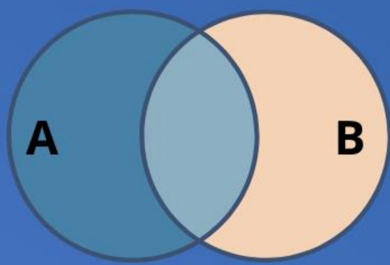
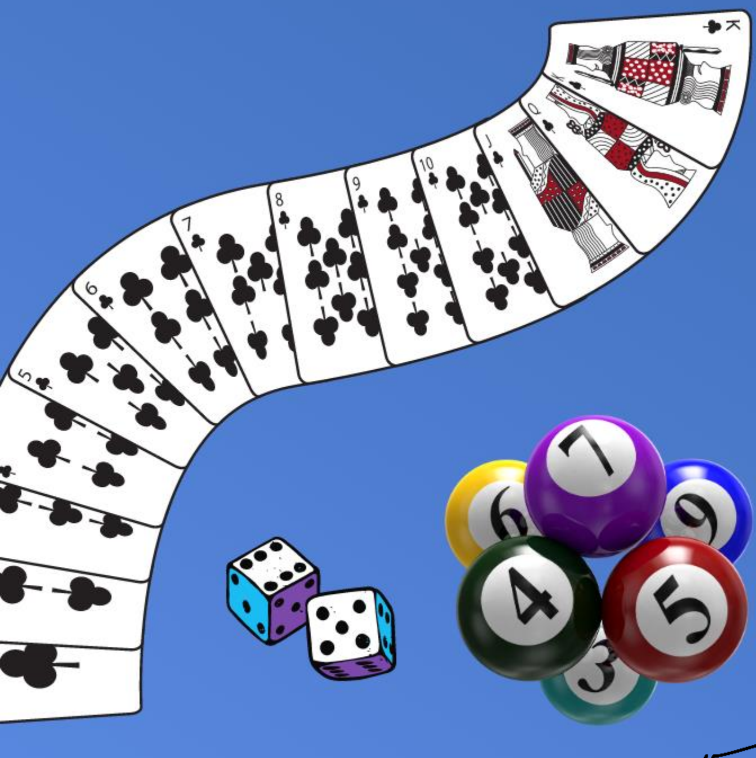
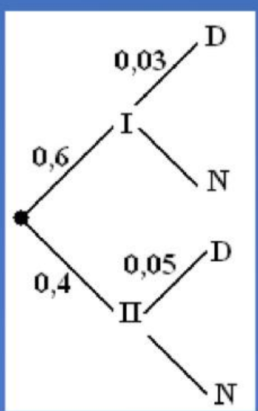
JANILSON PINHEIRO DE ASSIS
ISAAC REINALDO PINHEIRO DE LIMA
JOELMA DE ASSIS FRANÇA
ROBERTO PEQUENO DE SOUSA
ROBSON PEQUENO DE SOUSA
TELDE NATEL CUSTÓDIO
WALTER MARTINS RODRIGUES
JOAQUIM ODILON PEREIRA

ELEMENTOS DE PROBABILIDADE I







Pantanal Editora

2023

JANILSON PINHEIRO DE ASSIS
ISAAC REINALDO PINHEIRO DE LIMA
JOELMA DE ASSIS FRANÇA
ROBERTO PEQUENO DE SOUSA
ROBSON PEQUENO DE SOUSA
TELDE NATEL CUSTÓDIO
WALTER MARTINS RODRIGUES
JOAQUIM ODILON PEREIRA

ELEMENTOS DE PROBABILIDADE I

Pantanal Editora

2023

Copyright© Pantanal Editora
Editor Chefe: Prof. Dr. Alan Mario Zuffo
Editores Executivos: Prof. Dr. Jorge González Aguilera e Prof. Dr. Bruno Rodrigues de Oliveira

Diagramação: A editora. **Diagramação e Arte:** A editora. **Imagens de capa e contracapa:** Canva.com. **Revisão:** O(s) autor(es), organizador(es) e a editora.

Ficha Catalográfica

Catalogação na publicação
Elaborada por Bibliotecária Janaina Ramos – CRB-8/9166

E38

Elementos de probabilidade 1 / Janilson Pinheiro de Assis, Isaac Reinaldo Pinheiro de Lima, Joelma de Assis França, et al. – Nova Xavantina-MT: Pantanal, 2023. 91p. ; il.

Outros autores: Roberto Pequeno de Sousa, Robson Pequeno de Sousa, Telde Natel Custódio, Walter Martins Rodrigues, Joaquim Odilon Pereira.

Livro em PDF

ISBN 978-65-81460-88-4
DOI https://doi.org/10.46420/9786581460884

1. Probabilidades. I. Assis, Janilson Pinheiro de. II. Lima, Isaac Reinaldo Pinheiro de. III. França, Joelma de Assis. IV. Título.

CDD 519.2

Índice para catálogo sistemático

I. Probabilidades

Pantanal Editora

Rua Abaete, 83, Sala B, Centro. CEP: 78690-000.
Nova Xavantina – Mato Grosso – Brasil.
Telefone (66) 99682-4165 (Whatsapp).
https://www.editorapantanal.com.br
contato@editorapantanal.com.br

PREFÁCIO

A teoria das probabilidades e suas leis, propriedades, regras, teoremas, axiomas e corolários é considerada do Ponto de Vista Moderno como difundida no início do Século Vinte e no decorrer deste cem últimos anos, e desempenha uma relevante influência e destaque em todos as ramos das atividades humanas, na ciência e no cotidiano das pessoas e empresas, do ponto de vista cientifico ela é parte básica dos argumentos, abstrações, induções bem como das conclusões obtidas de forma segura, confiável, robusta e com elevada precisão e exatidão, seja em aplicações nas ciências físicas, biológicas e sociais, etc.. Sendo que no decorrer deste período de tempo, ou seja, primeira metade do século vinte, esta teoria foi incorporada e utilizada para dá suporte a criação de novas ferramentas de análise estatística, de modelos, de uso frequente na pesquisa científica e na estatística matemática em geral, assim como na construção de modelos da inferência estatística, bem como no uso do estabelecimento da estatística multivariada, da análise de sobrevivência, dos modelos lineares generalizados, na geoestatística e na estatística Bayesiana. Sendo assim, toda esta demanda e importância faz com que a teoria probabilística seja componente do conteúdo programático da quase totalidade dos programas pedagógicos nos cursos de graduação e pós-graduação, além de cursos básicos e até mesmo no ensino fundamental, sendo amplamente ensinada na universidade nos cursos de graduação e de pós graduação, nos colégios, escolas, institutos, cursos técnicos, em cursos preparatórios e em treinamento, em especializações e aperfeiçoamentos, dentre outros. A teoria antiga e a moderna são pré requisitos na teoria geral de probabilidades e na inferência estatística podendo ser apresentada em vários níveis no contexto matemático. A origem da probabilidade é ligada ao cálculo das possibilidades de ocorrência de determinados resultados em jogos de azar, cartas, dados, moedas, roletas, casinos, bingos, rifas, casas de apostas, loterias, etc., propiciando um grande apelo intuitivo.

Este livro foi escrito com a finalidade de complementar o texto da disciplina estatística para alunos de qualquer área, especialmente voltado para estudantes de um bacharelado em Ciências Exatas, Ciências Físicas , Ciências da Terra, Ciências Sociais, Ciências Humanas, Economia, Administração, Engenharia, Medicina, Genética, Ecologia, Ciências Biológicas e Ciências Agrárias, Dentre outras.

A ideia de elaboração desse livro surgiu da experiência de mais de 30 anos lecionando a disciplina de estatística na ESAM e atualmente Universidade Federal Rural do Semi-Árido (UFERSA), o qual teve origem das inúmeras notas de aulas elaboradas e distribuídas ao longo desse período.

Essa obra exige do leitor conhecimentos elementares de matemática e estatística, pois procurou-se sempre que possível simplificar as demonstrações das propriedades, leis e teoremas, Axiomas, Corolários componentes da teoria, ou seja não utiliza formalismo matemático que vá além de noções básicas de cálculo. O capítulo trata do conceitos fundamentais, propriedades e medidas características das variáveis aleatórias e distribuições de probabilidades unidimensionais e bidimensionais, tanto de natureza discreta ou descontínua, como de natureza contínua, inclusive trata sobre covariância e correlação.

No final do capítulo apresentamos uma lista de exercícios propostos, os quais tem a finalidade de fixar no leitor os fundamentos teóricos da probabilidade mediante a resolução de exemplos nas mais diferentes situações. No final do livro existe o apêndice A e o B contendo respectivamente o alfabeto grego com todas as letras em maiúsculo e minúsculo, além de sua pronúncia em português e a tabela com os valores do número "$e^{-\lambda}$" que é a base dos logaritmos neperianos (e = 2,1828) elevado ao valor negativo da média da distribuição de Poisson no intervalo contínuo o qual é o valor lambda, ou taxa média de Poisson que é o valor de λ.

Finalmente o autor agradece antecipadamente a todos aqueles que se manifestarem, emitindo sugestões para novas edições deste livro, críticas e correções, pois todas serão bem recebidas e só virão a contribuir para o aprimoramento e aperfeiçoamento dessa obra.

Peço desculpas antecipadamente pela presença de eventuais erros de qualquer natureza, pois todos são de inteira responsabilidade do autor.

Os Autores

Mossoró-RN, Brasil Junho de 2023

SUMÁRIO

APRESENTAÇÃO

A origem da moderna teoria das probabilidades é atribuída a dois famosos matemáticos franceses, Blaise Pascal (1623 - 1662) e Pierre Fermat (1601 - 1665), que durante o século XVII publicou e apresentou conceitos primitivos do que seria hoje a teoria moderna das probabilidades. Uma série de problemas, muitos dos quais relativos a "jogos de azar", levaram a uma troca proveitosa de correspondência entre Pascal e Fermat, na qual, os princípios fundamentais da Teoria das Probabilidades foram formulados.

Pouco tempo depois, em 1657, o holandês Christian Huygens (1629 – 1695) publicou o primeiro livro de probabilidades intitulado De Ratiociniis in Ludo Aleae. As maiores contribuições deste período devem-se a Bernoulli (1654 - 1705) e a de Moivre (1667 - 1754).

Em 1812 Pierre de Laplace (1749-1827) introduziu um conjunto de novas idéias e técnicas no seu livro Theore Analytique des Probabilites e pela primeira vez a teoria das probabilidades deixou o campo dos "jogos de azar" e foi aplicada a muitos problemas práticos e científicos. A teoria dos erros, o cálculo atuarial e a mecânica estatística são exemplos de algumas importantes aplicações da teoria das probabilidades desenvolvidas no século XIX.

Como tantos outros campos da matemática, o desenvolvimento da teoria das probabilidades tem sido estimulado pela variedade das suas aplicações. E simultaneamente, cada avanço da teoria tem permitido o alargamento da sua esfera de influência. A estatística matemática é um dos importantes ramos da teoria das probabilidades; outras aplicações ocorrem em campos tão diversos como a genética, a engenharia, a economia, a medicina, as ciências sociais, etc.

A teoria da probabilidade é o ramo da matemática que estuda os experimentos aleatórios, estocásticos ou casuais, exprimindo a chance de ocorrência de eventos aleatórios associados aos espaços amostrais desses ensaios, nas mais diferentes áreas das ciências físicas, biológicas e sociais associadas ao conhecimento humano, como por exemplo a química, a física, a biologia, a agricultura, a medicina, a genética, a atuaria, o setor de seguros, finanças e economia, astronomia, engenharia, meteorologia, a área de tecnologia, pesquisa de opinião pública, etc.

Essa parte da matemática constitui a base teoria para a realização da chamada inferência estatística ou estatística indutiva ou ainda estatística analítica, haja visto que qualquer afirmação que se faça sobre a população ou universo estatístico é um raciocínio incerto, e essa incerteza é medida por um número denominado probabilidade.

A teoria de probabilidade é um importante ramo da matemática pura, que teve um começo bastante modesto. Suas raízes vêm de uma simples teoria matemática dos jogos de azar iniciada em 1654, quando o jogador De Mere propôs ao matemático Pascal suas famosas perguntas a respeito dos jogos de azar. Desde então, essa teoria percorreu um vasto caminho, inclusive o conceito passou por diversas fases. Igualmente, o campo de aplicação da teoria sofreu ampliações notáveis, como uma teoria que se ocupava com os jogos de azar passou a constituir o fundamento da inferência estatística.

Nas ciências físicas, biológicas e sociais as probabilidades existem em grande quantidade. Sempre existiram. Sua abundância e importância advêm dos elementos significativos não controláveis presentes nos sistemas biológico, econômico e social a que as atividades humanas estão associadas. Essa falta de controle sobre importantes variáveis de um sistema, que foge ao domínio do pesquisador exprime incerteza: Prevalecerá esta ou aquela condição? Por sua vez, incerteza exprime probabilidade: Qual a possibilidade de que prevaleça esta ou aquela condição?. A partir da probabilidade, combinada com o poder de controle parcial do sistema, tem-se que escolher mediante a tomada de decisão: Dadas as possibilidades e consequências desta ou daquela ocorrência de um evento aleatório, como o pesquisador deve atuar? Se não fosse estes riscos, a escolha não passaria de rotina ou de resultados determinísticos e não casuais, aleatórios ou estocásticos. Neste caso, não ocorreria a tomada de decisão, pois se a incerteza está ausente do sistema, o conhecimento perfeito dominaria os resultados do ensaio, sendo assim, estaria ausente o desafio da escolha. Daí a importância do apoio do Espaço Amostral ou Espaço de Resultados (S; Ω), como Resposta dos Possíveis Resultados da Experiência Aleatória.

A definição Clássica ou a Priori de probabilidade dada por Laplace como o Quociente ou Razão do número de casos favoráveis a ocorrência do Evento ou Acontecimento sobre o número de casos Totais ou Possíveis do Espaço Amostral foi à primeira definição formal de probabilidade, e apareceu pela primeira vez em forma clara na obra Líber de Ludo Aleae de Jerônimo Cardano (1501-1576).

Os experimentos que repetidos sob as mesmas condições produzem resultados geralmente diferentes serão chamados experimentos aleatórios, Casuais ou Estocásticos. Fenômenos aleatórios acontecem Frequentemente no Dias a Dia da vida do Homem. São frequentes perguntas tais como: choverá amanhã? Qual será a temperatura máxima em Mossoró-RN no próximo domingo? Qual será o número de ganhadores da Loteria Esportiva e da Mega Sena? Quantos habitantes terá o Brasil no ano 2030, a Bolsa de valores Sobe nos Próximos dias?, a Cotação do Dólar deve cair em Função dos Boatos do mercado financeiro?

A teoria das probabilidades é o ramo da matemática que cria, desenvolve e em geral pesquisa modelos que podem ser utilizados para estudar experimentos ou fenômenos aleatórios. Sendo utilizados para realizar generalizações de populações a partir de amostras aleatórias em condições de incerteza, com elevada confiabilidade e baixo risco ou erro ou elevada precisão.

O modelo matemático utilizado para estudar um fenômeno aleatório particular varia em sua complexidade matemática, dependendo do fenômeno estudado. Mas todos esses modelos têm ingredientes básicos comuns.

É interessante destacar uma referência a alguns dos mais famosos cientistas que tanto contribuíram para a teoria das probabilidades desde Laplace: são eles Chebyshev, Markov e Kolmogorov.

O que será mostrado nesse livro são as definições básicas, as propriedades, os teoremas, axiomas, corolários e leis em geral, com o objetivo de estudar uma série de fenômenos aleatórios relativamente simples e algumas vezes complexos referentes a ideias e noções que são bastante gerais na pesquisa científica. Sendo necessário e exigido do leitor conhecimentos prévios de Matemática e Estatística.

> A teoria do azar consiste em reduzir todos os acontecimentos do mesmo gênero a um certo número de casos igualmente possíveis, ou seja, tais que estejamos igualmente inseguros sobre sua existência, e em determinar o número de casos favoráveis ao acontecimento cuja probabilidade é buscada. A razão deste número para o de todos os casos possíveis é a medida dessa probabilidade, a qual é portanto uma fração cujo numerador é o número de casos favoráveis e cujo denominador é o número de todos os casos possíveis.
>
> Pierre Simon Laplace (1749-1827)
>
> Astrônomo, Matemático e Físico Francês
>
> Ensaio filosófico sobre as Probabilidades.

Os Autores

Mossoró-RN, Brasil Junho de 2023

VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÃO DE PROBABILIDADE UNIDIMENSIONAIS E MULTIDIMENSIONAIS DISCRETAS (DESCONTINUAS) E CONTÍNUAS

INTRODUÇÃO

Na prática dos ensaios ou experimentos aleatórios e da pesquisa em geral, é muitas vezes, mais racional, interessante, eficiente e elegante associar um número a um evento ou acontecimento aleatório referente a determinado fenômeno, e calcular a probabilidade da ocorrência desse número ou resposta, do que a probabilidade do evento ou ocorrência em si. Como exemplo tem-se num experimento aleatório de um lançamento de uma moeda honesta ou equilibrada que a probabilidade da face cara ficar voltada para a horizontal, isto, a $P(\text{ocorrência de cara}) = P(X = 1) = \frac{1}{2}$, onde o significado do valor de X é número de faces cara num lançamento de uma moeda honesta, é dado como aquele em que X pode assumir o valor 1 se a face que ocorre for cara, e 0 se a face voltada para cima for coroa. Nos próximos tópicos será explicado de maneira mais detalhada o conceito desses valores numéricos assumidos pelos eventos aleatórios o qual denominamos de variável aleatória, senão vejamos.

