Variáveis aleatórias discretas

Distribuição Binomial ou Modelo Binomial

Qualquer experiência aleatória com as caraterísticas seguintes pode ser representada pelo modelo binomial.

- É formada por “n” provas iguais.
- Em cada prova da experiência só são possíveis dois resultados sucesso ou insucesso.
- O resultado de cada prova é independente dos resultados das restantes provas.
- A probabilidade do sucesso que se representa por p, não varia de uma prova para outra.

À variável aleatória X que representa o número de sucessos nas “n” provas descritas anteriormente chama-se variável aleatória com distribuição binomial de paramêtros n e p e representa-se por

X ~ B(n,p).

Tarefa 77

Considera a variável aleatória:

X: “Número de meninas numa família com dois filhos”.

a) Quais ao valores que a variável pode tomar?

b) Define a distribuição de probabilidade.

Tarefa 78

Uma variável aleatória X tem a seguinte distribuição de probabilidade:

x_i	$P(X = x_i)$
a	2b
3	b
5	b

O valor médio desta variável aleatória é igual a 2. Qual é o valor de a?

Nota

$X \sim B(n; p)$

A função massa de probabilidade, da v.a. X, é dada por:

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1,\cdots n$$

Com $\binom{n}{x} = \dfrac{n!}{(n-x)!\,x!}$ e

n! = 1x2x3x...xn

Exemplos:

1. Pretendemos estudar o número de versos que saem numa moeda equilibrada, em 100 lançamentos.

$$X \sim B(n = 100; p = \frac{1}{2})$$

2. Queremos estudar o número Y de vezes que sai um múltiplo de 3 em 200 lançamentos de um dado perfeito com as faces numeradas de 1 a 6.

$$Y \sim B(n = 200; p = \frac{1}{3})$$

3. Queremos estudar o número de lâmpadas defeituosas, num lote de 500, fabricadas por uma máquina em que 5% das lâmpadas produzidas são defeituosas, admitindo que os acontecimentos são independentes $Z \sim B(n = 500; p = 0{,}05)$

A utilização da fórmula da distribuição Binomial evidenciada na nota ao lado origina cálculos trabalhosos e monótonos mesmo para valores relativamente pequenos de n e está fora do âmbito do programa.

Para simplificar esta tarefa estão disponíveis tabelas (anexos no fim do livro) que permitem obter as probabilidades associadas a *p* sucessos.

Também existem máquinas de calcular que nos permitem determinar os valores estabelecidos.

Exemplo:

Sabe-se que numa determinada associação 60% dos associados votaram a favor do presidente eleito, 10% votaram contra e 30% abstiveram-se. Determina a probabilidade de, num grupo de 10 associados, escolhidos ao acaso, quatro terem:

a) Votado a favor

Seja X a variável aleatória definida como:

X: “ Número de associados do grupo que votaram a favor”

$$X \sim B(10;0,6)$$

$$P(X=4)=0{,}1115=11{,}15\%$$

Esta representada uma parte da tabela onde se obtém o valor.

		p								
n	***x***	0,05	0,1	0,15	0,2	0,25	0,3	0,5	0,55	0,6
10	4	0,0010	0,0112	0,0401	0,0881	0,1460	0,2001	0,2051	0,1596	**0,1115**

b) Votado

Seja Y a variável aleatória definida como:

Y: “ Número de associados do grupo que votaram”

$$Y \sim B(10;0,7)$$

$$P(Y=4)=0{,}0368=3{,}68\%$$

Propriedades:

Se X é uma variável aleatória com distribuição Binomial, B(n, p), então,

i) o valor médio de X é E(X)=μ = np;

ii) a variância de X é Var(X) = np(1− p).

Tarefa 79

Um exame tem 20 questões de escolha múltipla, cada uma com quatro respostas possíveis.

Um aluno respondeu a todas as perguntas ao acaso.

Seja X a variável aleatória correspondente ao número de respostas que o aluno acertou.

Justifica que X é uma variável aleatória com distribuição binomial e indica os parâmetros n e p .

Tarefa 80

Num saco há quatro bolas vermelhas e duas azuis indistinguíveis ao tato.

Considera a seguinte experiência: “tirar uma bola do saco, tomar nota da cor e repor a bola”.

