad j = 1, 2, 3, \ldots,$$

pois se $X = j$, nas primeiras $j-1$ repetições A não ocorre, ocorrendo na j-ésima.

(a) Prove que $\sum_{j=1}^{\infty} P(X=j) = 1$.

(b) Mostre que $E(X) = 1/\text{p}$ e $\text{Var}(X) = (1-p)/\text{p}^2$..

[Sugestão: $E(X) = \sum_{j=1}^{\infty} j \cdot p(X=j) = p \sum_{j=1}^{\infty} j \cdot (1-p)^{j-1} = p \sum_{j=1}^{\infty} \frac{d}{dq} q^j$, com $1-p=q$.]

(c) Se s e t são inteiros positivos, então

$$P(X > s+t \mid X > s) = P(X > t).$$

Essa propriedade nos diz que a distribuição geométrica não tem memória. Essa propriedade é compartilhada pela distribuição exponencial, a ser estudada no Capítulo 7.

56. (Meyer, 1965). O custo de realização de um experimento é \$1.000,00. Se o experimento falha, um custo adicional de \$300,00 tem de ser imposto. Se a probabilidade de sucesso em cada prova é 0,2, se as provas são independentes e continuadas até a ocorrência do primeiro sucesso, qual o custo esperado do experimento?

57. Distribuição de Pascal. Considere a mesma situação experimental do Problema 55, só que agora o experimento é continuado até que o evento A ocorra pela r-ésima vez. Defina a v.a. Y = número de repetições necessárias para que A ocorra exatamente r vezes. Note que, se $r = 1$, obtemos a distribuição geométrica. Mostre que

$$P(Y = j) = \binom{j-1}{r-1} p^r q^{j-r},\ j = r,\ r+1,\ \ldots$$

58. A Desigualdade de Jensen. Vimos, na fórmula (6.4), que se $h(x) = ax + b$, então $E[h(X)] = h[E(X)]$, ou seja, $E(aX + b) = aE(X) + b$.
Esta fórmula pode não valer se $h(x)$ não for linear. O que vale é o seguinte resultado, denominado Desigualdade de Jensen. Se $h(x)$ for uma função convexa e X uma v.a., então

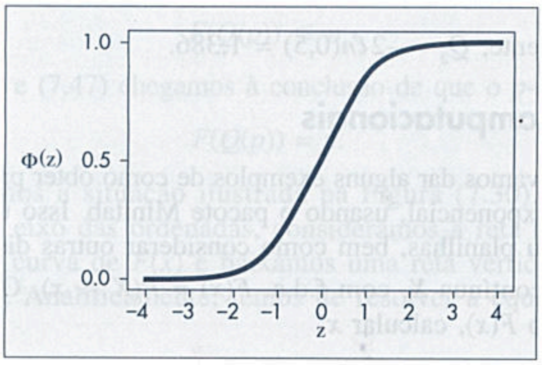
$$E[h(X)] \geqslant h[E(X)],$$

com igualdade se e somente se h for linear (ou se a variância de X for zero).
Por exemplo, se $h(x) = x^2$, então $E(X^2) \geqslant [E(X)]^2$, do que decorre que $\text{Var}(X) = E(X^2) - [E(X)]^2 \geqslant 0$.
Lembremos que uma função h é convexa se $h((x+y)/2) \leqslant (h(x) + h(y))/2$, para todo par x, y no domínio de h. Em termos geométricos, h é convexa se o ponto médio da corda que une dois pontos quaisquer da curva representando h está acima da curva. A função h é côncava se $-h$ for convexa. Por exemplo, $\log x$ é uma função côncava.

59. Use o problema anterior para verificar as relações entre:
(a) $E(e^X)$ e $e^{E(X)}$;
(b) $E(\log X)$ e $\log [E(X)]$, para $X > 0$;
(c) $E(1/X)$ e $1/E\{X)$, para $X \neq 0$.

dos no intervalo [–4, 4], podemos considerar um vetor de valores $z = [-4{,}0;\ -3{,}9;\ -3{,}8;\ \ldots;\ 3{,}8;\ 3{,}9;\ 4{,}0]$ e obter os valores da f.d.a. com o comando CDF. Depois, pedir para plotar os pares $(z_i, F(z_i))$. O gráfico está na Figura 7.31.

Figura 7.31: Gráfico da f.d.a. da $N(0, 1)$. Minitab.



7.10 Problemas e Complementos

28. Numa determinada localidade, a distribuição de renda (em reais) é uma v.a. X com f.d.p.

$$f(x)=\begin{cases}\dfrac{1}{10}x+\dfrac{1}{10}, & 0\leqslant x\leqslant 2\\[2mm] -\dfrac{3}{40}x+\dfrac{9}{20}, & 2< x\leqslant 6\\[2mm] 0, & x<0 \text{ ou } x>6.\end{cases}$$

(*a*) Qual a renda média nessa localidade?