VARIÁVEL ALEATÓRIA UNIDIMENSIONAL

Seja "S" um espaço amostral associado a um experimento aleatório E. Então uma função X que associa a cada valor "s" de "S" um número real $X(s)$ é definida como variável aleatória unidimensional.

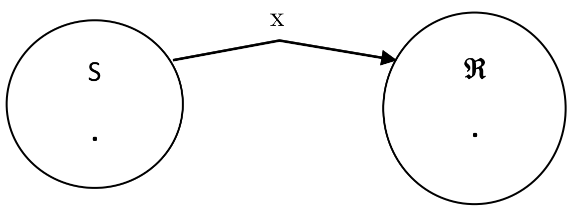


Figura 1. Espaços amostrais e variável aleatória, associados a um experimento aleatório.

TIPOS DE VARIÁVEIS ALEATÓRIAS

Variável Aleatória Discreta (V.A.D.)

É aquela que deve assumir valores em um conjunto finito ou em um conjunto infinito, porém enumerável (resulta de contagem). Seu campo de Definição é o conjunto dos números naturais. Pode – se ainda afirmar que é aquela variável na qual a diferença entre dois possíveis valores desta variável for finito e não nulo

Exemplos:

- Número de sementes por vagem
- Número de insetos por planta
- Número de acidentes, etc.

Variável Aleatória Contínua (V.A.C.)

É aquela que deve assumir valores em um conjunto infinito não enumerável (resulta de mensuração). Seu campo de Definição é o conjunto dos números Reais.

Exemplos:

- Diâmetro do caule de uma árvore em centímetro
- Peso de raízes por planta em gramas
- Rendimento de uma Cultura agrícola em toneladas por hectares, etc.

ESTUDO DA VARIÁVEL ALEATÓRIA DISCRETA (V.A.D.)

Distribuição de probabilidades de uma variável aleatória discreta

Observação: Dados dois conjuntos A e B não vazios, chama-se função (ou aplicação) de A em B, representada por $f: A \to B; y = f(x)$, a qualquer relação binária que associa a cada elemento de A, um único elemento de B. Portanto, para que uma relação de A em B seja uma função, exige-se que a cada $x \in A$ esteja associado um único $y \in B$, podendo entretanto existir $y \in B$ que não esteja associado a nenhum elemento pertencente ao conjunto A. Na notação $y = f(x)$, entendemos que y é imagem (contradomínio ou campo de valores da função) de x pela função f, ou seja: y está associado a x através da função f, sendo x o domínio da função $f(x)$ (campo de definição da função). Exemplos: $f(x) = 4x + 3$; então $f(2) = 4.2 + 3 = 11$ e portanto, 11 é imagem de 2 pela função f; $f(5) = 4.5 + 3 = 23$, portanto 23 é imagem de 5 pela função f, $f(0) = 4.0 + 3 = 3$, etc.

Função de probabilidade ou de massa de probabilidade: É a função que associa a cada valor assumido pela variável aleatória, a probabilidade de evento correspondente, isto é $P(x = x_i) = P(A_i)$, $i = 1, 2, \dots, n$, sendo $0 \leq P(x_i) \leq 1 \ \forall x_i$.

Distribuição de probabilidades: É o conjunto formado pelo valor da variável aleatória e junto com sua função de probabilidade ou valor da probabilidade isto é, $\left(x_i, P(x_i)\right), i = 1, \dots, n$ sendo necessário que $\sum_{i=1}^{n} P(x_i) = 1$.

Função de repartição ou de distribuição ou distribuição de probabilidade acumulada $[\boldsymbol{F(x)}]$

Introdução

A função de probabilidade acumulada que se obtém somando-se as probabilidades de todos os valores da variável aleatória, menores ou iguais a um determinado valor de x_i, define a função de distribuição de variável x_i. Ela é representada por:

Distribuição de probabilidades acumuladas ou função de repartição $[F(x)]$

Definição

$$F(x) = P(x \leq x_i) = \sum_{x=0}^{x \leq x_i} P(x_i) = \begin{cases} 0 & \text{se} \quad -\infty < x < x_1 \ (x < x_1) \\ P(x_1) & \text{se} \quad x_1 \leq x < x_2 \\ P(x_1) + P(x_2) & \text{se} \quad x_2 \leq x < x_3 \\ \vdots & \\ P(x_1) + P(x_2) + \ldots + P(x_n) & \text{se} \quad x_n \leq x < \infty \ (x \geq x_n) \end{cases}$$

sendo $F(x) = P(x \leq x_i) = \sum_{x=0}^{n} P(x_i) = 1$.

O gráfico de $F(x)$ é dado pela Figura 2, abaixo.

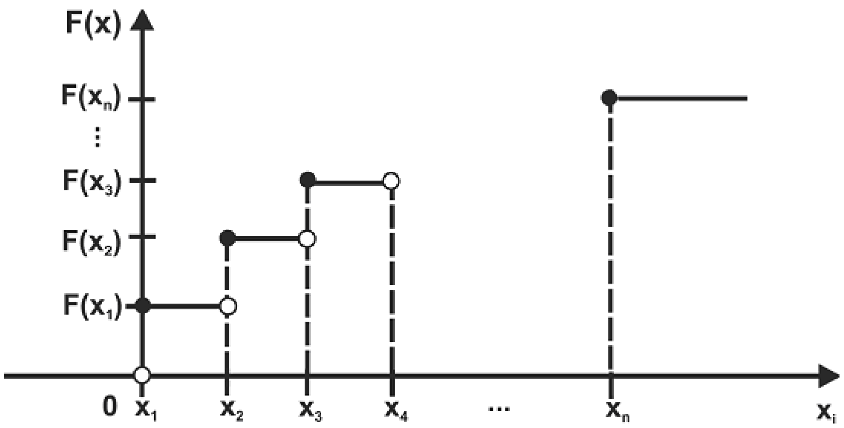


Figura 2. Distribuição de probabilidade acumulativa de uma variável aleatória discreta.

Propriedades de F(x)

i) $0 \leq F(x) \leq 1$

ii) $F(-\infty) = 0$

iii) $F(+\infty) = 1$

iv) $F(x)$ é descontínua nos pontos $x = x_0$ onde $P(x = x_0) \neq 0$

v) $F(x)$é continua à direita dos pontos $x = x_0$, onde $P(x = x_0) \neq 0$

vi) $P(a < x \leq b) = F(b) - F(a)$

vii) $P(a \leq x \leq b) = F(b) - F(a) + P(x = a)$

viii) $P(a < x < b) = F(b) = F(a) - P(x = b)$

ix) $F(x)$ é uma função não decrescente, como $P(a < x \leq b) = F(b) - F(a) \geq 0 \Rightarrow F(b) \geq F(a)$. Logo $F(x)$ é não decrescente.

Exemplo: Seja um experimento aleatório

$$E = \text{plantarduassementesdefeijãomacassareaguardaragerminação};$$

$$S = \{(G,G); (G,\bar{G}); (\bar{G},G); (\bar{G},\bar{G})\},$$

onde ($G = \text{Germinar}$; $\bar{G} = \text{Não germinar}$). E seja X a variável aleatória discreta que representa o número de sementes germinadas:

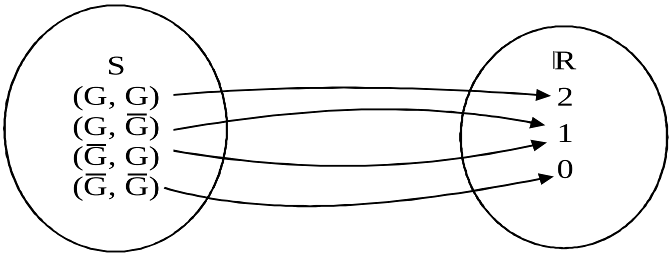
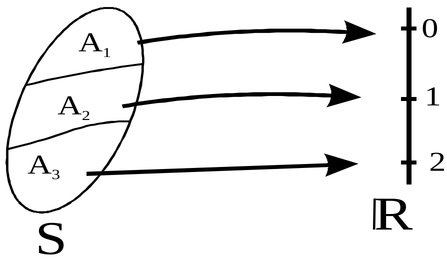


Figura 3. Espaço amostral e conjunto dos números reais, ilustrando a associação de variáveis aleatórias unidimensionais discretas aos diversos resultados ou eventos.

Logo, $X = 0,1,2$

X	Evento correspondente
0	$A_1 = \{(\bar{G},\bar{G})\}$
1	$A_2 = \{(G,\bar{G});(\bar{G},G)\}$
2	$A_3 = \{(G,G)\}$



Espaço amostral e conjunto dos números reais, ilustrando a associação de variáveis aleatórias unidimensionais discretas ao campo de definição da variável aleatória para os diversos resultados ou eventos do experimento aleatório.

A distribuição de probabilidade é então dada por:

TABELA 1. DISTRIBUIÇÃO DE PROBABILIDADE DO NÚMERO DE SEMENTES DE FEIJÃO GERMINADAS

	X_i	$P(X_i)$
$P(x=0)=P(A_1)=\frac{1}{4}$	0	1/4
$P(x=1)=P(A_2)=\frac{2}{4}$	1	2/4
$P(x=2)=P(A_3)=\frac{1}{4}$	2	1/4
	Soma	1,0

E o gráfico da função de probabilidade é mostrada na Figura 4 abaixo.

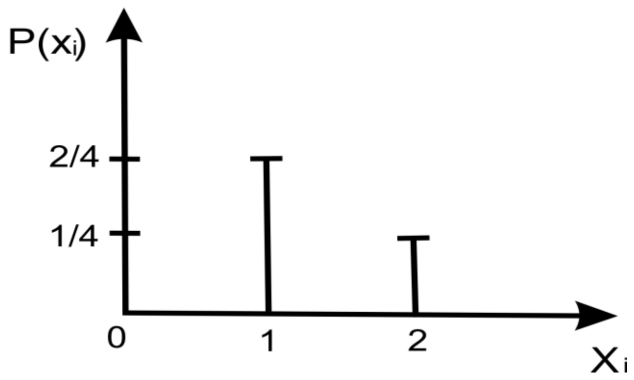


Figura 4. Gráfico em hastes ou bastão da distribuição de probabilidade discreta do número de sementes de feijão germinadas.

A função de distribuição de x ou função de repartição [F(x)] é dada por:

$$F(x) = P(x \leq x_i) = \sum_{x_i \leq x} P(X_i) = \sum_{x=0}^{x \leq x_i} P(X_i)$$

Temos então,

$$F(0) = P(x \leq 0) = P(x=0) = \frac{1}{4}$$

$$F(1) = P(x \leq 1) = P(x=0) + P(x=1) = \frac{1}{4} + \frac{2}{4} = \frac{3}{4}$$

$$F(2) = P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$$

$$F(2) = F(1) + P(x=2) = \frac{3}{4} + \frac{1}{4} = 1$$

Pode-se ainda calcular também: Neste caso biologicamente absurdo mas matematicamente normal.

$$F(1{,}34) = P(x \leq 1{,}34) = P(x \leq 1)$$

$$P(x \leq 1) = F(1) = \frac{3}{4}$$

$$F(-3) = P(x \leq -3) = 0$$

Com esses resultados podemos escrever

$$F(x) = \begin{cases} 0 & \text{se} \quad x < 0 \\ 1/4 & \text{se} \quad 0 \leq x < 1 \\ 3/4 & \text{se} \quad 1 \leq x < 2 \\ 1 & \text{se} \quad x \geq 2 \end{cases}$$

E a função de repartição é dada por:

TABELA 2. DISTRIBUIÇÃO DE PROBABILIDADE ACUMULADA OU FUNÇÃO DE REPARTIÇÃO DA VARIÁVEL ALEATÓRIA X, NÚMERO DE SEMENTES DE FEIJÃO MACASSAR GERMINADAS.

X_i	$P(X_i)$	$F(X_i)$
0	1/4	1/4
1	2/4	3/4
2	1/4	4/4
Soma	1,0	-

O gráfico de $F(x)$ é dado pela Figura 5, mostrada a seguir.

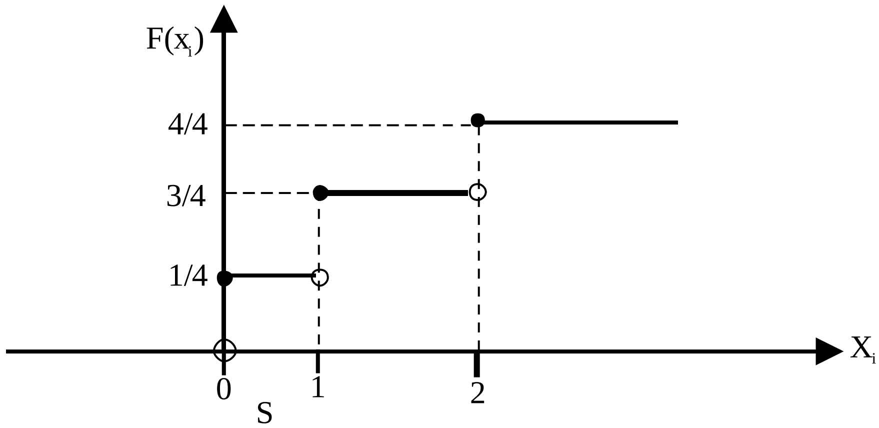


Figura 5. Gráfico em escada representativo da distribuição de probabilidade acumulada da variável aleatória X número de sementes germinadas de feijão macassar.

O domínio de $F(x)$ é ℝ e o contra domínio é o conjunto $\{1/4, 3/4, 4/4\}$.

ESTUDO DA VARIÁVEL ALEATÓRIA CONTÍNUA [V.A.C]

Definição 1: Uma variável aleatória X é contínua se existir uma função f, denominada função densidade de probabilidade (f.d.p.) de X, que satisfaça às seguintes condições:

$f(x) \geq 0, \forall x \in \mathbb{R}_x$ (para rodo X em S (campo de definição ou domínio da função)

$$\int_{\mathbb{R}_x} f(x)dx = \int_{-\infty}^{+\infty} f(x)dx = 1$$

Para quaisquer a, b com $-\infty < a < b < +\infty$, teremos,

$$P(a < x < b) = \int_a^b f(x)dx$$

Definição 2: Distribuição de probabilidade de variável aleatória contínua se para uma variável aleatória contínua X se conhece a probabilidade de que assuma valores em cada um dos intervalos de seu campo de variação, de tal maneira que a soma das probabilidades seja igual a 1, dizemos que está definida uma distribuição de probabilidade de X (V.A.C). Para caracterizar a distribuição obviamente teremos de definir uma função, $f(x)$. esta função $f(x)$ recebe o nome de função de densidade de probabilidade da variável aleatória contínua.

A probabilidade de que a variável X assuma um valor dentro do intervalo $a \leq X \leq b$ será dada pela integral $P(a < x < b) = \int_a^b f(x)dx$. Como já foi mencionado para que uma função de variável contínua possa ser considerada como uma função de densidade de probabilidade, ela deve obedecer às seguintes propriedades:

a) $f(x) \geq 0$ para todo X em S (campo de definição ou domínio da função).

b) $\int_a^b f(x)dx = P(a < X \leq b), b > a$

c) $\int_{-\infty}^{+\infty} f(x)dx = 1$

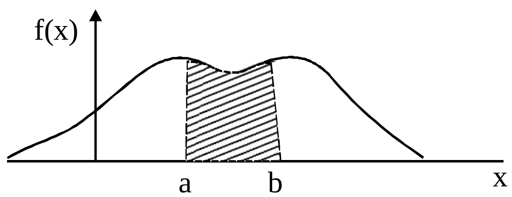


Figura 6. Área ou densidade sob a curva da distribuição densidade de probabilidade $f(X)$ da variável aleatória contínua X.

Em relação aos itens a b e c mencionados anteriormente pode concluir que a representação gráfica da função densidade de probabilidade possui as seguintes propriedades:

(i) Se $f(x)$ é a função de densidade para uma variável aleatória X, podemos representar $Y = f(x)$ graficamente por uma curva como;

(ii) como $f(x) \geq 0$ a curva não pode estar abaixo do eixo-x.

(iii) A área total delimitada pela curva e pelo eixo-x deve ser 1, em virtude da propriedade $\left[\int_{-\infty}^{\infty} f(x)\, dx = 1\right]$.

(iv) Geometricamente a, probabilidade de X estar entre a e b, isto é, $P(a < X < b)$, é então dada pela área sombreada na figura abaixo:

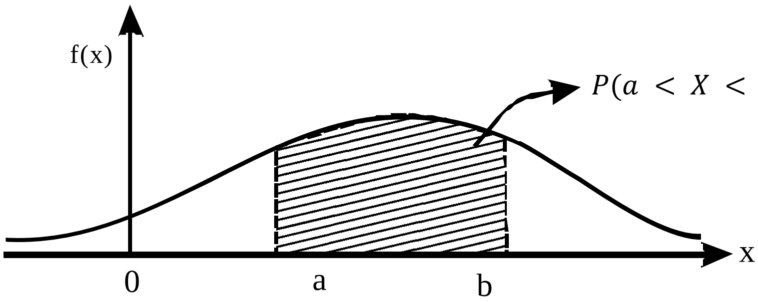


Figura 7. Distribuição contínua de probabilidade.

Observações importantes

i) $P(a < X < b)$ representa a área sob a curva da função densidade de probabilidade entre a e b, isto é: $P(a < X < b) = \int_a^b f(x)dx$.