Determina a probabilidade de repetindo a experiência 10 vezes, tirar duas e só duas bolas azuis.

Exemplo:

Consideramos que a probabilidade de umas calças de ganga sair com defeito da linha de produção de uma determinada fabrica é de 5%.

a) Determina o número de calças defeituosas esperado num lote de 1000 calças

X: " Nº de calças defeituosas"

n=1000

p=0,05

$$E[X] = np = 1000 \times 0,05 = 50$$

O número esperado de calças defeituosas é 50.

b) Determina o desvio padrão da distribuição.

Primeiro determinamos a variância.

$$Var\left[X\right] = \sigma^2 = np(1-p) = 47,5$$

O desvio é $\sigma = \sqrt{47,5} = 6,89$

Tarefa 81

A probabilidade de uma peça sair defeituosa de uma determinada fábrica é 8%. Determina:

a) O número de peças defeituosas esperado num lote de 500 peças.

b) O desvio padrão da distribuição.

Modelo de Poisson

A distribuição de Poisson, permite descrever uma grande variedade de situações com aplicações em muitas áreas do conhecimento.

Para que uma experiência aleatória possa ser explicada usando o Modelo de Poisson tem que verificar as seguintes condições:

- O número de ocorrências em intervalos de tempo não sobrepostos são variáveis aleatórias independentes.
- A probabilidade de um certo número de ocorrências se verificar é a mesma para intervalos da mesma dimensão, isto é, aquela probabilidade depende apenas da amplitude do intervalo e não da posição em que se situa esse intervalo.
- A probabilidade de se registarem duas, ou mais, ocorrências num intervalo suficientemente pequeno é irrelevante, quando comparada com a probabilidade e se verificar apenas uma ocorrência.

Alguns exemplos de situações que se adequam a uma distribuição de Poisson, são:

- nº de chamadas telefónicas que chegam, em certo período de tempo, a uma central telefónica;
- nº de doentes que chegam a um hospital, num certo intervalo de tempo;
- nº de avarias que ocorrem numa máquina, num certo intervalo de tempo.

Uma variável aleatória X cujo domínio é N_0 que admite distribuição de Poisson escreve-se $X \sim P(\lambda)$ onde λ é o parâmetro caracterizador desta distribuição.

Da mesma forma que no Modelo Binomial a utilização da fórmula da distribuição de Poisson origina cálculos trabalhosos e monótonos e está fora do âmbito do programa.

Para simplificar esta tarefa estão disponíveis tabelas (anexos ao fim do livro) que permitem obter as probabilidades associadas a *p* sucessos.

Também existem máquinas de calcular que nos permitem determinar os valores estabelecidos.

Exemplo:

Admita que o número de camiões que, por hora, atravessam uma ponte segue uma distribuição de Poisson com parâmetro igual a 8.

Qual a probabilidade de que, numa hora, exatamente 4 camiões atravessem a ponte?

X: "Nº de camiões que por hora travessam a ponte"

$$X \sim P(8)$$

	λ									
x	7,1	7,2	7,3	7,4	7,5	7,6	7,7	7,8	7,9	8
4	0,0874	0,0836	0,0799	0,0764	0,0729	0,0696	0,0663	0,0632	0,0602	0,0573

$$P(X = 4) = 0{,}0573$$

Propriedades:

Se, $X - P(\lambda)$ então,

i) o valor médio de X é E(X) = λ;

ii) a variância de X é Var(X) = λ.

Tarefa 82

O número de reclamações que uma lavandaria recebe por dia é uma variável aleatória seguindo uma distribuição de Poisson com parâmetro 2,5.

Determina a probabilidade da lavandaria num dia:

a) Não receber reclamações.

b) Receber 2 reclamações.

Variáveis aleatórias contínuas

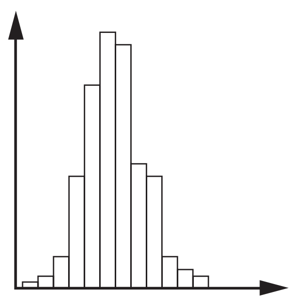
Distribuição Normal

A distribuição Normal é sem dúvida uma das distribuições mais utilizadas na Estatística. São inúmeras as variáveis aleatórias que descrevem fenómenos físicos ou caraterísticas humanas (peso, altura, etc.) que seguem distribuição Normal.