(*b*) Escolhida uma pessoa ao acaso, qual a probabilidade de sua renda ser superior a \$3.000,00?

(*c*) Qual a mediana da variável?

29. Se X tiver distribuição uniforme com parâmetros α e β, mostre que:

(*a*) $E(X) = \dfrac{\alpha+\beta}{2}$.

(*b*) $\mathrm{Var}(X) = (\beta-\alpha)^2/12$.

(*c*) $F(x)=\begin{cases}0, & x<\alpha\\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha\leqslant x\leqslant\beta\\ 1, & x>\beta.\end{cases}$

30. Complete a tabela abaixo, que corresponde a alguns valores da função

$$G(u)=P(0\leqslant U\leqslant u),$$

definida na seção 7.4.1, com U uma v.a. uniforme no intervalo (–1/2, 1/2).

Probabilidades p, tais que $p = P(0 \leq U \leq u)$

Primeira decimal de u	Segunda decimal de u				Primeira decimal de u
	0	1	…	9	
0,0					0,0
0,1					0,1
0,2					0,2
0,3					0,3
0,4					0,4
0,5					0,5

31. Dada a v.a. X, uniforme em $(5, 10)$, calcule as probabilidades abaixo, usando a tabela do problema anterior.
    (a) $P(X < 7)$
    (b) $P(8 < X < 9)$
    (c) $P(X > 8{,}5)$
    (d) $P(|X - 7{,}5| > 2)$

32. Se $X \sim N(\mu, \sigma^2)$, calcular $E(X)$ e $\mathrm{Var}(X)$.

    [Sugestão: Fazendo a transformação de variáveis $x = \mu + \sigma t$, obtemos que $E(X) = \frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2/2}\,dt + \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} te^{-t^2/2}\,dt$. A primeira integral resulta μ (por quê?) e a segunda anula-se, pois o integrando é uma função ímpar. Para obter a variância, obtenha $E(X^2)$ por integração por partes.]

33. As notas de Estatística Econômica dos alunos de determinada universidade distribuem-se de acordo com uma distribuição normal, com média 6,4 e desvio padrão 0,8. O professor atribui graus A, B e C da seguinte forma:

Nota	Grau
$x < 5$	C
$5 \leq x < 7{,}5$	B
$7{,}5 \leq x \leq 10$	A

    Numa classe de 80 alunos, qual o número esperado de alunos com grau A? E com grau B? E C?

34. O peso bruto de latas de conserva é uma v.a. normal, com média 1.000 g e desvio padrão 20 g.
    (a) Qual a probabilidade de uma lata pesar menos de 980 g?
    (b) Qual a probabilidade de uma lata pesar mais de 1.010 g?

35. A distribuição dos pesos de coelhos criados numa granja pode muito bem ser representada por uma distribuição normal, com média de 5 kg e desvio padrão de 0,8 kg. Um abatedouro comprará 5.000 coelhos e pretende classificá-los de acordo com o peso, do seguinte modo: 20% dos leves como pequenos, os 55% seguintes como médios, os 15% seguintes como grandes e os 10% mais pesados como extras. Quais os limites de peso para cada classe?

36. Uma enchedora automática de garrafas de refrigerantes está regulada para que o volume médio de líquido em cada garrafa seja de $1.000\ cm^3$ e o desvio padrão de $10\ cm^3$. Pode-se admitir que a variável volume seja normal.
    (a) Qual é a porcentagem de garrafas em que o volume de líquido é menor que $990\ cm^3$?
    (b) Qual é a porcentagem das garrafas em que o volume líquido não se desvia da média em mais que dois desvios padrões?
    (c) O que acontecerá com a porcentagem do item *(b)* se a máquina for regulada de forma que a média seja $1.200\ cm^3$ e o desvio padrão $20\ cm^3$?

37. O diâmetro de certo tipo de anel industrial é uma v.a. com distribuição normal, de média $0{,}10$ cm e desvio padrão $0{,}02$ cm. Se o diâmetro de um anel diferir da média em mais que $0{,}03$ cm, ele é vendido por \$5,00; caso contrário, é vendido por \$10,00. Qual o preço médio de venda de cada anel?

38. Uma empresa produz televisores e garante a restituição da quantia paga se qualquer televisor apresentar algum defeito grave no prazo de seis meses. Ela produz televisores do tipo A (comum) e do tipo B (luxo), com lucros respectivos de \$1.000,00 e \$2.000,00, caso não haja restituição, e com prejuízos de \$3.000,00 e \$8.000,00, se houver restituição. Suponha que o tempo para a ocorrência de algum defeito grave seja, em ambos os casos, uma v.a. com distribuição normal, respectivamente, com médias 9 meses e 12 meses, e variâncias 4 meses^2 e 9 meses^2. Se tivesse de planejar uma estratégia de *marketing* para a empresa, você incentivaria as vendas dos aparelhos do tipo A ou do tipo B?