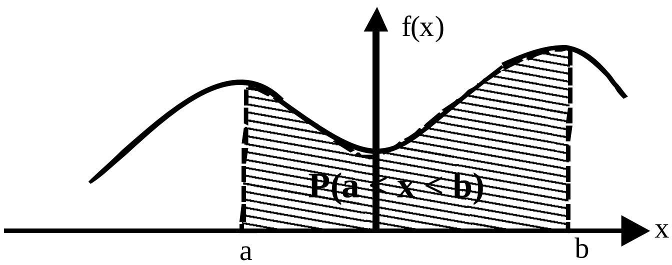


Figura 8. Área, densidade ou probabilidade sob a curva da distribuição densidade de probabilidade.

ii) Para qualquer valor especificado de X, por exemplo, $X = c$, temos:

$$P(x = c) = \int_c^c f(x)dx = F(c) - F(c) = 0$$

Portanto, no caso contínuo, $P(A) = 0 \xrightarrow{\text{não}} A = \varnothing$. Tendo em vista esta observação, podemos escrever:

$$P(c \leq X \leq d) = P(c < X \leq d) = P(c \leq X < d) = P(c < X < d)$$

Sendo assim as probabilidades acima são todas iguais, se X for variável aleatória contínua.

Os eventos impossíveis têm probabilidade igual a 0 (zero), mas, os eventos de probabilidade igual a 0 (zero) não são forçosamente impossíveis.

iii) Dado uma função g, tal que:

iii 1) $g(x) \geq 0, \forall x \in X(W) \text{ou} \forall x \in \mathbb{R}$

iii 2) $\int_{-\infty}^{\infty} g(x)dx = k$, onde k é um número real positivo (não necessariamente igual a 1) $\forall x \in \mathbb{R}$.

Logo, g não é uma função de densidade de probabilidade.

Para tornar g uma função densidade de probabilidade, devemos fazer o seguinte:

Tomar $\frac{g(x)}{k} = f(x)$ implica $g(x) = kf(x), \forall \in X(W) \text{ou} \forall x \in \mathbb{R}$

$$\int_{-\infty}^{\infty} g(x)dx = k \int_{-\infty}^{\infty} f(x)dx = k \text{ implica } \frac{1}{k}\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$$

iv) Se o conjunto de valores de X, estiver contido no intervalo $[a, b]$, então para $x \notin [a, b]$ implica $f(x) = 0$.

v) $f(x)$ para o caso de variável aleatória contínua, não representa probabilidade de coisa alguma, somente quando $f(x)$ for integrada entre 2 limites ela produzirá uma probabilidade que corresponde a área delimitada pela função $f(x)$, eixo dos X e pelas retas $x = a$ e $x = b$, sendo $a < b$.

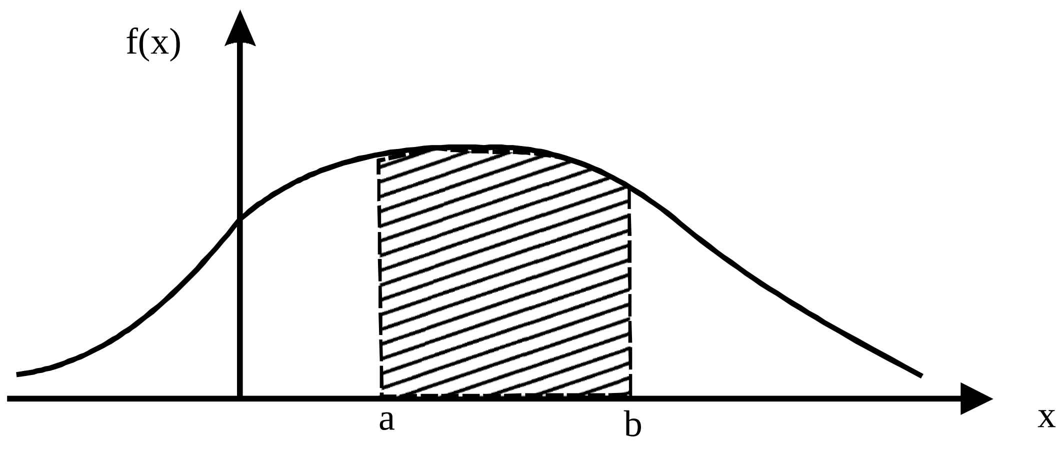


Figura 9. Área ou probabilidade delimitada pela curva, pelo eixo-x e pelas retas $x = a$ e $x = b$. $P(a \leq x \leq b) = \text{áreadaregiãosombreada}$

FUNÇÃO DE REPARTIÇÃO, FUNÇÃO DE DISTRIBUIÇÃO OU DISTRIBUIÇÃO DE PROBABILIDADES ACUMULADAS [F(X)]

A função F(x)

$$\mathrm{F}(x) = \mathrm{P}(-\infty < x < x) = \mathrm{P}(x \leq x) = \int_{-\infty}^{x} dF(x) = \int_{-\infty}^{x} f(t)\,\mathrm{dt}$$

que dará a probabilidade de que X assuma ou tome qualquer valor menor ou igual a x.

$F(x) = P(X \leq x)$ é uma função monotonicamente crescente, que cresce de 0 a 1, e é representada por uma curva como na figura a seguir:

Uma função é monotonicamente crescente quando: Na medida em que procedemos da esquerda para a direita, a função de distribuição permanece a mesma, ou aumenta, tomando valores de 0 a 1. Traduz-se matematicamente este fato dizendo-se que a função é monotonicamente crescente.

Gráficos de F(x):

O gráfico genericamente é

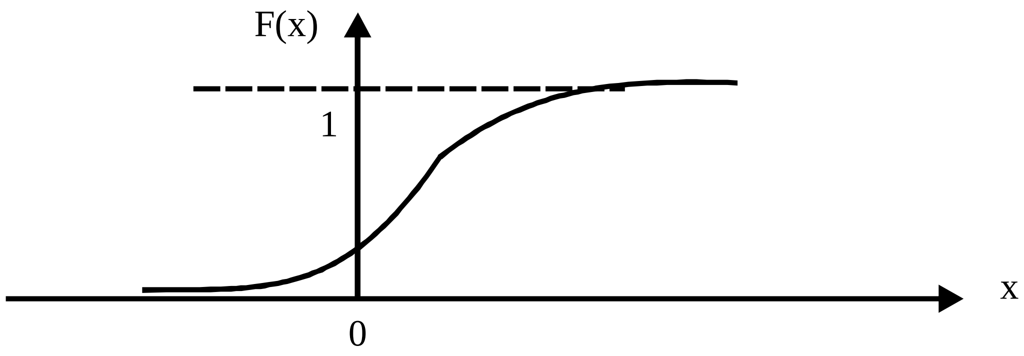


Figura 10. Função de repartição contínua.

E no caso particular de $F(a) = P[-\infty \leq x \leq a]$

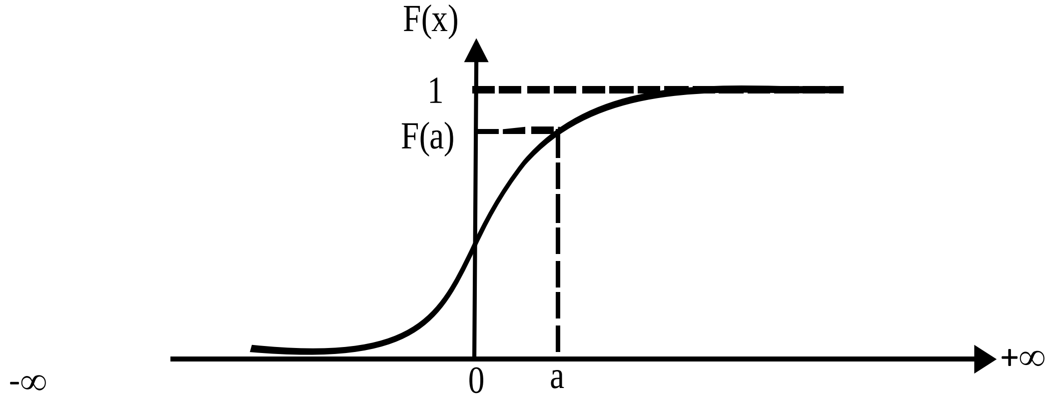


Figura 11. Cota de probabilidade acumulada.

Propriedades da Função de Repartição F(x)

i) $F(-\infty) = 0$

ii) $F(+\infty) = 1$

iii) $F(a) = P[-\infty \leq x \leq a]$

Exercícios de aplicação

Exemplo 1. A variável aleatória contínua x desse exemplo representa o consumo em gramas de ração de crescimento diário, em uma granja avícola. Seja

$$f(x) = \begin{cases} Kx & \text{se } 0 \leq x \leq 1\{ \text{ ou se } x \in [0,1]\} \\ 0 & \text{se } x < 0 \text{ ou } x > 1\{ \text{ ou se } x \notin [0,1]\} \end{cases}$$

Determinar:

i) K a fim de que $f(x)$ seja uma função densidade de probabilidade [f.d.p.].

ii) O gráfico de $f(x)$

iii) $P\left(0 \leq x \leq \frac{1}{2}\right)$

iv) $E(x)$

v) $VAR(x)$

vi) $P\left(x = \frac{1}{2}\right)$

vii) $F(x)$

viii) O gráfico de $F(x)$

ix) $\int_{-\infty}^{\infty} f(x)dx = 1, \int_0^1 Kxdx = 1 \rightarrow K\left(\frac{x^2}{2}\right)\int_0^1 = 1 \therefore K = 2$

x) $f(x) = \begin{cases} 2x & \text{se } 0 \leq x \leq 1 \\ 0 & \text{se x} < 0 \text{ ou x} > 1 \end{cases}$

TABELA 13. DISTRIBUIÇÃO DE PROBABILIDADE CONTÍNUA DA VARIÁVEL ALEATÓRIA X: CONSUMO EM GRAMAS DE RAÇÃO DE CRESCIMENTO DIÁRIO, EM UMA GRANJA AVÍCOLA.

X	$f(x)$
0	0
1/2	1
1	2

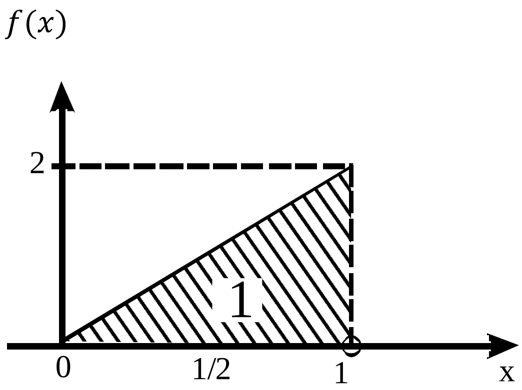


Figura 12. Área ou densidade sob a curva da distribuição densidade de probabilidade da variável aleatória contínua X: consumo em gramas de ração de crescimento diário, em uma granja avícola.

xi) $\mathrm{P}\left(0 \leq x \leq \frac{1}{2}\right) = \int_{0}^{1/2} \;\; 2x dx = (x^2)\Big|_{1/2}^{0} = \frac{1}{4}$

xii) $\mathrm{E}(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{0}^{1} x 2x dx = \int_{0}^{1} 2x^2 dx = 2\left(\frac{x^3}{3}\right)\Big|_{0}^{1} = \frac{2}{3}$

xiii) $\mathrm{E}(x^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{0}^{1} x^2 2x dx = \int_{0}^{1} 2x^3 dx = 2\left(\frac{x^4}{3}\right)\Big|_{0}^{1} = \frac{1}{2} \therefore$

$$\mathrm{VAR}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

xiv) $\mathrm{P}\left(x = \frac{1}{2}\right) = \int_{1/2}^{1/2} \;\; 2x\mathrm{d}x \;=\; 2\left[\frac{x^2}{2}\right]_{1/2}^{1/2} = 2\left[\frac{(1/2)^2}{2} - \frac{(1/2)^2}{2}\right] = 0$

xv) $\mathrm{F(x)} = \int_{-\infty}^{\infty} 0 ds = 0$, para $x \leq 0$

$$\mathrm{F(x)} = \int_{-\infty}^{\infty} f(s) ds = \int_{0}^{x} 2s ds = 2\left(\frac{S^2}{2}\right)\Bigg|_{0}^{x} = S^2|_0^x = x^2 - \; 0^2 = x^2 \text{ para } 0 \; < x < \; 1$$

$$\mathrm{F(x)} = \int_{0}^{1} 2\mathrm{sds} + \int_{1}^{x} 0\mathrm{d}\,s = 2\left(\frac{S^2}{2}\right)\Bigg|_{0}^{x} = S^2|_0^1 = 1^2 - 0^2 = 1, \text{ para } x \geq \; 1$$

Logo, $F(x)$ é dada por:

viii) $\mathrm{F(x)} = \begin{cases} 0 & \text{se} \;\; x \leq 0 \\ x^2 & \text{se} \;\; 0 \; < \; x \; < \; 1 \\ 1 & \text{se} \;\; x \geq 1 \end{cases}$

É importante lembrar a seguinte observação: $\frac{d}{dx} F(x) = 2x$ para $0 \; < \; x \; < \; 1$

$F(X_i)$

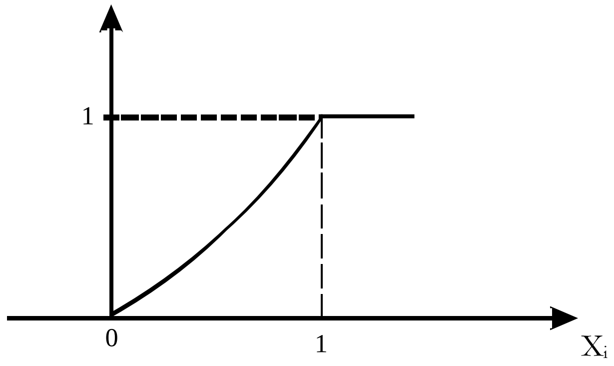


Figura 13. Gráfico da distribuição de probabilidade acumulada da variável aleatória X: consumo em gramas de ração de crescimento diário, em uma granja avícola.

Exemplo 2. Determine a constante C de modo que a função.

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{em caso contrário} \end{cases}$$

i) Seja uma função de densidade de probabilidade.

ii) Calcule $P(1 < x < 2)$

iii) Determine a função de distribuição da variável aleatória.

A resposta:

i) $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \therefore \int_0^3 cx^2\mathrm{d}x = \left.\frac{cx^3}{3}\right]_0^3 = 1$

$9c = 1 \therefore \quad c = \frac{1}{9}$

ii) $\mathrm{P}(1 < x < 2) = \int_1^2 \frac{1}{9}\, x^2\, dx = \left.\frac{x^3}{27}\right]_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$

iii) $\mathrm{F}(x) = \mathrm{P}(x \leq x) = \mathrm{P}(-\infty < x \leq x) = \int_{-\infty}^{x} f(u)du$

Se $x < 0$ então $F(x) = 0$

Se $0 \leq x < 3$ então $\mathrm{F}(x) = \int_0^x f(x)\mathrm{d}x = \int_0^x \frac{1}{9}x^2\mathrm{d}x = \frac{x^3}{27}$

Se $x \geq 3$ então

$$\mathrm{F}(x) = \int_0^3 f(x)\mathrm{d}x + \int_3^x f(x)\mathrm{d}x = \int_0^3 \frac{1}{9}x^2\mathrm{d}x + \int_3^x 0\mathrm{d}x = 1$$

Assim a função de distribuição ou de repartição desejada é:

$$F(x) = \begin{cases} 0 & \text{para} \quad x < 0 \\ \dfrac{x^3}{27} & \text{para} \quad 0 \leq x < 3 \\ 1 & \text{para} \quad x \geq 3 \end{cases}$$

Exemplo 3. Seja X o tempo durante o qual um motor elétrico de um sistema de irrigação é usado em carga máxima, num certo período de tempo em minutos. A função densidade de probabilidade de X é dada por:

$$F(x) = \begin{cases} \dfrac{1}{1500^2}x \text{, se } 0 \leq x < 1500 \\ \dfrac{1}{1500^2}(3000 - x), \text{ se } 1500 \leq x \leq 3000 \end{cases}$$

Calcular $E(x)$, ou seja, o tempo médio em que o equipamento será utilizado em carga máxima.

$$E(x) = \int_0^{1500} \frac{1}{1500^2} x \cdot x\mathrm{d}x + \int_{1500}^{3000} \frac{1}{1500^2}(3000 - x) \cdot x\mathrm{d}x$$

$$\boldsymbol{E}(x) = \frac{1}{1500^2}\left\{\left(\frac{x^3}{3}\right)\Big|_0^{1500} + (1500x^2 - \frac{x^3}{3}\Big|_{1500}^{3000}\right\} = 1500 \text{ minutos}$$

Graficamente temos:

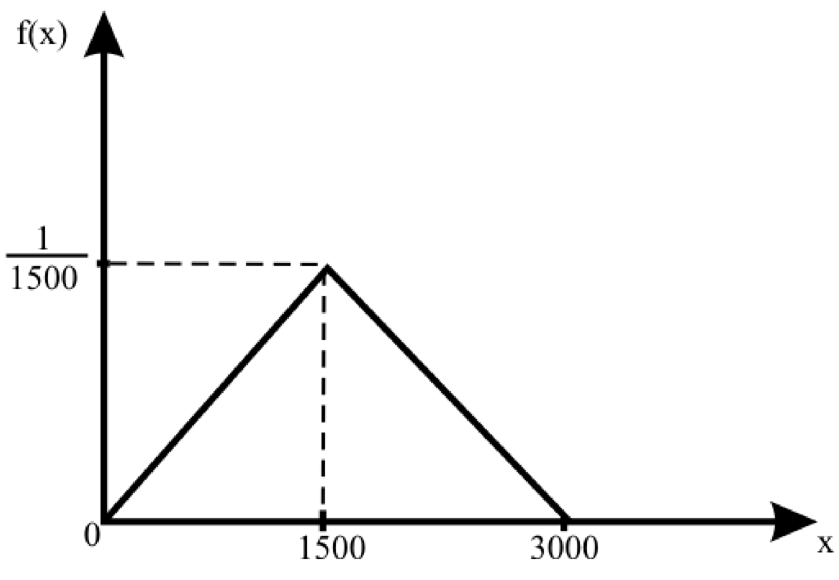


Figura 14. Curva da distribuição densidade de probabilidade do tempo X durante o qual um motor elétrico de um sistema de irrigação é usado em carga máxima, num certo período de tempo em minutos.