Consideremos que queríamos estudar o peso dos alunos de ensino secundário em Timor-Leste.

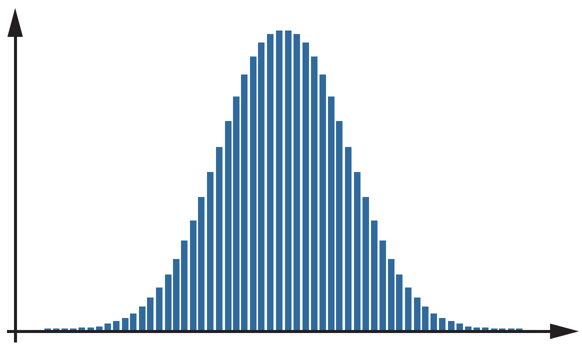
Seja X a variável aleatória que representa o peso de um aluno, escolhido ao acaso de entre os alunos timorenses, do ensino secundário. A variável aleatória X é contínua.

Para estudar a população recolhemos uma amostra e representamos os dados num histograma



Considerando cada vez mais observações e consequentemente mais classes para o histograma, este, no limite, vai aproximar-se de uma área limitada por uma curva representativa da chamada função **densidade de probabilidade**.

A função densidade que se representa com a forma de sino é para a população, o equivalente ao histograma para a amostra.



Tarefa 83

O número de chamadas que chegam num período de cinco minutos à central telefónica de uma empresa é uma variável aleatória com distribuição de Poisson de parâmetro $\lambda = 10$.

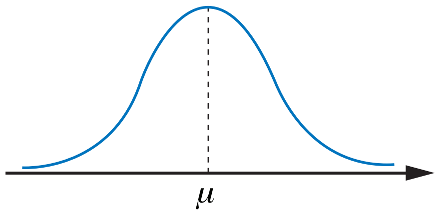
a) Calcula a probabilidade de num período de cinco minutos, chegarem exatamente 8 chamadas.

b) Calcula a probabilidade de num período de cinco minutos, não chegar nenhuma chamada.

c) Determina o número médio de chamadas que chegam num período de cinco minutos à empresa.

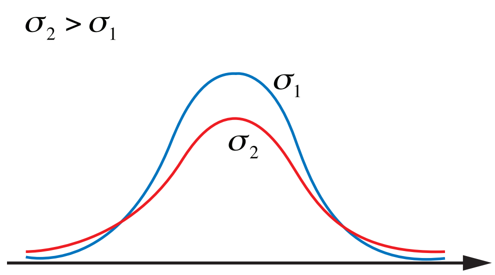
Caraterísticas da curva normal

É simétrica em relação ao valor médio μ da variável.

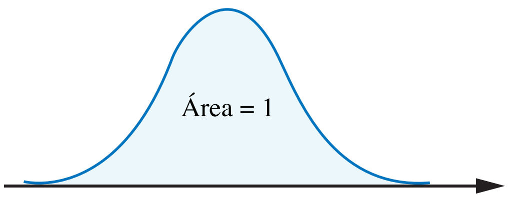


Tem um máximo para $x=\mu$

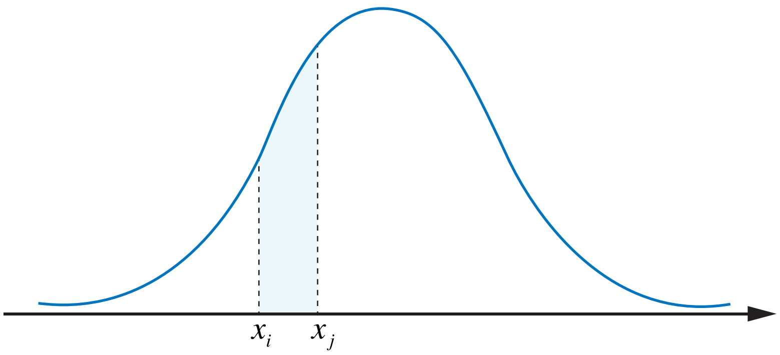
Quanto maior for o desvio padrão, σ, mais achatada é a curva.



A área compreendida entre acurva e o eixo Ox é igual a 1.



A probabilidade de que a variável tome valores no intervalo $[x_i, x_j]$ é igual à área definida pelo eixo Ox, pelo gráfico da função densidade e pelas retas $x = x_i$ e $x = x_j$.