39. Determine as médias das v.a. X, Y e Z:
    (a) X uniforme em $(1, 3)$, $Y = 3X + 4$, $Z = e^X$.
    (b) X tem f.d.p. $f(x) = e^{-x}$, $x > 0$, $Y = X^2$, $Z = 3/(X+1)^2$.

40. Suponha que X tenha distribuição uniforme em $[-a, 3a]$. Determine a média e a variância de X.

41. Se T tiver distribuição exponencial com parâmetro β, mostre que:
    (a) $E(T) = \beta$. *(b)* $\mathrm{Var}(T) = \beta^2$.

42. Os dados a seguir representam uma amostra de firmas de determinado ramo de atividade de uma região. Foram observadas duas variáveis: faturamento e número de empregados.

Nº de empregados	Nº de empresas
0 ⊢ 20	35
20 ⊢ 50	75
50 ⊢ 100	45
100 ⊢ 200	30
200 ⊢ 400	15
400 ⊢ 800	8
> 800	2
Total	210

Faturamento	Nº de empresas
0 ⊢ 10	18
10 ⊢ 50	52
50 ⊢ 100	30
100 ⊢ 200	26
200 ⊢ 400	24
400 ⊢ 800	20
800 ⊢ 1600	16
1600 ⊢ 3200	14
3200 ⊢ 6400	6
> 6400	4
Total	210

(a) Calcule a n
(b) Supondo n
pela amostr
com o obse

43. Suponha que a
mentar. Faça Y =
(a) Determine I
(b) Determine a
(c) Calcule $E(X$
(d) Calcule $E(Y$

44. Dada a v.a.

determine a méc

45. *(a)* Prove que, s
(b) Prove que Γ
(c) Calcule $\Gamma(1)$
(d) Prove que a
em (7.32)) sć

46. Distribuição de P
conexão com prc
Dizemos que a v.
for dada por

Aqui, b pode repr
proporção de indi

(a) Calcule a média e a variância para cada variável.

(b) Supondo normalidade para cada uma dessas variáveis, com parâmetros estimados pela amostra, calcule os valores esperados para cada intervalo de classe e compare com o observado.

43. Suponha que a v.a. X tenha densidade $f(x) = 1$, para $0 < x < 1$ e igual a zero no complementar. Faça $Y = X^2$.

(a) Determine $F_Y(y) = P(Y \leqslant y)$, y real.

(b) Determine a f.d.p. de Y.

(c) Calcule $E(X^2)$, utilizando a f.d.p. de X.

(d) Calcule $E(Y)$, utilizando a f.d.p. de Y, e compare com (c).

44. Dada a v.a.

$$Z = \frac{X - \mu_x}{\sigma_x},$$

determine a média e a variância de Z, sabendo-se que a f.d.p. de X é

$$f(x) = e^{-x}, x > 0.$$

45. (a) Prove que, se α for inteiro positivo, $\Gamma(\alpha) = (\alpha - 1)!$.

(b) Prove que $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$.

(c) Calcule $\Gamma(1)$ e $\Gamma(1/2)$.

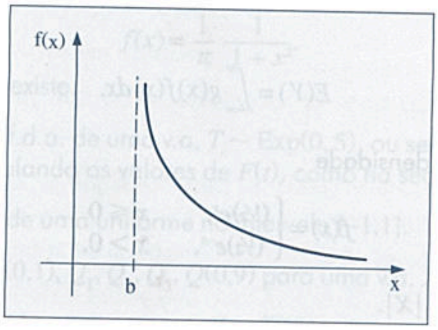
(d) Prove que a média e a variância de uma v.a. X com distribuição gama (densidade em (7.32)) são, respectivamente, $\alpha\beta$ e $\alpha\beta^2$.

46. Distribuição de Pareto. Esta é uma distribuição freqüentemente usada em Economia, em conexão com problemas de distribuição de renda.

Dizemos que a v.a. X tem *distribuição de Pareto* com parâmetros $\alpha > 0$, $b > 0$ se sua f.d.p. for dada por

$$f(x) = \begin{cases} \alpha/b\ (b/x)^{\alpha+1}, & x \geqslant b \\ 0, & x < b. \end{cases}$$

Aqui, b pode representar algum nível mínimo de renda, x é o nível de renda e $f(x)\,\Delta x$ dá a proporção de indivíduos com renda entre x e $x + \Delta x$. O gráfico de $f(x)$ está na figura abaixo.