MEDIDAS CARACTERÍSTICAS DAS DISTRIBUIÇÕES DE PROBABILIDADES

Esperança matemática ou valor esperado ou média de uma variável aleatória

$$E(X) = \mu(X) = \mu X = \mu = m$$

Para variável aleatória discreta

Para uma variável aleatória discreta X que pode tomar os valores $X_1, \ldots, X_n$ com as respectivas probabilidades $P_1, \ldots, P_n$, defini-se a esperança de X como:

$$E(X) = \mathrm{X_1 P_1 + X_2 P_2 \ldots .. + X_n P_n} = \sum_{\mathrm{i}=1}^{n} X_i P_i$$

Sob o ponto de vista científico a esperança matemática corresponde ao que se espera que aconteça, em média.

A esperança matemática pode ser pensada como uma média aritmética ponderada, senão vejamos:

$$E(x) = x_1 f(x_1) + x_2 f(x_2) + \cdots + X_n f(x_n) \text{ ou } E(x) = \sum_{\mathrm{i}=1}^{\mathrm{n}} X_i f(x_i)$$

Isto é, $E(x)$ é a média ponderada dos possíveis valores de X_i cada um ponderado por sua probabilidade.

Existe uma interpretação física para a média e a variância. Suponha que, em cada ponto X_i dos eixos dos X, seja colocada uma quantidade de massa $[f(x_i) = P(x_i)]$. Então, a média é o centro de gravidade do sistema e a variância, o momento de inércia do sistema.

Esperança condicional $\{\boldsymbol{E}[(\boldsymbol{Y}/\boldsymbol{X})]\}$

A esperança condicional da variável aleatória discreta Y dada a variável aleatória X, que denotaremos por $E[(Y/X)]$, é a função cujo valor para $X = x$ é dada por:

$$E[(Y/X)] = \sum_{y} YP(Y = y/X = x)$$

onde a soma se estende a todos os valores da variável Y. E a definição é análoga para $\boldsymbol{E}[(\boldsymbol{Y}/\boldsymbol{X})]$.

Para variável aleatória continua

Para uma variável aleatória contínua X, com função de densidade de probabilidade, define-se a esperança como:

Se $\mathrm{P}(a \leq \mathrm{X} \leq b) = \int_a^b f(x)\mathrm{d}x$, então x é considerada uma variável aleatória contínua.

A esperança matemática de x, simbolizada por $[E(x)]$ é definida por:

$$E(x) = \int_{\Re} xf(x)dx \text{ ou } \int_{-\infty}^{\infty} x\mathrm{f}(x)\mathrm{d}x$$

Pode acontecer que esta integral (imprópria) não convirja. Consequentemente diremos que $E(x)$ existirá se e somente se $\int_{-\infty}^{\infty} |x|f(x)\mathrm{d}x$ for finita.

Esperança condicional ou média condicional $\{E[(Y/X)]\}$

A esperança condicional, ou média condicional, de Y dado $X = x$ é definida por:

$$E[(Y/X)] = \int_{-\infty}^{+\infty} Yf(Y/X) \ dy$$

E definição análoga para $E[(X/Y)]$.

Verifica-se que a esperança condicional de Y dado $X = x$ é a esperança de Y com relação à distribuição condicional de Y dado $X = x$. Em geral, quando falamos da esperança condicional de Y dado X, nos referimos à variável aleatória $E[(Y/X)]$ que para $X = x$ tem o valor dado pela expressão $E[(Y/X = x)]$.

Propriedades da esperança matemática

1) $E(K) = K$, sendo K uma constante
2) $E(KX) = KE(X)$, sendo K uma constante
3) $E(X \pm K) = E(X) \pm K$
4) $E(X \pm Y) = E(X) \pm E(Y)$
5) $E(XY) = E(X)E(Y)$. Se X e Y são independentes.
6) $E\left[\left(X - \mu_x\right)\right] = E(X) - E\left[\mu_X\right] = \mu_x - \mu_x = 0$, isto é a média ou esperança de uma variável aleatória centrada na média é sempre nula.

Dedução das propriedades da esperança matemática: Propriedades da Esperança Matemática $\boldsymbol{E(X) = \mu(X) = \mu X = \mu = m}$

(1) A média de uma constante é igual a própria constante

Sendo K uma constante, então: $E(K) = K$, pois temos que: $E(K) = [(KP_1) + (KP_2) + \cdots + (KP_n)] = K(P_1 + P_2 + \cdots + P_n) = (K)(1)$, pois, $(\mathrm{P}_1 + \mathrm{P}_2 + \cdots + \mathrm{P}_\mathrm{n}) = \sum_{i=1}^{n} P(X_i) = 1$.

(2) Se multiplicarmos os valores de uma variável aleatória por uma constante, a média fica multiplicada por essa constante

$$E(KX) = KE(X)$$

(2.1) Para o caso de variáveis aleatórias discretas temos que:

$$E(KX) = \sum_{i=1}^{n} KX_iP_i = K\sum_{i=1}^{n} X_iP_i = KE(X)$$

(2.2) Para o caso de variáveis aleatórias contínuas temos que:

$$E(KX) = \int_{-\infty}^{+\infty} KXf(X)\,dX = K\int_{-\infty}^{+\infty} Xf(X)dX = KE(X)$$

(3) Se somarmos ou subtrairmos um valor constante e arbitrário aos valores de uma variável aleatória, a média fica somada ou diminuída dessa mesma constante

$$E(X \pm K) = E(X) \pm K$$

Demonstraremos o caso de $E(X+K)$, mas ressaltando que vale para $E(X-K)$:

$$\mathrm{E(X+K)} = \sum_{i=1}^{n}(X_i+K).P_i = \sum_{i=1}^{n} X_i.P_i + \ K\sum_{i=1}^{n} P_i = E(X) + K,$$

pois $\sum_{i=1}^{n} P_i = 1$.

(4) A média de uma soma ou diferença de variáveis aleatórias é igual à soma ou diferença das médias dessas variáveis

$$E(X \pm Y) = E(X) \pm E(Y)$$

(4.1) Para o caso de variáveis aleatórias discretas temos que:

Suponhamos que as variáveis X e Y são definidas para o conjunto de valores $X_1, X_2, \dots X_m$, e $Y_1, Y_2, \dots, Y_n$. Indicaremos por $P(X_i, Y_j)$. A probabilidade de ocorrerem simultaneamente X_ie $Y_j(i = 1,2, \dots, m; j = 1,2, \dots, n)$. Assim $P(X_1Y_2)$ indica a probabilidade de ocorrer X_1 com Y_2, $P(X_2Y_1)$ a probabilidade de ocorrer X_2 com Y_1, etc.

Pela definição de esperança matemática, podemos escrever:

$$\begin{aligned} E(X+Y) = {} & (X_1+Y_1)P(X_1,Y_1) + (X_1+Y_2)P(X_1,Y_2) + \cdots + (X_1+Y_n)P(X_1,Y_n) \\ & + (X_2+Y_1)P(X_2,Y_1) + (X_2+Y_2)P(X_2,Y_2) + \cdots + (X_2+Y_n)P(X_2,Y_n) + \cdots \\ & + (X_m+Y_1)P(X_m,Y_1) + (X_m+Y_2)P(X_m,Y_2) + \cdots + (X_m+Y_n)P(X_m,Y_n) \end{aligned}$$

Esta soma pode ser colocada sob a seguinte forma:

$$\begin{aligned} \mathrm{E(X+Y)} = {} & X_1[\mathrm{P}(X_1,Y_1) + \mathrm{P}(X_1,Y_2) + \dots + \mathrm{P}(X_1,Y_n)] + X_2[P(X_2,Y_1) + P(X_2,Y_2) + \dots \\ & + P(X_2,Y_n)] + \dots + X_m[P(X_m,Y_1) + P(X_m,Y_2) + \dots + P(X_m,Y_n)] \\ & + Y_1[P(X_1,Y_1) + P(X_2,Y_1) + \dots + P(X_m,Y_1)] + Y_2[P(X_1,Y_2) + P(X_2,Y_2) \\ & + \dots + P(X_m,Y_2)] + \dots + Y_n[P(X_1,Y_n) + P(X_2,Y_n)] + \dots + P(X_m,Y_n)] \end{aligned}$$

Mas pelo teorema da soma de probabilidades temos:

$$P_1 = P(X_1) = P(X_1,Y_1) + P(X_1,Y_2) + \cdots + P(X_1,Y_n)$$
$$P_2 = P(X_2) = P(X_2,Y_1) + P(X_2,Y_2) + \cdots + P(X_2,Y_n)$$

……………………………………………………….

$$P_m = P(X_m) = P(X_m,Y_1) + P(X_m,Y_2) + \cdots + P(X_m,Y_n)$$

$$q_1 = P(Y_1) = P(X_1, Y_1) + P(X_2, Y_1) + \cdots + P(X_m, Y_1)$$
$$q_2 = P(Y_2) = P(X_1, Y_2) + P(X_2, Y_2) + \cdots + P(X_m, Y_2)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$q\_n = P(Y\_n) = P(X\_1, Y\_n) + P(X\_2, Y\_n) + \ldots + P(X\_m, Y\_n)$$

Sendo assim temos que:

$$E(X+Y) = X_1P(X_1) + X_2P(X_2) + \cdots + X_mP(X_m) + Y_1P(Y_1) + Y_2P(Y_2) + \cdots + Y_nP(Y_n)$$

$$E(X+Y) = E(X) + E(Y)$$

(4.2) Para o caso de variáveis aleatórias contínuas essas propriedades são válidas de forma semelhante:

A média do produto de duas variáveis aleatórias independentes é igual ao produto das médias dessas variáveis. Sendo assim temos

$$E\,(XY) = E(X)E(Y)$$

(5) Para o caso de variáveis aleatórias discretas temos que:

Para o caso do produto de duas variáveis aleatórias, consideramos as variáveis X e Y, definidas para os conjuntos de valores abaixo.

$$X_1, X_2, \ldots X_m, \text{ e } Y_1, Y_2, \ldots, Y_n.$$

Sendo assim a esperança matemática do produto de duas variáveis aleatórias é definida por.

$$E(XY) = X_1Y_1P(X_1, Y_1) + X_1Y_2P(X_1, Y_2) + \cdots + X_1Y_nP(X_1, Y_n) + X_2Y_1P(X_2, Y_1) + X_2Y_2P(X_2, Y_2) + \cdots + X_2Y_nP(X_2, Y_n) + \cdots + \cdots + X_mY_1P(X_m, Y_1) + X_mY_2P(X_m, Y_2) + \cdots + X_mY_nP(X_m, Y_n)$$

onde $P(X_i, Y_j)$ é a probabilidade de ocorrerem simultaneamente X_i e Y_j $(i = 1,2,\ldots,m\,;\ j = 1,2,\ldots,n)$.

Nesse caso temos que impor a condição de que os eventos sejam independentes. Para isto devemos ter $P(X_i, Y_j) = P(X_i)P(Y_j)$, onde $P(X_i)$e $P(Y_j)$ são as probabilidades individuais ou marginais de ocorrerem um particular valor de X ou de Y, respectivamente. Obtém-se assim:

$$E(XY) = X_1Y_1P(X_1)P(Y_1) + X_1Y_2P(X_1)P(Y_2) + \cdots + X_1Y_nP(X_1)P(Y_n) + X_2Y_1P(X_2)P(Y_n) + X_2Y_2P(X_2, Y_2) + \cdots + X_2Y_nP(X_2)P(Y_n) + \cdots + \cdots + X_mY_1P(X_m)P(Y_1) + X_mY_2P(X_m)P(Y_2) + \cdots + X_mY_nP(X_m)P(Y_n)$$

Esta expressão pode ser colocada na seguinte forma.

$$\mathrm{E(XY)} = X_1\mathrm{P}(X_1)[(Y_1)\mathrm{P}(Y_1) + Y_2\mathrm{P}(Y_2) + \ldots + Y_n\mathrm{P}(Y_n)] + X_2P(X_2)[P(Y_1) + Y_2P(Y_2) + \ldots + Y_nP(Y_n)] + \ldots + \ldots + X_mP(X_m)[Y_1\,P(Y_1) + Y_2\,P(Y_2) + \ldots + Y_nP(Y_n)$$

$$E(XY) = [X_1\,P(X_1) + X_2\,P(X_2) + \ldots + X_m\,P(X_m)][Y_1\,P(Y_1)] + Y_2\,P(Y_2) + \ldots + Y_n\,P(Y_n)]$$

$$E(XY) = E(X)E(Y)$$

(5.1) Para o caso de variáveis aleatórias contínuas essas propriedades são válidas de forma semelhante:

A média de uma variável aleatória centrada é zero ou nula, isto é, $E\ (erro) = E[(X - \mu\ )] = 0$.

(5.2) Para o caso de variáveis aleatórias discretas temos que:

$$E(e) = E[(X - \mu)] = (X_1 - \mu)P_1 + (X_2 - \mu)P_2 + \cdots + (X_n - \mu)P_n$$

$E(e) = E[(X - \mu)] = [(X_1P_1) + (X_2P_2) + \cdots + (X_nP_n)] - \mu(P_1 + P_2 + \cdots + P_n)$, e como,

$$[(X_1P_1) + (X_2P_2) + \cdots + (X_nP_n)] = E(X) = \mu, e(P_1 + P_2 + \cdots + P_n) = 1,$$

Temos finalmente que: $E(e) = \mu - \mu = 0$.

(5.3) Para o caso das variáveis aleatórias contínuas o procedimento é semelhante. Então fica assim:

$$\mathrm{E(e)} = \int_{-\infty}^{+\infty} (X - \mu)\, dP = \int_{-\infty}^{+\infty} X dP - \mu \int_{-\infty}^{+\infty} dP$$

E como $\int_{-\infty}^{+\infty} X\, dP = \mu$, $\int_{-\infty}^{+\infty} dP = 1$, obtemos finalmente que:

$$E(e) = \mu - \mu = 0$$

como no caso anterior.

(6) $E(e_1)E(e_2) = 0$

Sendo X_1 e X_2 duas variáveis casuais independentes, de médias μ_1 e μ_2, os desvios correspondentes são: $e_1 = X_1 - \mu_1$, e $e_2 = X_2 - \mu_2$, e já sabemos que

$E(e_1) = E(X_1) - E(\mu_1) = \mu_1 - \mu_1 = 0$ e que $\mathrm{e_2} = E(X_2) - E(\mu_2) = \mu_2 - \mu_2 = 0$

Demonstraremos que $E(e_1e_2) = 0$.

Com efeito, pelo teorema da esperança matemática do produto, temos:

$$E(e_1e_2) = E(e_1)E(e_2) = (0)(0) = 0$$

(6.1) Funções de variáveis aleatórias

Seja X uma variável aleatória. Então $Y = g(X)$ é também uma variável aleatória, ou seja, o caso agora é de uma variável aleatória que depende de X, por exemplo $g(X) = X^2$ ou $g(X) = (X - \mathrm{\mu})^2$ segue-se que:

$$P(Y = y) \sum_{[x \backslash g(x) = y]} P(X = x)$$

O valor esperado ou esperança matemática de $g(X)$ também pode ser obtido a partir da função de probabilidade ou de densidade de probabilidade de X.

(6.1.1) Esperança matemática ou média para a função de variável aleatória $g(X)$ de uma variável aleatória discreta $X\left\{E\left[g\left(X\right)\right]\right\}$.

Seja X uma variável aleatória discreta com função de probabilidade $f(X) = P(X_i)$, então

$$E[g(X)] = g(x_1)f(x_1) + g(x_2)f(x_2) + \cdots + g(x_n)f(x_n)$$

$$E[g(X)] = \sum_{i=1}^{n} g(x_i)\, f(x_i)$$

(6.1.2) Esperança matemática ou média para a função de variável aleatória $g(X)$ de uma variável aleatória contínua $X\left\{E\left[g\left(X\right)\right]\right\}$.

Se X é uma variável aleatória contínua com função densidade de probabilidade $f(X)$, então.

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)\, f(x)dx$$

VARIÂNCIA $\left\{\mathrm{VAR}(x) \ = \ \mathrm{V}(x) \ = \ \Sigma^2_{(x)} = \sigma^2\right\}$

Define-se variância de uma variável aleatória como sendo.

(1) Variável aleatória discreta

$$\sigma^2_{(x)} = E[(x_i - \mu)^2] = \ \Sigma x_i^2 \ . \mathrm{P}(x_i) \text{ - } \mu^2 = \ \mathrm{E}(x_i^2) \text{ - } \mu^2$$

onde

$$\begin{cases} E(x^2) \ = \ \displaystyle\sum_{i=1}^{n} x_i^2 f(x_i)\, , m \ = \ \mu = E(x) \end{cases}$$

$$m \ = \ \mu = E(x)$$

$$\sigma^2_{(x)} = E[(x - \mu)^2] = E(x^2) - [E(x)]^2$$

(2) Variável aleatória contínua

$$\sigma^2_{(x)} = \ \mathrm{E}[(x - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\mathrm{dx} = \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x - \mu^2$$

onde

$$\begin{cases} E(x^2) \ = \ \displaystyle\int_{-\infty}^{+\infty} x^2 f(x)\mathrm{dx} \end{cases}$$

$$V(x) = \sigma^2_{(x)} = E(x^2) - [E(x)]^2$$

Propriedades da variância

a) $V(K) = 0$, sendo K uma constante

b) $V(KX) = K^2V(X)$, sendo K uma constante

c) $V(X \pm K) = V(X)$, sendo K uma constante

d) $V(X \pm Y) = V(X) + V(Y)$, se X, Y são variáveis aleatórias independentes.

e) $V(X \pm Y) = V(X) + V(Y) \pm 2.COV.(X,Y)$, se X, Y são variáveis aleatórias dependentes.