Referência Histórica

A distribuição normal foi reconhecida pela primeira vez pelo francês **Abraham de Moivre (1667-1754)**.

Posteriormente, **Carl Friedrich Gauss (1777-1855)** elaborou desenvolvimentos mais profundos e formulou a equação da curva; é por isso que também é conhecida, como o "Sino de Gauss".

Tarefa 84

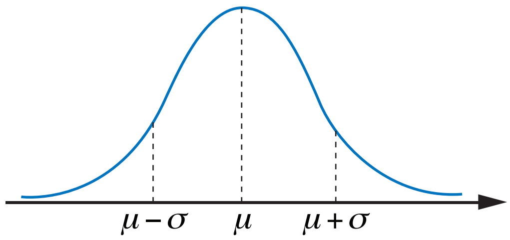
As massas dos parafusos produzidos por uma fábrica distribuem-se normalmente com μ=56g e σ=6g.

Determina a percentagem de parafusos cuja massa:

a) Pertence ao intervalo]50,56[em gramas.

b) É inferior a 56 gramas.

A concavidade da curva muda de sentido para $x_1 = \mu - \sigma$ e $x_2 = \mu + \sigma$.



Tarefa 85

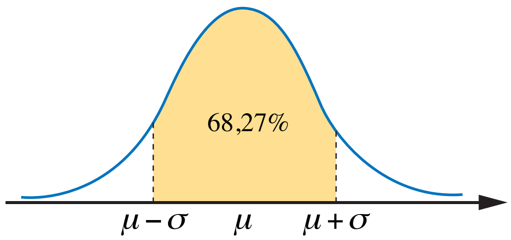
A distribuição dos pesos dos alunos de uma escola segue uma distribuição normal com μ=64 e σ=10 em Kg. Determina a percentagem de alunos que pesam:

a) Entre 54 Kg e 74 Kg.

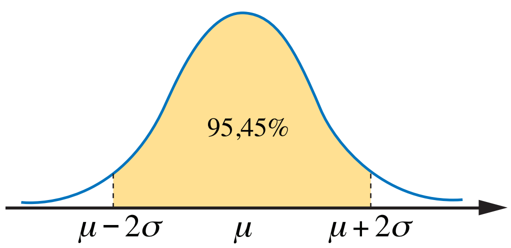
b) Entre 34 Kg e 94 Kg.

O eixo das abcissas é assíntota da curva.

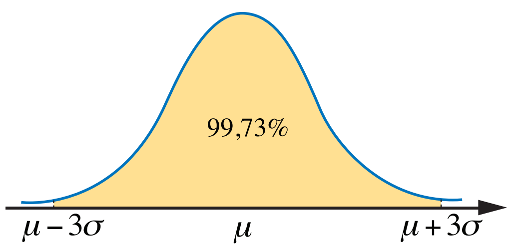
A área abaixo da curva distribui-se em intervalos da seguinte forma:



$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0{,}6827$$



$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0{,}6827$$



$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0{,}9973$$

Uma distribuição normal fica totalmente caraterizada se conhecemos o valor médio μ e o desvio padrão σ. Representa-se por N(μ,σ).

Exemplo:

Considera uma distribuição normal relativa ao peso, em Kg , das crianças de um infantário, de valor médio 11kg e desvio padrão 1kg.

X:" Peso das crianças de um infantário"

$$X \sim N(\mathbf{1} \quad ,1)$$

Pelas propriedades anteriores verifica-se:

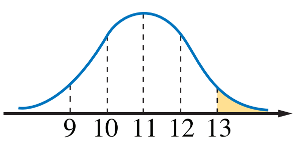
68,27% da população têm peso entre 10Kg e 12Kg.

95,45% da população têm peso entre 9Kg e 13 Kg.

99,73% da população têm peso entre 8Kg e 14Kg.

Tendo em conta que a distribuição é simétrica, podem obter-se outras probabilidades, por exemplo:

Determinar $P(X > 13)$



$$P(9 < X < 13) \approx 0,9545$$

$$P(X < 9) \quad + \quad P(X > 13) \approx 1 - 0,9545 \approx 0,0455$$

$$P(X > 13) = \frac{1}{2} \times 0,0455 \approx 0,0228$$