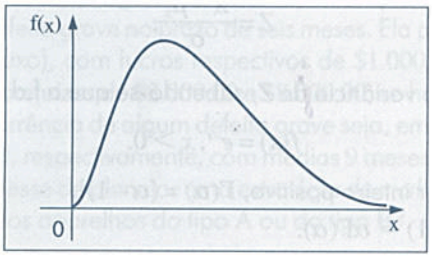
(a) Prove que $\int_{-\infty}^{\infty} f(x)dx = 1$.

(b) Mostre que, para $\alpha > 1$, $E(X) = \dfrac{\alpha b}{\alpha - 1}$ e para $\alpha > 2$, $\text{Var}(X) = \dfrac{\alpha b^2}{(\alpha-1)^2(\alpha-2)}$.

47. Distribuição lognormal. Outra distribuição usada quando se têm valores positivos é a distribuição lognormal. A v.a. X tem distribuição lognormal, com parâmetros μ e σ^2, $-\infty < \mu < \infty$, $\sigma^2 > 0$, se $Y = \ell nX$ tiver distribuição normal com média μ e variância σ^2. A f.d.p. de X tem a forma

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{\ell nx - \mu}{\sigma}\right)^2}, & \text{se } x > 0 \\ 0, & \text{se } x \leqslant 0. \end{cases}$$

O gráfico de $f(x)$ está na figura abaixo.



(a) Prove que $E(X) = e^{\mu + \sigma^2/2}$.

(b) Se $E(X) = m$, prove que $\text{Var}(X) = m^2(e^{\sigma^2} - 1)$.

48. Suponha que X tenha distribuição exponencial com parâmetro β. Prove que

$$\frac{P(X > t + x)}{P(X > x)} = P(X > t), \forall t, x \geqslant 0.$$

Essa propriedade nos diz que a distribuição exponencial não tem memória. Por exemplo, se X for a vida de um componente eletrônico, a relação acima diz que, se o componente durou até o instante x, a probabilidade de ele não falhar após o intervalo $t + x$ é a mesma de não falhar após o instante t. Nesse sentido, X "esquece" a sua idade, e a eventual falha do componente não resulta de uma deterioração gradual e sim de alguma falha repentina.

49. Se X for uma v.a. contínua, com f.d.p. $f(x)$, e se $Y = g(X)$ for uma função de X, então Y será uma v.a com

$$E(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Suponha que X tenha densidade

$$f(x) = \begin{cases} (½)e^{x}, & x \leqslant 0 \\ (½)e^{-x}, & x > 0. \end{cases}$$

Obtenha $E(Y)$, se $Y = |X|$.

50. Se X for uniforme no intervalo $[0, 1]$, obtenha a média da v.a. $Y = (½)X^2$.

51. Distribuição de Weibull. Um modelo que tem muitas aplicações na *teoria da confiabilidade* é o modelo de Weibull, cuja f.d.p. é dada por

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1}e^{-\alpha x^\beta}, & x \geq 0 \\ 0, & x < 0, \end{cases}$$

onde α e β são constantes positivas. A v.a. X pode representar, por exemplo, o tempo de vida de um componente de um sistema.

(*a*) Se $\beta = 1$, qual a f.d.p. resultante? (*b*) Obtenha $E(X)$ para $\beta = 2$.

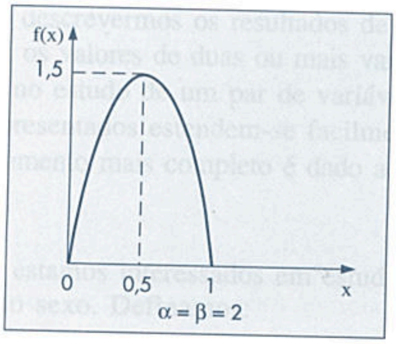
52. Distribuição Beta. Uma v.a. X tem distribuição beta com parâmetros $\alpha > 0$, $\beta > 0$, se sua f.d.p. for dada por

$$f(x) = \begin{cases} \dfrac{1}{B(\alpha, \beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{caso contrário.} \end{cases}$$

Aqui, $B(\alpha, \beta)$ é a função beta, definida por

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx.$$

É possível provar que $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. A figura abaixo mostra a densidade da distribuição beta para $\alpha = \beta = 2$. Para esse caso, calcule $P(X \leq 0{,}2)$. Calcule a média e a variância de X para $\alpha = \beta = 2$.



53. Se na distribuição t de Student colocarmos $v = 1$, obteremos a *distribuição de Cauchy*,

$$f(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}.$$

Mostre que $E(X)$ não existe.

54. Obtenha o gráfico da f.d.a. de uma v.a. $T \sim \text{Exp}(0, 5)$, ou seja, $E(T) = 2$, considerando 20 valores de T e calculando os valores de $F(t)$, como na seção 7.9.