Dedução das propriedades da variância: Propriedades da Variância $[\boldsymbol{V(X)}] = \boldsymbol{Var(X)} = \boldsymbol{\sigma(X)^2 = \sigma X^2 = \sigma^2}$

(a) A variância de uma constante é nula

$$Var(K) = 0, Var(K) = E[K - E(K)]^2 = E[K - K]^2 = E(K2) - [E(K)]^2$$

$$E(K^2) = \sum_{i=1}^{n} K^2P(K) = K^2 \sum_{i=1}^{n} P(K) = K^2 \cdot 1 = K^2$$

$$[E(K)] = \sum_{i=1}^{n} KP(K) = K \sum_{i=1}^{n} P(K) = K \cdot 1 = K$$

então

$$Var(K) = K^2 - [K]^2 = K^2 - K^2 = 0$$

(b) Se for multiplicado todos os valores de uma variável aleatória por uma constante, sua variância fica multiplicada pelo quadrado dessa constante.

Com efeito temos

$$Var(KX) = E\{[(KX - E(X)]\text{^}2\},$$

Onde

$$E(KX) = KE(X) = K\mu,$$

portanto:

$$Var(KX) = E[KX - E(KX)]^2 = E[KX - KE(X)]^2 = EK[X - E(X)]^2 = Var(KX)$$
$$= EK^2[X - E(X)]^2 = K^2 \underbrace{E[X - E(X)]^2}_{Var(X)}$$

logo,

$$Var(KX) = K^2Var(X)$$

(c) Se somarmos ou subtrairmos uma constante aos valores de uma variável aleatória, sua variância permanece inalterada

$$Var(Xi \pm K) = Var(X)$$

$$Var(Xi + K) = Var(Xi) + Var(K)$$

$$Var(Xi + K) = Var(Xi) + 0,$$

como

$$Var(Xi) = Var(X),$$

temos que:

$$Var(Xi + K) = Var(X)$$

(d) A variância de uma soma ou diferença de variáveis aleatórias independentes é igual à soma das variâncias dessas variáveis

$$Var\ (X + Y) = Var\ (X) + Var(Y),$$

se X e Y são variáveis casuais independentes mostraremos que

$$Var(X + Y) = Var(X) + Var(Y),$$

de fato:

$$Var(X + Y) = E[(X + Y - \mu)]^2,$$

onde

$$\mu = E[X + Y] = E(X) + E(Y) = \mu_1 + \mu_2,$$

onde:

$$\mu_1 = E(X) \text{ e } \mu_2 = E(Y).$$

Temos, pois, com

$$e_1 = X - \mu_1 e e_2 = Y - \mu_2$$

$$Var(X + Y) = E[(X + Y - \mu_1 - \mu_2)^2]$$

$$Var(X + Y) = E[(e_1 + e_2)^2]$$

$$Var(X + Y) = E[(e_1^2 + 2e_1e_2 + e_2^2)] =$$

$$Var(X + Y) = E(e_1^2) + 2E(e_1)E(e_2)] + E(e_2^2) =$$

$$Var(X + Y) = Var(X) + Var(Y),$$

pois, como foi visto, temos sempre que,

$$E(e_1) = E(e_2) = 0$$

Esta propriedade pode ser facilmente generalizada para o caso da soma de "n" variáveis aleatórias independentes. Obtemos então:

$$Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$$

Já este teorema, por sua vez, pode ser combinado com o anterior, para fornecer a variância de uma função linear de variáveis aleatórias independentes, da seguinte forma:

$$Var(K_1X_1 + K_2X_2 + \cdots + K_nX_n) = Var(K_1X_1) + Var(K_2X_2) + \cdots + Var(K_nX_n)$$
$$= K_1^2 Var(X_1) + K_2^2 Var(X_2) + \cdots + K_n^2 Var(X_n)$$

Se essas variâncias forem todas iguais, isto é, se tivermos:

$$Var(X_1) = Var(X_2) = \cdots = Var(X_n) = \sigma^2,$$

obteremos o seguinte:

$$Var(K_1X_1) + Var(K_2X_2) + \cdots + Var(K_nX_n) = (K_1^2 + K_2^2 + \cdots + K_n^2)\sigma^2 = \sigma^2 \sum_{i=1}^{n} K^2$$

Como caso particular temos o da variância da média aritmética.

$$Var(\overline{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \sum_{i=1}^{n} \frac{1}{n^2} Var(X_i) = \frac{n}{n^2}\sigma^2 = \frac{\sigma^2}{n}$$

(e) $Var(X - Y) = Var\ (X) + Var(Y)$, se X e Y são independentes.

Prova:

$$Var(X - Y) = Var(X) + (-1)^2 Var(Y) = Var(X) + Var(Y)$$

(e.1) Se X e Y não são variáveis aleatórias independentes, então,

$$Var\ (X \pm Y) = Var\ (X) + Var\ (Y) \pm 2[\underbrace{E(XY)E(X)E(Y)}_{COVARIÂNCIA}]$$

$$Var(X + Y) = E[(X + Y) - E(X + Y)]^2 = Var(X + Y) = E\big[\big(X - E(X)\big) + \big(Y - E(Y)\big)\big]^2 = Var(X + Y)$$
$$= E[X - E(X)]^2 + [Y - E(Y)]^2 + 2[X - E(X)][Y - E(Y)] = Var(X + Y)$$
$$= E[X - E(X)]^2 + E[Y - E(Y)]^2 + 2E[X - E(X)][Y - E(Y)] = Var(X + Y)$$
$$= Var(X) + Var(Y) + 2Cov(X, Y)$$

(e.2) $Var(X_i - \mu) = \frac{(n-1)}{n}\sigma^2$

Consideremos o desvio, erro, afastamento, discrepância ou resíduo: $e_1 = \left(X_1 - \overline{X}_1\right)$, em relação à média aritmética. Qual será a sua variância? Então vejamos.

Temos:

$$e_1 = X_1 - \overline{X} = X_1 - \frac{(X_1 + X_2 + \cdots X_n)}{n} = \frac{(n-1)}{n}X_1 - \frac{1}{n}X_2 - \cdots - \frac{1}{n}X_n,$$

logo:

$$Var(e_1) = \left(\frac{n-1}{n}\right)^2 Var(X_1) + \left(\frac{1}{n}\right)^2 Var(X_2) + \cdots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

Se tivermos $Var(X_i) = \sigma^2$, para $i = 1,2, \ldots, n$, obteremos:

$$Var(e_1) = Var(X_1 - \overline{X}) = \left(\frac{n-1}{n}\right)^2 \sigma^2 + \left(\frac{n-1}{n^2}\right)\sigma^2 = \left(\frac{n-1}{n}\right)\sigma^2$$

Por outro lado, pela definição de variância temos:

$$Var(X_1 - \overline{X}) = E\left\{\left[\left(X_1 - \overline{X}\right) - \left(E(X_1 - \overline{X}\right)\right]^2\right\} = E\left(\frac{n-1}{n}X_1 - \frac{1}{n}X_2 - \cdots - \frac{1}{n}X_n\right) =$$

$$= \left(\frac{n-1}{n}\right)E(X_1) - \left(\frac{1}{n}\right)E(X_2) - \cdots - \left(\frac{1}{n}\right)E(X_n)$$

Suponha que para todas as variáveis aleatórias X_i, com $i = 1,2,\ldots,n$, temos $E(X_i) = \mu$, $Var(X_i) = \sigma^2$.

Então fica assim:

$$E\left(X_1 - \overline{X}\right) = \left(\frac{n-1}{n}\right)\mu - \left(\frac{n-1}{n}\right)\mu = 0,$$

$$Var(X_1 - \overline{X}) = E\left[\left(X_1 - \overline{X}\right)^2\right],$$

Logo tem-se:

$$E\left[\left(X_1 - \overline{X}\right)^2\right] = \left(\frac{n-1}{n}\right)\sigma^2$$

Como é sabido que a fórmula da variância de um conjunto de valores é dada pela equação a seguir, sendo assim temos:

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1},$$

Pela demonstração acima verifica-se que:

$$E(S^2) = \left(\frac{1}{n-1}\right)E\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] = \left(\frac{1}{n-1}\right)\sum_{i=1}^{n}E\left[X_i - \overline{X}\right]^2 = \left(\frac{1}{n-1}\right)n\left(\frac{n-1}{n}\right)\sigma^2 = \sigma^2$$

(i) $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2.\sum_{i<j}^{n} Cov(X_i, Y_j)$, para $E(X_i^2) < \infty$, sendo $1 \leq i \leq n$

(ii) $Var(aX \pm b) = a^2 Var(X)$, se a e b são constantes.

Demonstração:

$$Var\,(aX \pm b) = Var(aX) + Var(b) \pm 2Cov(aX, b)$$

Como

$$Cov(aX, b) = E\{[aX - E(aX)][b - E(b)]\} = 0,$$

temos enfim que,

$$Var\left(aX \pm b\right) = a^2 Var(X)$$

DESVIO PADRÃO $[D.P. = \sigma_{(X)} = \sigma_x]$

O desvio padrão de uma variável aleatória é definido como a raiz quadrada positiva da variância, isto é $\sigma_{(x)} = \sqrt{VAR(X)}$.

Exercício de aplicação

Exercício 1. Um experimento aleatório consiste em colocar para germinar duas sementes de algodão herbáceo. Seja o espaço amostral referente a esse ensaio:

$$S = \{(G,G); (G,\bar{G}); (\bar{G},G); (\bar{G},\bar{G})\}$$

$$X = \text{Número de sementes que germinará}$$

$$X = \{0, 1, 2\}$$

TABELA 4. DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA DISCRETA X: DUAS SEMENTES DE ALGODÃO.

X	0	1	2	Soma
$P(X)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{4}{4}=1{,}0$

(i) Calcular a esperança matemática ou o número médio de sementes $[E(x)]$ da Distribuição.

	X	$P(X)$	$XP(X)$	X^2	$X^2P(X)$
	0	$1/4$	0	0	0
	1	$2/4$	$2/4$	1	$2/4$
	2	$1/4$	$2/4$	4	$4/4$
Σ	-	-	1,0	-	$\frac{6}{4} = 1{,}5$

$$E(X) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{2}{4} + 2 \cdot \frac{1}{4} = 1{,}0$$

$E(x) = 1{,}0$ sementes. Isto significa que se for colocado pra germinar duas sementes de algodão em um grande número de ensaios, espera-se que germine em média uma semente.

(ii) Calcular a variância $[V(x)]$ da Distribuição.

$$E(x^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{2}{4} + 2^2 \cdot \frac{1}{4} = \frac{2}{4} + \frac{4}{4} = \frac{6}{4} = 1{,}5$$

$$V(x) = 1{,}5 - (1{,}0)^2 = 0{,}5$$

$$V(x) = 0{,}5(semente)^2$$

(iii) Calcular o desvio padrão (σ_x) da Distribuição.

$\sigma_x = \sqrt{V(x)} = \sqrt{0{,}5} = 0{,}707$ semente. Este valor deve ser comparado com outro valor de uma outra distribuição semelhante, o que for menor apresenta menor variabilidade ou maior homogeneidade.

Exercício 2. Qual seria o preço a pagar (preço justo) por uma rifa que nos dará um primeiro prêmio no valor de R$ 25.000,00, com a probabilidade de 0,0001 e, um segundo prêmio, no valor de R$ 10.000,00 com a probabilidade de 0,0002?

$$E(x) = 25{,}000 \,.\, 0{,}0001 + 10.000 \,.\, 0{,}0002 = R\$\, 4{,}50$$

Exercício 3. Um jogador lança um dado. Se aparecerem os números 1,2 ou 3, recebe R$ 10,00. Se no entanto, aparecer 4 ou 5, recebe R$ 5,00. Se aparecer 6 ganha R$ 20,00. Qual o ganho médio do jogador?

TABELA 5. DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA X: VALORES EM REAIS.

X	10	10	10	5	5	20
Faces	1	2	3	4	5	6
$P(x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$E(x) = 10 \,.\, \frac{1}{6} + 10 \,.\, \frac{1}{6} + 10 \,.\, \frac{1}{6} + 5 \,.\, \frac{1}{6} + 5 \,.\, \frac{1}{6} + 20 \,.\, \frac{1}{6}$$

$$E(x) = 10(3/6) + 5(2/6) + 20(1/6)$$

$$E(x) = \frac{30}{6} + \frac{10}{6} + \frac{20}{6} = \frac{60}{6} = 10$$

$$E(x) = R\$\, 10{,}00$$

Exercício 4. Qual seria o preço a pagar (preço justo) por uma rifa que nos dará um primeiro prêmio no valor de R$ 5.000,00, com probabilidade de 0,0001, e um segundo prêmio, no valor de R$ 1.000,00 com a probabilidade de 0,0002?

$$E(x) = 5.000 \,.\, 0{,}0001 + 10.000 \,.\, 0{,}0002$$

$$E(x) = 0{,}5 + 0{,}2 = R\$\, 0{,}70$$

Exercício 5. Ao realizar uma operação financeira um investigador possui 75% de probabilidade de ganhar R$ 50.000,00 e 25% de probabilidade de perder R$ 30.000,00. Qual sua esperança de lucro ou sua esperança matemática?

$$E(x) = 50.000{,}00 \cdot \frac{75}{100} + (-30.000) \cdot \frac{25}{100}$$

$$E(x) = 37.500 + (-7500)$$

$$E(x) = R\$\ 30.000{,}00$$

Exercício 6. Determine a esperança de soma de pontos obtidos na jogada de um par de dados. $S = \{(1,1), (1,2), \ldots (6,6)\}$

Sejam X e Y os pontos que aparecem nas faces dos dois dados. Temos

$$E(x) = \mathrm{E}(y) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$$

$$E(x) = E(y) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = \frac{7}{2}$$

$$E(x+y) = E(x) + E(y) = \frac{7}{2} + \frac{7}{2} = \frac{14}{2} = 7$$

Exercício 7. Um jogador lança duas moedas não viciadas. Ele ganha R$ 5,00 se ocorrerem 2 caras, R$ 2,00 se ocorrer 1 cara e R$ 1,00 se não ocorrer cara.

$$(i)\ Ache\ seu\ ganho\ esperado.$$

$$S = \{CC, CT, TC, TT\},$$

onde $C =$ ocorre face cara e $T =$ ocorre face coroa.

TABELA 6. DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA X: VALORES RECEBIDOS EM REAIS.

X_i	+5	+2	+1
$P(X_i)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$$E(x) = 5 \cdot \frac{1}{4} + 2 \cdot \frac{2}{4} + 1 \cdot \frac{1}{4} = 2{,}50$$

Isto é o ganho esperado é R$ 2,50.

Exercício 8. Uma variável aleatória contínua tem a seguinte função de densidade de probabilidade

$$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ 3(x-1)^2 & \text{para } 0 \le x \le 1 \\ 0 & \text{para valores de } x > 1 \end{cases}$$

(i) Qual é a $E[x]$ e $Var[x]$

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{0}^{1} x 3(x - 1)^2 \, dx = 3 \int_{0}^{1} (x^2 - 2x + 1) \, x dx$$

$$E(x) = 3 \int_{0}^{1} (x^2 - 2x^2 + x) \, dx = \left[\frac{3}{4} x^4 - \frac{6}{4} x^3 + \frac{3}{2} x^2\right]_0^1 = \frac{1}{4} = \mu$$

$$V(x) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) \, dx - \mu^2$$

$$V(x) = \int_{0}^{1} x^2 \, 3 \, (x - 1)^2 \, dx - \mu^2 = \int_{0}^{1} (3x^4 - 6x^3 + 3x^2) \, dx - \mu^2$$

$$V(x) = \left[\frac{3}{5} x^5 - \frac{6x^4}{4} + \frac{3}{3} x^3\right]_0^1 - \mu^2 = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{3}{8}$$

$$VAR(x) = V(x) = \sigma^2_{(x)} = 0{,}0375$$

COVARIÂNCIA $[\boldsymbol{Cov(x, y)}]$

Dadas duas variáveis aleatórias X e Y, a covariância entre X e Y é:

Determinação (cálculo)

$$COV\ (X, Y) = \sum_i \sum_j (x_1 - \mu_x)(y_i - \mu_y) P(x_1, y_j)$$

$$COV\ (X, Y) = E[(x - \mu_x)(y - \mu_y) = E(xy) - \mu_x \, \mu_y = E(xy) - E(x)E(y)$$

ou ainda:

$$COV(X, Y) = E[x - E(x)] \, [y - E(y)] = E[(x - \mu_x)(y - \mu_y)] = E(xy) - \mu_x \mu_y$$

$$COV(X, Y) = E(xy) - E(X)E(Y)$$

Definição de covariância

O número $E\{[X - E(X)] \, [Y - E(Y)\}$ é chamado covariância de X e Y e é denotado por $COV(X, Y)$. Para que haja covariância é necessário que existam pelo menos duas variáveis. $COV(X, Y)$, traduz uma relação de dependência entre estas duas variáveis. A magnitude da covariância depende das unidades em que as variáveis aleatórias X e Y forem medidas. Se essas unidades mudam, então o valor da covariância também muda. Assim a informação contida na covariância é principalmente sobre o sinal da relação entre X e Y e não sobre a sua intensidade.

A covariância dá uma ideia da dispersão dos valores da variável aleatória bidimensional (X, Y) em relação ao ponto $\{E[X]; E[Y]\}$.

Vê-se da definição dessa medida que a covariância é a esperança dos produtos dos desvios dos valores de cada uma das duas variáveis em relação a suas médias. Sendo assim se as duas variáveis tendem a variar no mesmo sentido, isto é, valores de X acima de sua média estão associados a valores de Y acima de sua média, o mesmo ocorrendo para valores de ambos inferiores à média, então a covariância será positiva. Caso contrário, se valores acima da média de uma variável estão associados a valores inferiores à média da outra, então a covariância será negativa.

Propriedades da covariância

$$COV(X,Y) = E(XY) - E(X)E(Y)$$

$$COV(X,Y) = COV(Y,X),$$

logo a covariância é simétrica.

Se $V(X) = 0$ (ou $V(y) = 0$) então $COV(X,Y) = 0$.

$COV(ax,y) = aCOV(x,y)$ para $a \in \mathbb{R}$ então a $COV(X,Y)$ é bilineal

$$COV(X + Z,Y) = COV(X,Y) + COV(Z,Y).$$

$$COV(KX,Y) = COV(X,KY) = K\,COV(X,Y),$$

se K for uma constante.

$$COV(K,X) = COV(X,K) = 0$$

se K for uma constante.

Demonstração das Propriedades da Covariância [COV (X)]

(i) $COV(X,Y) = E(XY) - E(X)E(Y)$

Demonstração:

$$COV(X,Y) = E\{[X - E(X)][Y - E(Y)]\} = E\{XY - YE(X) - XE(Y) + E(X)E(Y)\} =$$

$$COV(X,Y) = E(XY) - E[YE(X)] - E[XE(Y)] + E[E(X)E(Y)] =$$

$$COV(X,Y) = E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) =$$

$$COV(X,Y) = E(XY) - E(X)E(Y)$$

(ii) $COV(X,Y) = COV(Y,X)$, logo a covariância é simétrica.

Demonstração:

$$COV(X,Y) = E(X,Y) - E(X)E(Y)$$

$$COV(Y,X) = E(Y,X) - E(Y)E(X)$$

Logo a $COV(X,Y) = COV(Y,X)$.

(iii) Se $V(X) = 0$ $[ou\, V(Y) = 0]$ então $COV\,(X,Y) = 0$

Demonstração:

$$COV(X,Y) = E\{[X - E(X)][Y - E(Y)]\}$$

$$VAR(X) = E[X - E(X)]^2$$

Se a $Var(X) = 0$, implica que $X = E(X)$, logo,

$$COV(X,Y) = E\{[X - X][Y - E(Y)]\} = E\{[0][Y - E(Y)]\}$$

$$COV(X,Y) = 0$$

(iv) $COV(X + Z, Y) = COV(X,Y) + COV(Z,Y)$

$$\begin{aligned} COV(X+Z,Y) &= E\{[(X+Z) - E(X+Z)][Y - E(Y)]\} \\ &= E\{[((X - E(X)) - (Z - E(Z))][Y - E(Y)]\} \\ &= E\{[(X - E(X))][Y - E(Y)] + [Z - E(Z)][Y - E(Y)\} \\ &= E\{[(X + E(X))][Y - E(Y)]\} + E\{[Z - E(Z)][Y - E(Y)]\} \\ &= COV(X,Y) + COV(Z,Y) \end{aligned}$$

(v) $COV(KX,Y) = COV(X,KY) = KCOV(X,Y)$, para $K \in \mathbb{R}$, isto é a covariância é Bilineal

$$COV(KX,Y) = E\{[KX - E(KX)][Y - E(Y)]\}$$

$$COV(KX,Y) = KE\{[X - E(X)][Y - E(Y)]\}$$

$$COV(KX,Y) = KCOV(X,Y)$$

Ou ainda,

$$COV(KX,Y) = E(KX,Y) - KE(X)E(Y) = KE(X,Y) - KE(X)E(Y)$$
$$= K\{E(X,Y) - E(X)E(Y)\} = KCOV(X,Y)$$

(vi) $COV(K,X) = COV(X,K) = 0$

$$COV(K,X) = E\{[K - E(K)][X - E(X)]\} = E\{[K - E(K)][X - E(X)]\}$$
$$= E\{[K - K][X - E(X)]\} = E\{[0][X - E(X)]\} = 0.[X - E(X)] = 0$$

COEFICIENTE DE CORRELAÇÃO $[r_{xy}]$

Introdução

Se estivermos estudando a dependência entre as variáveis X: peso de um reprodutor bovino em kg e Y: peso do 1º filho em kg, ao calcularmos a covariância, teremos uma medida ao quadrado (kg^2). Além disso a covariância pode assumir quaisquer valores reais no seu campo de variação, ou seja $-\infty < COV(X,Y) < +\infty$. Ou seja como já foi mencionado anteriormente a magnitude da covariância depende das unidades em que as variáveis aleatórias X e Y forem medidas. Se essas unidades mudam, então o valor da covariância também muda. Assim a informação contida na covariância é principalmente sobre o sinal da relação entre X e Y e não sobre a sua intensidade. Isto reduz o interesse da sua aplicação e leva a que se considere uma outra medida que não dependa das unidades.

Introduziremos o conceito de coeficiente de correlação que supera esses problemas. Sendo assim o coeficiente de correlação é uma medida de dependência das variáveis X e Y.

$$r_{xy} = \frac{COV(X,Y)}{\sqrt{V(X)V(Y)}} = \frac{E(\mathrm{XY}) - \mu_x\mu_y}{\sigma_x\sigma_y} = \frac{E(\mathrm{XY}) - E(X)E(Y)}{\sigma_x\sigma_y}$$

O coeficiente de correlação é uma espécie de covariância que não depende das unidades de medida. Pode aliás mostrar-se que o coeficiente de correlação entre X e Y é equivalente à covariância entre $X' = \frac{X-\mu_X}{\sigma_X}$ e $Y' = \frac{Y-\mu_Y}{\sigma_Y}$, que são as formas padronizadas ou estandartizadas de X e Y e admensionais, visto que os numeradores e denominadores têm as mesmas unidades. O coeficiente de correlçao é uma quantidade adimensional $O\ -1 \leq \sigma \leq +1$. Se $\sigma = 0$ (isto é se a covariância é zero), dizemos que as variáveis X e Y são não-correlacionadas em tal caso, porém, as variáveis podem ou não ser independentes. Onde as esperenaças matemáticas conjuntas das duas variáveis aleatórias (XY) são dadas conforme as variáveis aleatórias sejam bidimensionais discretas ou contínuas respectivamente pelas seguintes expresssões:

$$E(XY) = \textstyle\sum_i \sum_j X_i Y_j P(X_i Y_j),$$

para (XY) discreta

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} XYf(X,Y)dxdy,$$

para (XY) contínua.

Propriedades do coeficiente de correlação:

i) $r(X,Y) = r(Y,X)$

ii) $-1 \leq r \leq 1$

iii) $r\,(X,Y) = r(ax + b, cY + d), se\ a,c \neq 0$

iv) $r(X,Y) = 1, r(X,-X) = -1$

É importante destacar que pares de variáveis aleatórias com distribuições marginais (idênticas) podem ter covariâncias e correlações distintas, assim $COV(X,Y)$ e $r(X,Y)$ são medidas do inter-relacionamento entre X e Y.

Devemos frisar que, embora $COV(X,Y) = 0$ sempre que X e Y são variáveis aleatórias independentes, o inverso não é verdadeiro, isto é, se $COV(X,Y) = 0$ não podemos concluir que X e Y são variáveis aleatórias independentes.

Portanto, se X e Y são duas variáveis aleatórias independentes, então $COV(X,Y) = 0$.

De outro modo, se X e Y são independentes, isto implica X e Y não são correlacionadas. A recíproca não é verdadeira, isto é $COV(X,Y) = 0$ não implica X e Y independentes. Entretanto, é possível demonstrar que se as variáveis tem distribuição normal, o fato de a covariância ser igual a zero é condição suficiente para podermos afirmar que são variáveis independentes. Senão, vejamos o exemplo a seguir.

Na Tabela abaixo, apresentamos uma distribuição conjunta em que $COV(X,Y) = 0$ e as variáveis não são independentes, pois $P(X_i,Y_j) \neq P(X_i)P(Y_j)$.

TABELA 7. VALORES DE $P(X_i,Y_j)$ PARA A DISTRIBUIÇÃO CONJUNTA DE DUAS VARIÁVEIS DEPENDENTES COM $COV(X,Y) = 0$.

Y	X			$P(Y)$
	-1	0	+1	
-1	0,10	0,30	0,10	0,50
+1	0,25	0	0,25	0,50
P(X)	0,35	0,30	0,35	1,00

$$E(X) = (-1 \cdot 0{,}35) + (0 \cdot 0{,}30) + (1 \cdot 0{,}35) = 0$$

$$E(Y) = (-1 \cdot 0{,}50) + (1 \cdot 0{,}50) = 0$$

$$E(XY) = (-1)(-1)(0{,}10) + (-1)(0)(0{,}30) + (-1)(1)(010) + (1)(-1)(025) + (1)(0)(0) + (1)(1)(0{,}25)$$

$E(XY) = 0,$ $\quad$ $P(X = -1) = 0{,}35$

$COV(XY) = 0 - 0 \cdot 0 = 0$ $\quad$ $P(Y = +1) = 0{,}50$

$$P(X = -1;\, Y = +1) = 0{,}25 \neq P(X = -1)P(Y = +1) = 0{,}35 \cdot 0{,}50 = 0{,}175$$

Logo X e Y não são independentes

$$r_{xy} = \frac{COV(X,Y)}{\sqrt{V(X)V(Y)}} = \frac{0}{\sqrt{0{,}70 \cdot 1}} = 0$$

Sendo assim as variáveis X e Y não são correlacionadas.

Verifique abaixo as distribuições marginais de probabilidade X e Y, e determine as variâncias de X e de Y.

TABELA 8. DISTRIBUIÇÃO MARGINAL DE X.

X	$P(X)$	$XP(X)$	X^2	$X^2P(X)$
-1	0,35	-0,35	1	0,35
0	0,30	0,00	0	0,00
1	0,35	0,35	1	0,35
SOMA	1,00	0,00	-	0,70

$$E(X) = 0{,}00$$

$$E(X^2) = 0{,}70$$

$$V(X) = 0{,}70 - 0^2 = 0{,}70$$

TABELA 9. DISTRIBUIÇÃO MARGINAL DE Y.

Y	$P(Y)$	$YP(Y)$	Y^2	$Y^2P(Y)$
-1	0,50	-0,50	1	0,50
1	0,50	0,50	1	0,50
SOMA	1,00	0,00	-	1,00

$$E(Y) = 0$$

$$E(Y^2) = 1$$

$$V(Y) = 1 - 0^2 = 1$$

OUTRAS MEDIDAS CARACTERÍSTICAS DAS DISTRIBUIÇÕES DE PROBABILIDADES

Mediana

É um ponto definido segundo a idéia de dividir a distribuição de probabilidade em duas partes equiprováveis denotá-la-emos por md. Será, pois, um ponto tal que $p(X < md) = P(X > md) = 0{,}5$, quando tal ponto existir.

Isto ocorre sempre no caso contínuo em que a mediana pode também ser definida como o ponto tal que $F(md) = 0{,}50$. No caso discreto (às vezes, também, no caso contínuo) quando a condição acima subsistir, haverá todo um intervalo cujos pontos satisfazem a ela, convencionando-se em geral adotar um ponto médio desse intervalo. Caso a condição acima não subsistir, a mediana será o menor valor para qual $F(md) > 0{,}50$.

A mediana representa uma forma alternativa de caracterização do centro da distribuição.

A idéia de mediana pode ser generalizada, imaginando-se a distribuição dividida em várias partes equiprováveis. Tem-se, assim, os quartis (4 partes), decis (10 partes), percentis (100 partes), etc.

Moda

É (são) o(s) ponto(s) de maior probabilidade, no caso discreto, ou maior densidade de probabilidade, no caso contínuo. É, portanto, um parâmetro que indica a região mais provável da distribuição. Designaremos a moda por Mo.

Coeficiente de variação [C.V]

É a razão entre o desvio padrão e a média da distribuição de probabilidade.

$$\text{CV} = \frac{\sigma}{\mu} 100\ (\%)$$

Amplitude

É dada pela diferença entre o maior e o menor valores possíveis da variável. É um parâmetro de dispersão pouco útil no cálculo de probabilidade. Denotá-lo-emos por R.

Parâmetros de assimetria e achatamento (curtose)

Outros parâmetros às vezes usados são os chamados parâmetros de assimetria e achatamento, que dizem respeito à caracterização da forma da distribuição.

Momentos e função geradora de momentos

Momentos

Na estatística descritiva o conceito de momento amostral (simples e centrado) é mostrado e discutido de forma bastante detalhada. No caso das distribuições populacionais, define-se de forma análoga os momentos simples e centrados e a sua interpretação geométrica é semelhante à dos momentos amostrais. Sendo assim é recomendável ao leitor recapitular a assunto sobre momentos amostrais e comparar as definições de momentos amostrais com as de momentos populacionais que será apresentado a seguir.

Momento simples de ordem r

(i) Para a variável aleatória discreta:

O momento simples de ordem r $(r = 0,1,2,\dots)$ é a constante

$$\mu_r' = E(X^r) = \sum_x X^r f(X)$$

(ii) Para a variável aleatória contínua:

$$\mu_r' = E(X^r) = \int_{-\infty}^{+\infty} X^r f(X) dx$$

O momento simples de ordem r é portanto, a média da variável aleatória X^r, logo o momento simples de ordem 1 é a média $\mu_1' = \mu = E(X)$. Deve-se destacar ainda que o momento de ordem 0 é 1.

Momento centrado de ordem r

(i) Para a variável aleatória discreta

Seja $\mu = E(X) = \mu_1'$. O momento centrado de ordem r $(r = 0,1,2,\dots)$ é a constante

$$\mu_r = E[(X-\mu)^r] = \sum_x (X-\mu)^r f(X)$$

(ii) Para a variável aleatória contínua

O momento centrado de ordem r $(r = 0,1,2,\dots)$ é a constante

$$\mu_r = E[(X-\mu)^r] = \int_{-\infty}^{+\infty} (X-\mu)^r f(X) dx$$

A variância é o momento centrado de ordem 2, isto é,

$$\sigma^2 = E[(X-\mu)^2] = \mu_2$$

A assimetria e a curtose estão relacionadas, respectivamente, com os momentos centrados de ordem 3 e 4. No entanto, como os momentos têm dimensões, é habitual usar-se os coeficientes α_1 e α_2 para caracterizar a assimetria e a curtose.

Coeficiente de Assimetria

Seja $E[(X-\mu)^r]$ o momento centrado de ordem r da variável aleatória X e $\sigma^2 = Var(X)$.

O coeficiente de assimetria é a constante

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mu_3}{\sigma^3}$$

O coeficiente de assimetria mede a falta de simetria da função de probabilidade, se a variável aleatória é discreta, ou da função densidade de probabilidade, se a variável aleatória é contínua.

Coeficiente de Curtose

Seja $E[(X-\mu)^r]$ o momento centrado de ordem r da variável aleatória X e $\sigma^2 = Var(X)$.

O coeficiente de curtose é a constante

$$\alpha_4 = \frac{\mu_4}{(\mu_2)^2} = \frac{\mu_4}{\sigma^4}$$

O coeficiente de curtose, embora mais difícil de interpretar, mede o peso das caudas (o grau de achatamento das curvas) da função de probabilidade ou da função densidade de probabilidade, conforme o caso.

FUNÇÃO GERADORA DE MOMENTOS

Os momentos de uma variável aleatória podem ser obtidos a partir do conhecimento da função geradora de momentos, quando esta existe.

Seja X uma variável aleatória. Chama-se função geradora de momentos (f.g.m.) da variável aleatória ao valor esperado da variável aleatória e^{tX} , se este valor existir para todo o valor de t em algum intervalo aberto centrado na origem.

Função Geradora de Momentos para uma variável aleatória discreta

A função geradora de momentos (f.g.m.) da variável aleatória discreta representa-se por $m_x(t)$, sendo calculada por:

$$m_x(t) = E(e^{tx}) = \sum_x e^{tx} f(x)$$

Função Geradora de Momentos para uma variável aleatória contínua

A função geradora de momentos (f.g.m.) da variável aleatória contínua representa-se por $m_x(t)$, é determinada por:

$$m_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

Quando não houver ambiguidade, $m_x(t)$ representa-se somente por $m(t)$. Se a função geradora de momentos existir, então $m(t)$ é indefinidamente derivável em torno da origem e, no caso da variável aleatória ser contínua, a r-ésima $(r = 1,2,\dots)$ derivada da função geradora de momentos em relação a t é:

$$\frac{d^r m(t)}{dt^r} = m^{(r)}(t) = \frac{d^r}{dt^r} E(e^{tX}) = \int_{-\infty}^{+\infty} x^r e^{tX} f(x)dx = E(X^r e^{tX})$$

Para o caso da variável aleatória ser discreta, obtém-se o mesmo resultado.