55. Idem, para 30 valores de uma uniforme no intervalo $[-1,1]$.

56. Obtenha os quantis $Q(0{,}1)$, Q_1, Q_2, Q_3, $Q(0{,}9)$ para uma v.a. $X \sim N(10; 16)$.

57. Resolva a mesma questão para uma v.a. $Y \sim \chi^2(5)$.

58. Para uma v.a. com distribuição qui-quadrado, com ν graus de liberdade e ν *par*, vale a seguinte fórmula:

$$P(\chi^2(\nu) > c) = e^{-c/2} \sum_{j=0}^{\nu/2-1} \frac{(c/2)^j}{j!}.$$

Calcule essa probabilidade para os seguintes casos e compare com os valores tabelados na Tabela IV:

(a) $\nu = 4$, $c = 9{,}488$; (b) $\nu = 10$, $c = 16$.

59. Usando a aproximação normal a uma variável qui-quadrado, calcular:

(a) $P(\chi^2(35) > 49{,}76)$; (b) o valor y tal que $P(\chi^2(40) > y) = 0{,}05$.

60. Se $X \sim N(\mu\ ,\ \sigma^2)$, com densidade $f(x)$ dada por (7.17), provemos que a integral $I = \int_{-\infty}^{\infty} f(x)\,dx = 1$. Como esta integral é sempre positiva, mostremos que $I^2 = 1$. Novamente, como no Problema 32, fazemos a transformação $x = \mu + \sigma t$ e obtemos $I^2 = \frac{1}{2\pi}\int\int e^{-(t^2+s^2)/2} dsdt$, onde os limites de integração são $-\infty$ e ∞. Agora fazemos outra transformação, passando de coordenadas cartesianas para polares: $s = r\cos\theta$, $t = r\ \text{sen}\ \theta$, de modo que $dsdt = r\ drd\theta$. Segue-se, integrando primeiro com relação a r e depois com relação a θ, que

$$I^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-r^2/2} r dr d\theta = \frac{1}{2\pi}\int_0^{2\pi}[-e^{-r^2/2}]_0^{\infty} d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta = 1.$$

Variáve
Multidi

8.1 Distribuição

Em muitas situaç
a um mesmo ponto a
tulo, iremos nos conc
os conceitos e result
variáveis aleatórias. U
nas seções 8.1 a 8.4.

Exemplo 8.1. Suponl
com três crianças, qu

X = número de n

$Y = \begin{cases} 1, \text{ se o prim} \\ 0, \text{ se o prim} \end{cases}$

Z = número de ve
dentro da me

Com essas inforn
ma probabilidade, ol
que o primeiro filho

As distribuições d
e são dadas na Tabel

Problemas

15. Usando um pacote de sua preferência, gere:
   (a) 100 valores de uma distribuição binomial, com parâmetros $n = 15$, $p = 0{,}7$.
   (b) 500 valores de uma distribuição de Poisson com parâmetro $\lambda = 1{,}3$.
   Em cada caso, faça um histograma e veja se este corresponde à distribuição em questão.

16. Usando um pacote computacional de sua preferência, gere:
   (a) 500 valores de uma normal padrão;
   (b) 1.000 valores de uma distribuição qui-quadrado com cinco graus de liberdade;
   (c) 800 valores de uma distribuição exponencial com parâmetro 3.
   Em cada caso, faça um histograma, um ramo-e-folhas e um *box plot*. Comente.

17. Usando o método de Box-Müller, gere cinco valores de uma distribuição normal padrão.

9.5 Problemas e Complementos

18. O método dos quadrados centrais de von Neumann opera do modo descrito a seguir. Considere um inteiro n_0 com m dígitos e seu quadrado n_0^2, que terá $2m$ dígitos (eventualmente acrescentando zeros à esquerda). Tome os dígitos centrais de n_0^2 e divida o número obtido por 10^m para se obter um NA, u_0, entre 0 e 1. Continue, tomando n_1 como o número inteiro central desse passo.

   Esse método pode não funcionar bem, como o exemplo abaixo de Kleijnen e van Groenendaal (1994) mostra.

   Suponha $m = 2$ e considere $n_0 = 23$. Então, $n_0^2 = 0529$, e o primeiro NA é $u_0 = 0{,}52$. Agora, $n_1 = 52$, $n_1^2 = 2704$ e $u_1 = 0{,}70$. Sucessivamente, obtemos $u_2 = 0{,}90$, $u_3 = 0{,}10$, $u_4 = 0{,}10$ etc. Ou seja, a partir de u_4, os NA se repetem.

   Obtenha números aleatórios, com $m = 3$, usando esse método.

19. Uma distribuição binomial de parâmetros n e p pode ser simulada também do seguinte modo. Considere a recursão

$$p_{j+1} = \frac{n-j}{j+1}\,\frac{p}{1-p}\,p_j,$$

   com $p_j = P(X = j)$, $j = 0, 1, \ldots, n$.

   Chame j o valor atual, $pr = P(X = j)$, $F = F(j) = P(X \leqslant j)$ e o algoritmo:

   Passo 1. Gere o NA u;

   Passo 2. $r = p/(1-p)$, $j = 0$, $pr = (1-p)^n$, $F = pr$;

   Passo 3. Se $u < F$, coloque $X = j$;

   Passo 4. $pr = \dfrac{r(n-j)}{j+1}\,pr$, $F = F + pr$, $j = j + 1$.

   Passo 5. Volte ao passo 3.

   Usando esse algoritmo, gere cinco valores da v.a. $X \sim b(5; 0{,}3)$.

tros $n = 15$, $p = 0,7$.
tro $\lambda = 1,3$.
à distribuição em questão.

nco graus de liberdade;
netro 3.
box plot. Comente.
distribuição normal padrão.

do modo descrito a seguir.
e terá $2m$ dígitos (eventual-
rais de n_0^2 e divida o número
inue, tomando n_1 como o

abaixo de Kleijnen e van

meiro NA é $u_0 = 0,52$.
emos $u_2 = 0,90$, $u_3 = 0,10$,

.

lada também do seguinte

lgoritmo:

,3).

20. Simulação de uma distribuição de Poisson. Se $N \sim P(\lambda)$, então $P(N = j) = p_j$ é dada por

$$P(N = j) = \frac{e^{-\lambda}\lambda^j}{j!}, j = 0, 1, \ldots \tag{9.7}$$

A geração de valores de uma distribuição de Poisson parte da seguinte relação recursiva, que pode ser facilmente verificada:

$$p_{j+1} = \frac{\lambda}{j+1}\, p_j, \quad j \geqslant 0. \tag{9.8}$$

Seja, também, $F(j) = P(N \leqslant j)$ a função de distribuição acumulada (f.d.a.) de N. Considere j o valor atual gerado e queremos gerar o valor seguinte. Chamemos simplesmente $p = p_j$ e $F = F(j)$. Então o algoritmo para se gerar os sucessivos valores é o seguinte:

Passo 1. Gere o NA u;

Passo 2. Faça $j = 0$, $p = e^{-\lambda}$ e $F = p$;

Passo 3. Se $u < F$, coloque $N = j$;

Passo 4. Faça $p = \dfrac{\lambda}{j+1} p$, $F = F + p$ e $j = j + 1$;

Passo 5. Volte ao *Passo* 3.

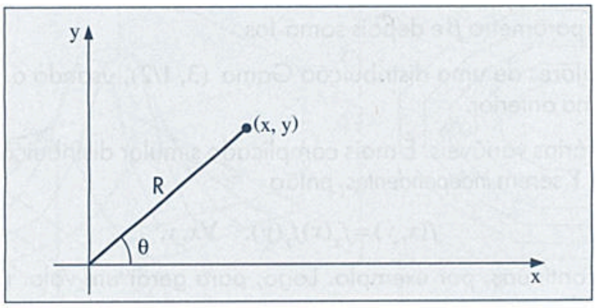
Note que, no *Passo* 2, se $j = 0$, $P(N = 0) = p_0 = e^{-\lambda}$ e $F(0) = P(N \leqslant 0) = p_0$.

21. Usando o procedimento recursivo do Problema 20, gere cinco valores de uma v.a. com distribuição de Poisson de parâmetro $\lambda = 2$.

22. Transformação de Box-Müller. Considere as v.a. X e Y, independentes e ambas $N(0, 1)$. Observando a Figura 9.10, vemos que $R^2 = X^2 + Y^2$ e $\text{tg}\,\theta = Y/X$. A densidade conjunta de X e Y é

$$f(x, y) = \frac{1}{2\pi}\, e^{-(x^2 + y^2)/2}.$$

Figura 9.10: Transformação de variáveis $(x, y) \to (R, \theta)$.



Considere a transformação de variáveis

$$r = x^2 + y^2$$
$$\theta = \text{arctg}(y/x).$$

A densidade conjunta de r e θ é obtida usando o resultado (8.28). Temos que $x = \sqrt{r}\cos\theta$, $y = \sqrt{r}\operatorname{sen}\theta$ e o Jacobiano da transformação é $|J| = 1/2$. Segue-se que a densidade de r e θ é

$$f(r, \theta) = 1/2\pi \cdot e^{-r^2} \cdot 1/2, 0 < r < \infty, 0 < \theta < 2\pi.$$

Dessa relação podemos concluir que $r = R^2$ e θ são *independentes*, com

$$R^2 \sim \text{Exp}(2),\ \theta \sim \mathcal{U}(0, 2\pi).$$

Portanto, podemos escrever que

$$X = R\cos\theta = \sqrt{-2\log U_1}\cos(2\pi U_2)$$
$$Y = R\operatorname{sen}\theta = \sqrt{-2\log U_1}\operatorname{sen}(2\pi U_2)$$

Aqui, usamos o fato de que, se $R^2 \sim \text{Exp}(2)$, gerado um NA U_1, vem que $-2\log U_1 \sim \text{Exp}(2)$ e se $\theta \sim \mathcal{U}(0, 2\pi)$, então gerado um NA U_2, vem que $2\pi U_2 \sim \mathcal{U}(0, 2\pi)$.