Se nesta última igualdade fizermos $t = 0$, obtemos.

$$E(X^r) = \frac{d^r m(t)}{dt^r}\Big|_{t\,=\,0} = \mu_1^{'}$$

Deste modo, a derivada de ordem r da função geradora de momentos de X nada mais é do que o momento de ordem r da variável aleatória X, sendo esta a razão da atribuição do nome função geradora de momentos à quantidade $m(t)$. Em particular, concluímos que.

$$E(X) = \frac{dm(t)}{dt}\Big|_{t\,=\,0} = \mu_1^{'}$$

Exemplo de aplicação

Seja X uma variável aleatória que toma todos os valores inteiros não negativos $(0,1,2,3,\dots)$, sendo $P(X = K) = \frac{[e^{-\lambda}][\lambda^k]}{[k!]}(\lambda > 0)$. Esta distribuição de probabilidade chama-se distribuição de Poisson. Determine a função geradora de momentos de X.

Vejamos a resposta para este problema.

Por definição temos:

$$m(t) = \sum_{k=0}^{\infty} \frac{e^{kt} e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = exp(\,\lambda(e^t - 1)\,)$$

Função característica

A função geradora de momentos possui o inconveniente de não existir para todas as distribuições de probabilidade, pois a esperança matemática de e^{tX} nem sempre é definida. Devido a esse fato, surgiu a necessidade de definir outra função, possuindo propriedades análogas e denominada função caraterística. No entanto o emprego desta função necessita do conhecimento da teoria das funções de variáveis complexas, pois trata-se da esperança matemática de e^{itX}. Para o caso de variáveis aleatórias discretas ou contínuas temos as seguintes funções respectivamente.

$$E\left(e^{itX}\right) = \sum_{x=0}^{\infty} e^{itX} P_x \text{ ou } E\left(e^{itX}\right) = \int_{-\infty}^{+\infty} e^{itX} f(X) dX$$

Sobre esse assunto poderão ser encontradas informação complementares nas obras de HOEL (1965), MOOD e GRAYBILL (1963) E WILKS (1962).

VARIÁVEIS ALEATÓRIAS BIDIMENSIONAIS E DISTRIBUIÇÃO CONJUNTA DE PROBABILIDADE DE DUAS VARIÁVEIS ALEATÓRIAS (BIDIMENSIONAL)

Variável aleatória bidimensional $[(X, Y)]$

Sejam E um experimento aleatório e S um espaço amostral associado a E sejam $x = x(s)$ e $y = y(s)$ duas funções, cada uma associando um número real a cada resultado s de S; denominaremos (X, Y) uma variável aleatória bidimensional.

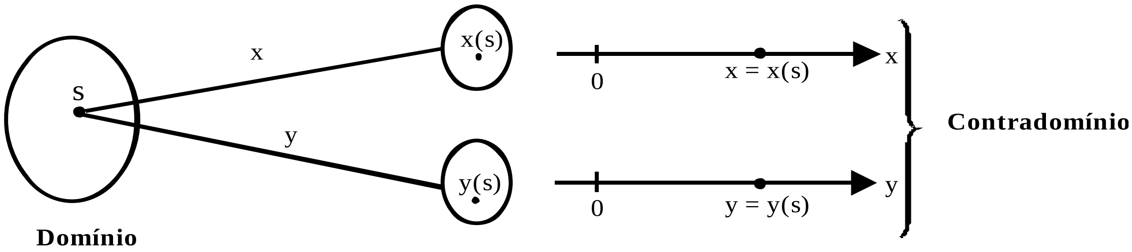


Figura 15. Espaço amostral e conjunto de definição de duas variáveis aleatórias X e Y.

Por exemplo: Considere a tomada de observações em 3 plantas de milho.

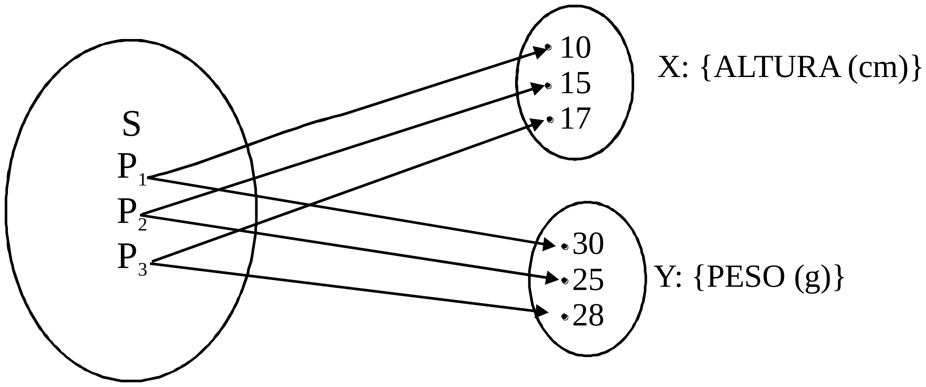


Figura 16. Espaço amostral e conjunto de definição de duas variáveis aleatórias X e Y: a tomada de observações de altura X em centímetro e o peso Y em gramas, em 3 plantas de milho.

$P_1 \rightarrow (10; 30)$

$P_2 \rightarrow (15; 25)$

$P_3 \rightarrow (17; 28)$

Então o par (X, Y) é definido como variável aleatória bidimensional.

Tipos de variáveis aleatórias bidimensionais

(i) Variável Aleatória Bidimensional Discreta (V.A.B.D.):

São aquelas cujo campo de definição no espaço bidimensional $\mathbb{R}^2$ e um conjunto de pontos finitos ou infinitos enumeráveis, por exemplo: número de grãos/ espiga; número de ramificações do pendão de milho, número de insetos por planta e número de folhas atacadas nessas plantas etc.

(ii) Variável Aleatória Bidimensional contínua (V.A.B.C.)

São aquelas definidas num sub-plano do espaço dimensional $\mathbb{R}^2$ cujos pontos são infinitos, como por exemplo, o peso da parte aérea e a altura de plantas de feijão, idade e peso de bovinos da raça Nelore, etc.

Variável aleatória bidimensional (x,y) discreta

Função de probabilidade conjunta

Seja X uma variável aleatória que assume os valores $x_1, x_2, \dots, x_m$ e y uma variável aleatória que assume os valores $y_1, y_2, \dots y_n$.

Definição. A função de probabilidade conjunta associa a cada par $(x_i, x_j) i = 1, \dots, m$, e $j = 1, \dots, n$ a probabilidade $P(x = x_i, y = y_j) = P(x_i, y_j)$.

Distribuição de probabilidade conjunta

Damos o nome de distribuição conjunta de probabilidade da variável bidimensional (X, Y) ao conjunto $(x_i, y_j), P(x_i, y_j), i = 1, \dots, m e j = 1, \dots, n$. Verifica-se que:

(i) $\sum_{i=1}^{m} \sum_{j=1}^{n} P(x = x_i, y = y_j) = 1$;

(ii) $P(X_i; Y_j) \geq 0, \forall (x, y) \in$ ao conjunto dos Reais ao Quadrado $\mathbb{R}^2$, sendo a probabilidade conjunta bidimensional um valor compreendido entre zero e um conforme o primeiro axioma do conceito moderno de probabilidade, ou seja, tem-se que $[0 \leq P(x_i, y_j) \leq 1]$.

Um exemplo de distribuição de probabilidade discreta bidimensional pode ser visto abaixo: onde X é o valor dos salários em reais e Y o tempo de serviço de empregados em uma agroindústria.

TABELA 10. DISTRIBUIÇÃO BIDIMENSIONAL DE PROBABILIDADE DOS VALORES X E Y ONDE X É O VALOR DOS SALÁRIOS EM REAIS E Y O TEMPO DE SERVIÇO DE EMPREGADOS EM UMA AGROINDÚSTRIA.

X \ *Y*	**4**	**5**	**6**	**Totais das linhas**
140	0	0	1/10	1/10
150	0	2/10	1/10	3/10
160	0	1/10	2/10	3/10
170	1/10	0	2/10	3/10
Totais das colunas	1/10	3/10	6/10	1

A representação gráfica da variável bidimensional discreta (X, Y) é:

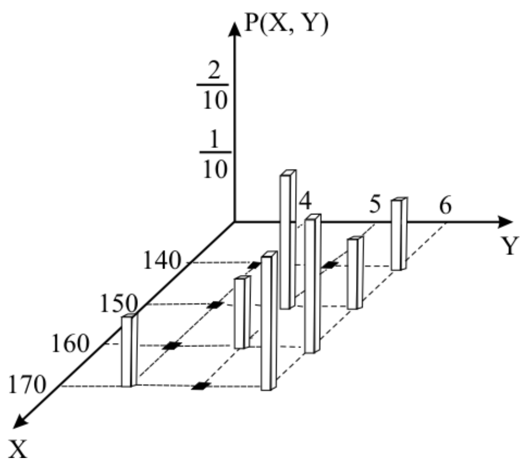


Figura 17. Gráfico representativo da distribuição de probabilidade conjunta da variável aleatória bidimensional discreta (XY).

Função de repartição conjunta discreta

$$F(x_i, y_j) = P(x \leq x_i, y \leq y_j) = \sum_{x=0}^{x \leq x_i} \sum_{y=0}^{y \leq y_j} P(x_i, y_j) \quad \forall \text{ (x,y) } \in \mathbb{R}^2$$

Que dará a probabilidade acumulada conjunta de x tomar valores menores ou iguais a um determinado valor x_i e y assumir valores até um especificado valor y_j como mostra a Figura abaixo:

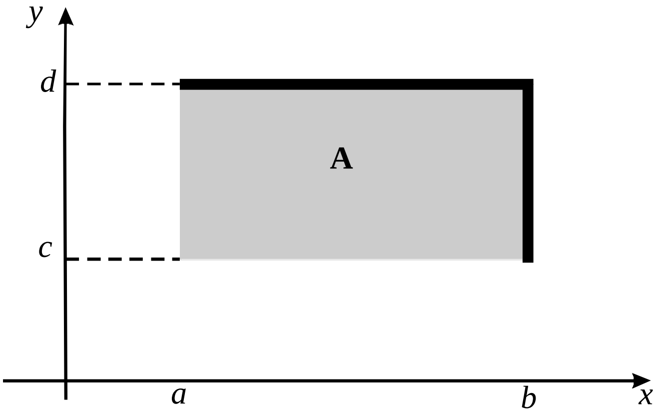


Figura 18. Probabilidade conjunta tal que $P(a < X \leq b, c < Y \leq d)$.

Variáveis aleatórias independentes

Seja (x, y) uma variável aleatória discreta bidimensional. Diremos que x e y são variáveis aleatórias independentes, se e somente se, $P(X_i, Y_j) = P(x_i)P(y_j)$ para quaisquer i e j isto é, $P(x = x_i; y = y_j) = P(x = x_i)P(y = y_j)$ para todo i e j.

Basta que o resultado anterior não se verifique para um par (x, y), para que x e y não sejam consideradas variáveis aleatórias independentes. Nesse caso diremos que x e y são variáveis aleatórias dependentes, essa definição pode ser estendida para mais de duas variáveis aleatórias.

Variável aleatória bidimensional (x, y) contínua

Distribuição de probabilidade conjunta de duas variáveis aleatórias contínuas:

$$[(XY);\ f(x, y)]$$

Função de densidade de probabilidade conjunta:

Seja (X, Y) uma variável aleatória bidimensional contínua, e $f(x, y)$ é uma função de densidade de probabilidade conjunta se:

i) $f(x, y) \geq 0$, $\forall (x, y) \in \mathbb{R}^2$

ii) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, dxdy = 1$

Como mostra a figura a seguir:

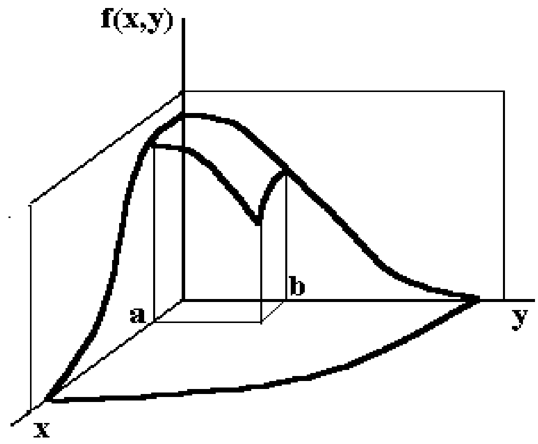


Figura 19. Gráfico representativo do volume da probabilidade conjunta da variável aleatória contínua (x, y) assumir valores menores ou iguais a a e b e $f(x, y)$.

Função de repartição ou de distribuição conjunta contínua

A probabilidade acumulada nesse caso será dada pela função abaixo:

$$F(X,Y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(s,t)dsdt, \forall (x,y) \in \mathbb{R}^2$$

Que fornecerá a probabilidade ou o volume da área delimitada por x e y.

Propriedades de $F(x, y)$

i) Valores compreendidos entre 0 e 1: $0 \leq F_{x,y}(x,y) \leq 1, \forall (x,y) \in \mathbb{R}^2$

ii) Função não decrescente: $F_{x,y}(x + \Delta x, y + \Delta y) \geq F_{x,y}(x,y)\ \forall \Delta x, \Delta y \geq 0$

iii) O limite superior é 1: $\lim_{x,y \to +\infty} F_{x,y}(X,Y) = 1$

iv) O limite inferior para qualquer uma das variáveis é zero: $\lim_{x,y \to -\infty} F_{x,y}(X,Y) = 0$ e $\lim_{x,y \to -\infty} F_{x,y}(x,y) = 0$

v) $F_{x,y}(x + \Delta x, y + \Delta y) - F_{x,y}(\Delta x, \Delta y) = F_{x,y}(x,y), \forall (\Delta x, \Delta y \geq 0)$,

$\frac{\partial F_{xy}(x,y)}{\partial x\, \partial y} = F_{x,y}(x,y), \forall (x,y) \in \mathbb{R}^2$

Variáveis aleatórias independentes

Seja (x, y) uma variável aleatória contínua bidimensional diremos que X e Y são variáveis aleatórias independentes se, e somente se, $f(x,y) = g(x)h(y)$ para todo (x, y), onde f é a função densidade de probabilidade [f.d.p.], e g e h são as funções densidades de probabilidade [f.d.p.] marginais de x e y, respectivamente.

DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL

Dada uma distribuição conjunta de duas variáveis aleatórias (x, y), podemos determinar as distribuições de x sem considerar y ou vice-versa, as quais são designadas funções de probabilidade marginais.

A palavra marginal é usada em virtude de, no caso discreto, os valores de $P(X = x)$ e de $P(Y = y)$ serem os totais das colunas e das linhas que se podem escrever nas margens da tabela de probabilidade conjunta.

Distribuição marginal de X para a variável aleatória discreta (x, y)

$$P(x = x_i) = P\{x = x, \text{para qualquer } y\} = P\{x = x, -\infty < Y < +\infty\} = \sum_{y \in D_y} P(X, Y)$$

$$P(x = x_i) = \sum_i P(x_i, y_i)$$

Distribuição marginal de Y para a variável aleatória discreta (x, y)

$$P(y = y_i) = P\{y = y, \text{para qualquer } x\} = P\{y = y, -\infty < X < +\infty\} = \sum_{x \in D_x} P(X, Y)$$

$$P(y = y_i) = \sum_i P(x_i, y_i)$$

Distribuição marginal de x para a variável aleatória (x, y) contínua.

$g(x) = \int_{-\infty}^{\infty} f(x, y)dy$, a qual é a função densidade marginal de x.

Distribuição marginal de y para a variável aleatória (x, y) contínua

$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$, a qual é a função densidade marginal de y.

Verifique que $g(x)$ e $h(y)$ são de fato densidades de probabilidade, pois são não negativas e sua integral na reta é igual a um.

Função de distribuição ou de repartição marginal de x para a variável aleatória (x, y) discreta

$$F(x) = P(x \leq x;\ y = qualquer) = \sum_{u \leq x} \sum_{y} P(U, Y)$$

Função de distribuição ou de repartição marginal de y para a variável aleatória (x, y) discreta

$$F(y) = P(y \leq y;\ x = \text{qualquer}) = \sum_{u \leq x} \sum_{x} P(X, U)$$

Função de distribuição ou de repartição marginal de x para a variável aleatória (x, y) contínua

$$F(X) = \int_{-\infty}^{X} \int_{-\infty}^{+\infty} f(u, y) \ dydu$$

Função de distribuição ou de repartição marginal de y para a variável aleatória (x, y) contínua

$$F(Y) = \int_{-\infty}^{Y} \int_{-\infty}^{+\infty} f(x, u) \ dxdu$$

Distribuições de probabilidade condicionais para variáveis aleatórias discreta

i) Para a variável aleatória X

Considere-se a variável aleatória bidimensional (X, Y) discreta, com função de probabilidade conjunta $f(X, Y)$. A função de probabilidade de X condicionada pela realização do acontecimento $\{Y = y\}$, com $P(Y = y) > 0$, é dada por:

$$P(X = x/Y = y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{P(X,Y)}{P(Y)} \quad (\text{Para } Y \text{ fixo}) \text{ e } P(Y) > 0$$

ii) Para a variável aleatória Y

Considere-se a variável aleatória bidimensional (X, Y) discreta, com função de probabilidade conjunta $f(X, Y)$. A função de probabilidade de Y condicionada pela realização do acontecimento $\{X = x\}$, com $P(X = x) > 0$, é dada por:

$$P(Y = y/X = x) = \frac{P(X=x, Y=y)}{P(X=x)} = \frac{P(X,Y)}{P(X)} \quad (\text{Para } X \text{ fixo}) \text{e } P(X) > 0$$

Distribuições de probabilidade condicionais para variáveis aleatórias contínuas

i) Para a variável aleatória X

A densidade condicional de X, dado $Y = y$ denominada por $f(x/y)$, é dada para cada x fixo por: $f(x/y) = \frac{f(x,y)}{f(y)}$, se $f(y) > 0$. Dessa igualdade obtém-se: $f(x, y) = f(y)f(x/y)$.