23. Usando um aplicativo estatístico, gere:
    (a) 100 valores de uma $N(5; 0{,}9)$ e faça o histograma dos valores gerados.
    (b) 200 valores de uma $\text{Exp}(1/2)$ e faça o histograma dos valores gerados.
    (c) 500 valores de uma Gama(α, β), com $\alpha = \beta = 2$, e faça o histograma.
    (d) 300 valores de uma $\chi^2(32)$ e faça o histograma.

    Os histogramas que você obteve estão de acordo com as definições dadas dessas distribuições? Comente.

24. Usando um pacote, gere:
    (a) 300 valores de uma distribuição $t(120)$.
    (b) 500 valores de uma distribuição $F(56, 38)$.
    (c) 300 valores de uma distribuição $B(20, 30)$.

    Faça um histograma dos valores simulados em cada caso e responda a mesma pergunta do problema anterior.

25. Simulação de uma distribuição gama. Pode-se demonstrar, usando resultados não estudados neste livro, que se a v.a. $X \sim$ Gama(r, β), com r inteiro, então $X = Y_1 + Y_2 + ... + Y_r$, onde cada $Y_i \sim \text{Exp}(\beta)$ e as v.a. Y_i são independentes. Logo, para gerar um valor de uma distribuição Gama(r, β), com $r > 0$, inteiro, basta gerar r valores de uma distribuição exponencial de parâmetro β e depois somá-los.

26. Simule cinco valores de uma distribuição Gama $(3, 1/2)$, usando o procedimento descrito no problema anterior.

27. Simulação de várias variáveis. É mais complicado simular distribuições bidimensionais. No caso de X e Y serem *independentes*, então

$$f(x, y) = f_X(x)f_Y(y), \quad \forall x, y,$$

se elas forem contínuas, por exemplo. Logo, para gerar um valor (x, y) da densidade conjunta $f(x, y)$, basta gerar o componente x da distribuição marginal de X e a componente y da distribuição marginal de Y, *independentemente*.

No caso de v.a. dependentes, temos que vale a relação:

$$f(x, y) = f_X(x)f_{Y|X}(y|x).$$

Logo, por essa relação, primeiramente geramos um valor x da distribuição marginal de X e fixado esse valor, x_0, digamos, geramos um valor da distribuição condicional de X, dado que $X = x_0$. Isso implica que devemos saber como gerar valores das distribuições $f_X(x)$ e $f_{Y|X}(y|x)$.

Vamos nos limitar a dar dois exemplos no caso de v.a. independentes.

Exemplo 9.18. *Distribuição uniforme bidimensional.*

Na seção 9.1 vimos que para calcular a área da figura F contida no quadrado Q de lado unitário (Figura 9.1), considerávamos o quociente N'/N. Como geramos, naquele exemplo, os N pontos uniformemente distribuídos sobre Q? Pelo que vimos acima, basta gerar valores de v.a. $U_1 \sim \mathcal{U}(0, 1)$ e $U_2 \sim \mathcal{U}(0, 1)$, independentemente. Então,

$$P((U_1, U_2) \in F) = \text{área}(F).$$

Ou seja, a v.a. (U_1, U_2) é uniformemente distribuída em Q.

No caso da Figura 9.1, consideramos 200 valores gerados para U_1 e U_2, de modo que a área $(F) = 24/100$.

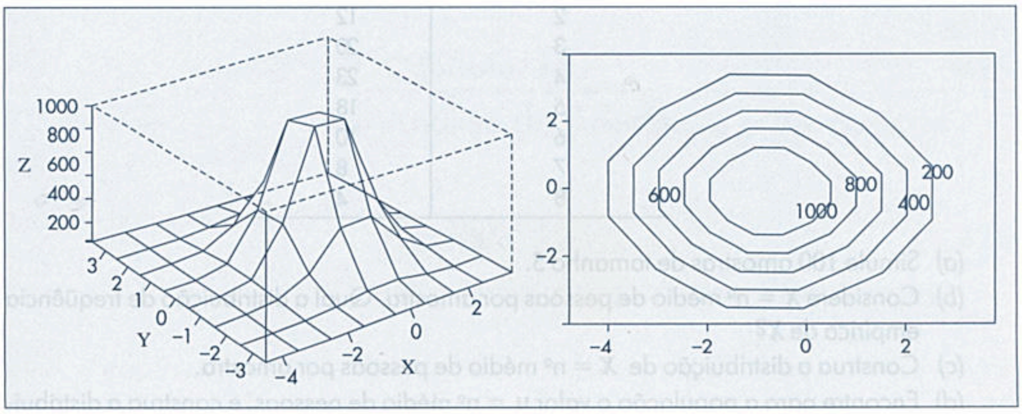
Exemplo 9.19. *Distribuição normal bidimensional.*

O método de Box-Müller gera valores de duas normais padrões independentes, Z_1 e Z_2. Logo, se quisermos gerar valores da distribuição conjunta de X e Y, independentes e normais, com $X \sim N(\mu_x, \sigma_x^2)$ e $Y \sim N(\mu_y \sim \sigma_y^2)$, basta considerarmos

$$X = \mu_x + \sigma_x Z_1, \quad Y = \mu_y + \sigma_y Z_2.$$

Na Figura 9.11 temos as curvas de níveis e o gráfico bidimensional obtidos gerando-se 10.000 valores cada uma de duas normais padrões independentes.

Figura 9.11: Distribuição normal padrão bidimensional gerada.



28. Usando um pacote computacional, gere:

(a) 1.000 valores de uma distribuição uniforme bidimensional no quadrado de lado unitário, supondo os componentes independentes;

(b) 1.000 valores de uma normal bi-dimensional (X, Y), com X e Y independentes, $X \sim N(10, 4)$ e $Y \sim N(15, 9)$.

29. Um time de futebol irá disputar 10 partidas num torneio de classificação.
    (a) Supondo que sua chance de vitória em cada jogo é de 60%, simule sua possível campanha.
    (b) Simule agora se é esperado o seguinte desempenho em cada jogo: 50% de vitória, 30% de empate e 20% de derrota.
    (c) Para a situação descrita em (b), simule 12 possíveis campanhas para o time, e estude a variável X = número de pontos obtidos (vitória = 3, empate = 1 e derrota = 0).
    (d) Proponha outros parâmetros para o time e repita a questão (c).

30. Suponha que uma moeda é viciada, de tal sorte que favoreça mais cara do que coroa. Para estimar a probabilidade de cara, você a pode lançar, digamos, 50 vezes.
    (a) Para simular um possível resultado do seu experimento, o que é que seria necessário?
    (b) Supondo que a probabilidade de ocorrer cara é $p = 0{,}6$, qual seria a sua simulação e sua estimativa de p?
    (c) Faça a simulação para 4 outras pessoas e dê suas respectivas estimativas. Alguém acertou o verdadeiro parâmetro?

31. Em uma população 20% das pessoas compram o produto C. Seleciona-se, com reposição, indivíduos dessa população até encontrar um comprador de C. A variável X indica o número de indivíduos entrevistados. Qual a distribuição simulada de X?

32. Uma pesquisa domiciliar irá entrevistar todos os moradores do domicílio e a distribuição do número de moradores por domicílio encontra-se abaixo. Será usada uma amostra de 5 domicílios:

Nº de moradores	Porcentagem
1	5
2	12
3	20
4	23
5	18
6	10
7	8
8	4

    (a) Simule 100 amostras de tamanho 5.
    (b) Considere X = nº médio de pessoas por amostra. Qual a distribuição de freqüência empírica de X?
    (c) Construa a distribuição de $\overline{X}$ = nº médio de pessoas por amostra.
    (d) Encontre para a população o valor μ = nº médio de pessoas, e construa a distribuição empírica de $\overline{X} - \mu$. Como pode ser interpretada essa distribuição?
    (e) Se o entrevistador recebe 2 u.m. por pessoa entrevistada, usando o resultado (b), qual a probabilidade de uma amostra custar mais de 12 u.m.?

33. A altura X c
variância σ
(a) Proponh
homens
(b) Com os
lisando
(c) Gere ur
amostra
populaç
(d) Como v
Dê exem

33. A altura X das pessoas segue aproximadamente uma curva normal com média μ e variância σ^2.

   (a) Proponha dois valores realísticos para μ e σ, e gere 10 alturas de uma população de homens. Calcule a média e o desvio padrão desta população.

   (b) Com os mesmos parâmetros gere uma outra amostra de 10 alturas. Olhando e analisando as duas amostras elas parecem vir de populações distintas?

   (c) Gere uma amostra de 10 alturas de uma população feminina. Compare com a amostra obtida em *(a)*, e diga se é possível afirmar que as duas amostras vêm de populações distintas.

   (d) Como você acha que os parâmetros influenciam para diferenciar bem as amostras? Dê exemplos.