Verifica-se assim, que conhecendo-se a densidade de probabilidade de Y e a densidade condicional de X dado Y, obtém-se a densidade conjunta de X e Y.

ii) Para a variável aleatória Y

A densidade condicional de Y, dado $X = x$ denominada por$f(y/x)$, é dada para cada x fixo por: $f(x/y) = \frac{f(x,y)}{f(x)}$, se $f(x) > 0$. Dessa igualdade obtém-se:$f(x, y) = f(x)f(y/x)$.

Verifica-se assim, que conhecendo-se a densidade de probabilidade de X e a densidade condicional de Y dado X, obtém-se a densidade conjunta de X e Y.

TEOREMA DE TCHEBYCHEFF (EM INGLÊS: BIENAYMÊ-TCHEBYCHEFF INEQUALITY)

Esse teorema é assim chamado em homenagem ao seu criador o matemático Russo, Pafnuty Lyovich Chebyshev (1821 – 1894).

Generalidades

A desigualdade de Tchebycheff é fundamental para o entendimento de como a variância de uma variável aleatória mede a variabilidade desta variável em torno da média (esperança matemática) $[E(x)]$. O conhecimento da média e da variância embora seja extremamente útil, não é suficiente para determinar sua distribuição, não permitindo assim responder a perguntas tais como "Quanto é $P(X \leq 5)$?" todavia como veremos a seguir é possível estabelecer cotas para a probabilidade, com o conhecimento apenas da média e da variância.

Para qualquer conjunto de dados (população ou amostra) e qualquer constante $k > 1$, a proporção dos dados que podem estar a menos de k desvios padrões da média (para qualquer dos dois lados) é pelo menos: $1 - \frac{1}{k^2}$, isto é,

$$P(\mu = \text{k}\sigma < X_i < \mu + \text{k}\sigma) \geq 1 - \frac{1}{k^2} \text{ ou } \text{P}(\bar{X} - \text{kS} < \text{X}_i < \bar{\text{X}} + \text{kS}) \geq 1 - \frac{1}{k^2}$$

Ou de outra maneira temos que para $k = a$ e o módulo $|x - \mu_x|$ a desigualdade de Tchebycheff afirma que: $P\left(|x - \mu_x| \leq a\sigma_x\right) \geq 1 - \dfrac{1}{a^2} \therefore a \geq 1 \therefore a = \frac{|x - \mu_x|}{\sigma_x}$.

Significa que, a probabilidade de que qualquer variável aleatória esteja dentro de a desvios padrões de sua média é no mínimo. $\left(1 - \frac{1}{a^2}\right)$ esse teorema é de grande importância em probabilidade e estatística, pois revela uma propriedade geral de variáveis aleatórias contínuas ou discretas com média e variância finitas.

O teorema pode ser também denominado de desigualdade de BIENAYMÊ-TCHEBYCHEFF a qual é uma propriedade da variância e do desvio padrão, que permite uma interpretação geral.

Seja qual for a distribuição da variável X de média m e de desvio padrão a e seja qual for a quantidade positiva k, é possível demonstrar que:

$$P(|X - m| \geq k\sigma) = P\left(\frac{|X - m|}{\sigma} \geq k\right) \leq \frac{1}{k^2}$$

Esta propriedade mostra, portanto, que a proporção de indivíduos que se afastam da média de k vezes o desvio padrão é sempre inferior a $1/k^2$.

Faremos a demonstração unicamente para o caso contínuo, mas vale também para o caso envolvendo variáveis aleatórias discretas. Designando por c uma constante positiva e por D o domínio no qual $(X-m)^2 \leq c$, temos:

$$\sigma^2 = E(X-m)^2 = \int_{-\infty}^{+\infty} (X-m)^2 f(x)dx \geq \int_D cf(x)dx = c\int_D f(x)dx$$

$$\sigma^2 = cP[(X-m)^2 \geq c]$$

Além disso, fazendo $c = k^2\sigma^2$, podemos escrever:

$$P[(x-m)^2 \geq c] = P(|X-m| \geq k\sigma) \geq \sigma^2/c = 1/k^2$$

Daí se deduz, por exemplo, que seja qual for a distribuição considerada, nunca pode haver mais de um quarto da probabilidade dos desvios em relação à média maiores do que duas vezes o desvio padrão, isto é, dos desvios reduzidos $(X-m)/\sigma$ maiores do que 2, em valor absoluto. Contudo, esta proporção pode ser muito mais fraca para certas distribuições ou para certos tipos de distribuições. Para uma variável uniforme variando entre 0 e 1, por exemplo, o desvio reduzido, máximo é $\sqrt{3}$, de tal modo que o valor 2 nunca é atingido nem ultrapassado neste caso:

$$P\left(\frac{|X-m|}{\sigma} \geq 2\right) = 0$$

De outra forma, também é possível mostrar que a seguinte desigualdade é válida para todas as distribuições contínuas simétricas possuindo um só máximo de densidade de probabilidade (distribuições unimodais).

$$P(|X-m| \geq k\sigma = P\left(\frac{|X-m|}{\sigma} \geq k\right) \leq 4/9k^2$$

Para este tipo de distribuições, a probabilidade de o valor $k = 2$ ser ultrapassada é sempre inferior a $1/9$, e não apenas a $1/4$.

Para ilustrar o teorema de Tchebichev, por exemplo, é possível afirmar que ao menos:

$$1 - \frac{1}{2^2} = \frac{3}{4} = 75\%$$

dos valores de qualquer conjunto de dados, devem estar a menos de dois desvios padrões da média, de qualquer lado dela. Esse teorema aplica-se a todos os conjuntos de dados e fornece apenas o resultado de uma proporção mínima de dados que devem estar entre dois valores. Para quase todos esses conjuntos, a porcentagem é quase sempre maior. Conhecendo-se um pouco mais de natureza da distribuição dos dados desses conjuntos, afirmações mais fortes e precisas dessa proporção podem ser feitas.

De acordo com FERREIRA (2005) uma outra relação interessante envolvendo a média, mediana e desvio padrão foi apresentada por O'CINNEIDE (1990). A mediana populacional nunca é maior que um desvio padrão da média, ou seja,

$$|\mu_d - \mu| \leq \sigma$$

Esta relação é um caso especial, quando $p = 0{,}5$, da relação:

$$\mu - \sigma\sqrt{\frac{1-p}{p}} \leq X_p \leq \mu + \sigma\sqrt{\frac{p}{1-p}}$$

proposta, posteriormente, por DHARMADHIKAR (1991).

Exercício de aplicação

Exercício 1. Fazendo $k = 2$ na desigualdade de Tchebichev, vemos que

$$P(|x - \mu| \geq 2\sigma \leq 0{,}25 \text{ ou } P(|x - \mu| < 2\sigma) \geq 0{,}75$$

Em palavras, a probabilidade de x diferir de sua média em mais de dois desvios padrões é inferior ou, no máximo, igual a 0,25; equivalente, a probabilidade de X estar compreendido no intervalo de dois desvios padrões a contar de sua média, é superior ou, no mínimo, igual a 0,75. Tal fato é tanto mais notável quanto se considerarmos que nem sequer especificamos a distribuição de probabilidade de X.

Exercício 2. Uma fabricação de fitas, temos que $\mu_x = 3m$ e $\sigma_x = 0{,}03m$

$$P(|x - 3| \leq a \,.\, 0{,}03) \geq 1 - \frac{1}{a^2}, \text{se } a = 2$$

$$P(|x - 3| \leq 0{,}06) \geq 1 - \frac{1}{4} \therefore P(|x - 3| \leq 0{,}06) \geq \frac{3}{4}$$

$$P(2{,}94 \leq X \leq 3{,}06) \geq \frac{3}{4} = 0{,}75$$

Quer dizer que a probabilidade de uma fita de $3\ m$ ter erro não superior a $\pm\, 0{,}06\ m$ é no mínimo $0{,}75$.

Teorema de Tchebycheff

Introdução

Uma outra forma de dedução para a demonstração do teorema de Tchebycheff também é realizado como descrito abaixo.

Suponha uma variável aleatória x que assume os valores $x_1, x_2, ...$, sem suas respectivas probabilidades de ocorrências $p_1, p_2, ...$. Considere ainda que essa variável tem uma média $m = E\,(x)$ e uma variância $\sigma^2 = E[(x-m)^2]$.

É comum adotar nas distribuições o desvio padrão σ como unidade de medida. Um desvio $x - m$, em valor absoluto e medido em relação a σ resulta num número real positivo a, isto é,

$$\frac{|x-m|}{\sigma} = \mathrm{a}\ ,\ \therefore\ |\mathrm{x}-\mathrm{m}| = \mathrm{a}\sigma$$

Com a seguinte equação $P[|x - m| < a\sigma]$ representa-se a probabilidade total de ocorrências cujos desvios em relação à média são, em valor absoluto, menores do que número real positivo a vezes o sigma σ. Uma restrição abaixo para essa probabilidade é dada pelo teorema de Tchebycheff, o qual afirma que a probabilidade de um desvio dado por $|x - m| < a\sigma$ é maior ou igual do que $1-(1/a^2)$, isto é,

$$P[|x - m| < a\ \sigma] \geq 1 - \frac{1}{a^2}$$

Demonstração:

A demonstração deste teorema é dada a seguir.

Sabe-se que

$$\sigma^2 = \mu_2 = E[(x-m)^2] = \Sigma(x_i - m)^2 p_i$$

Este somatório pode ser dividido em dois. Um, que indicaremos por Σ', referente aos termos para os quais tenhamos $|x_i - m| < a\sigma$; outro, representado por Σ'', incluindo os termos restantes, isto é, para os quais seja $|x_i - m| \geq a\sigma$. Sendo assim tem-se que:

$$\sigma^2 = \Sigma'(x_i - m)^2 p1 + \Sigma''(xi - m)^2 p_i,$$

logo

$$\sigma^2 \geq \Sigma''(x_i - m)^2 p_i$$

Como em Σ'' tem-se sempre $|x_i - m|^2 \geq a^2\sigma^2$, pode-se escrever ainda:

$$\sigma^2 \geq a^2\sigma^2\Sigma'' p_i,$$

$$\therefore\ \frac{1}{a^2} \geq \Sigma'' p_i$$

No entanto sabe-se que,

$$\Sigma'' p_i = P[|x-m| \geq a\sigma] = \Sigma'' p_i = 1 - p[|x-m| < a\sigma].$$

Obtém-se assim,

$$\frac{1}{a^2} \geq\ 1 - P[|x-m| < a\sigma],$$

Ou seja,

$$P[|x - m| < a\sigma\} \geq 1 - \frac{1}{a^2},$$

O que prova o teorema, que é válido também, com demonstração semelhante a esta, para o caso de variável aleatória contínua.

Desigualdade de Camp-Meidell

Se a distribuição de probabilidade possuir uma única moda isto e, for unimodal e também simétrica, então: $P(|X - \mu| \geq k\sigma) \leq \frac{4}{9}K^2$. (Desigualdade de Camp-Meidell). Estas desigualdades fornecem as probabilidades de que os valores de uma variável aleatória (qualquer) estejam num intervalo simétrico em torno da média de amplitude igual a k desvios padrões. Assim se $k = 2$, por exemplo, a desigualdade de Tchebycheff estabelece que o percentual de valores da variável aleatória que está compreendida no intervalo $\mu \pm 2\sigma$ é de pelo menos $1 - \frac{1}{4} = 75\%$. Conforme é mostrado na distribuição normal de probabilidades este percentual vale aproximadamente 95,44%. Mas como a normal é simétrica e unimodal, neste caso, um resultado mais próximo é dado pela desigualdade de Camp-Meidell, isto é, $1 - \frac{4}{9}k2 = 1 - \frac{1}{9} = 88{,}89\%$.

Uma aplicação do teorema de Tchebycheff

O teorema de THEBYCHEV que acabamos de ver é da mais alta importância, pela sua grande generalidade e pelas suas consequências de grande alcance, como mostram as considerações seguintes.

Seja uma variável casual x, da média m e variância σ^2 e para ela obtenhamos as observações $x_1, x_2, \ldots x_n$, às quais corresponderá a estimativa da média

$$\widehat{m} = \frac{x_1 + x_2 + \ldots + x_n}{n},$$

cuja variância será, $V(\widehat{m}) = \frac{\sigma^2}{\sqrt{n}}$, temos, pois, $\sigma(\widehat{m}) = \frac{\sigma}{\sqrt{n}}$ de sorte que o teorema da Tchebycheff nos dá:

$$P[|\widehat{m} - m| < a\sigma(\widehat{m})] \geq 1 - \frac{1}{a^2}$$

isto é,

$$P[|\widehat{m} - m| < a\frac{\sigma}{\sqrt{n}}] \geq 1 - \frac{1}{a^2}$$

Tomando $a = 100$, por exemplo, o que corresponde a fixar $1 - \frac{1}{a^2} = 0{,}9999 = 99{,}99\%$. Fixemos ainda, arbitrariamente $\frac{a\sigma}{\sqrt{n}} = \frac{100\sigma}{\sqrt{n}} = 0{,}1$ o que nos dá $n = 1.000.000$.

Isto significa que a probabilidade de que o desvio ou afastamento $\widehat{m} - m$, entre a média aritmética $\widehat{m}$ da distribuição seja, em valor absoluto, menor do que $0{,}1\sigma$, é maior do que $99{,}99\%$. E os valores dos desvios poderão ser diminuídos à vontade, e aumentada indefinidamente, aproximando-se de um cada vez mais a probabilidade, com a condição de que aumentem suficientemente o valor do número de observações n. É esta propriedade que favorece o uso da média aritmética das observações como estimativa da média de distribuição. É por isso que se denomina Lei dos Grandes Números, de maneira muito frequente ao teorema de Tchebycheff.

Exercícios de aplicação

1. Numa Empresa Agrícola localizada em Mossoró, RN, Brasil, verificou-se durante o ano de 2022, que o número de ausências semanais causadas pelos empregados tem uma média $m = 52{,}4$ e um desvio padrão $\sigma = 4{,}7$. Avaliar, pelo teorema de Tchebycheff, a probabilidade de o número de faltas numa semana não diferir da média mais de 12 unidades.

2. Na mesma firma do problema anterior, avaliar pelo teorema de Tchebycheff a probabilidade de termos na próxima semana, 65 faltas ou mais.

3. Na mesma firma mencionada nos dois problemas anteriores, avaliar a probabilidade de o número semanal médio de faltas, em 100 semanas, estar entre 49,4 e 55,4.

Exemplos de aplicação

Exemplo 1. Uma região (Nordeste brasileiro) foi subdivida em 10 sub-regiões. Em cada uma delas, foram observadas duas variáveis: número de poços artesianos (x) e número de riachos ou rios presentes na sub-região (y). Os resultados são apresentados na Tabela a seguir:

TABELA 11. DISTRIBUIÇÃO DO NÚMERO DE POÇOS ARTESIANOS (x) E NÚMERO DE RIACHOS OU RIOS PRESENTES NA SUB-REGIÃO (y).

Sub-região	1	2	3	4	5	6	7	8	9	10
X	0	0	0	0	1	2	1	2	2	0
Y	1	2	1	0	1	0	0	1	2	2

Considerando que escolhemos uma das sub-regiões ao acaso, isto é, cada sub-região tem mesma probabilidade $\frac{1}{10}$ de ser escolhida, podemos construir a distribuição conjunta de (X, Y).

TABELA 12. DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA BIDIMENSIONAL X E Y ONDE X É O NÚMERO DE POÇOS ARTESIANOS, E Y O NÚMERO DE RIACHOS OU RIOS PRESENTES NA SUB-REGIÃO.

(XY)	**PROBABILIDADE**
(0,0)	1/10
(0,1)	2/10
(0,2)	2/10
(1,0)	1/10
(1,1)	1/10
(2,0)	1/10
(2,1)	1/10
(2,2)	1/10
TOTAL	1

Note que pares idênticos foram agrupados e somamos as respectivas probabilidades. Uma forma equivalente de apresentar a distribuição conjunta, porém com maior apelo visual, é através da tabela de dupla entrada. E ainda as distribuições marginais também podem aparecer na tabela, bastando efetuar a soma das linhas para obter a marginal de X e, nas colunas, para a marginal de Y.

Por exemplo, para calcular a probabilidade de X ser igual a zero, temos:

$$P(x = 0) = P(x = 0, y = 0) + P(x = 0, y = 1) + P(x = 0, y = 2)$$

$$P(x = 0) = \frac{1}{10} + \frac{2}{10} + \frac{2}{10} = \frac{5}{10}$$

Repetindo os cálculos para outros valores se X e Y, obtemos a tabela a seguir:

TABELA 13. DISTRIBUIÇÃO DE PROBABILIDADE BIDIMENSIONAL CONJUNTA DA VARIÁVEL ALEATÓRIA (X, Y), ONDE X É O NÚMERO DE POÇOS ARTESIANOS E Y O NÚMERO DE RIACHOS OU RIOS PRESENTES NA SUB-REGIÃO.

X \ Y	**0**	**1**	**2**	$P(X = x)$
0	1/10	2/10	2/10	5/10
1	1/10	1/10	0	2/10
2	1/10	1/10	1/10	3/10
P(Y = y)	3/10	4/10	3/10	1,00

Portanto, as funções de probabilidades marginais (distribuições marginais de probabilidades) são as seguintes:

TABELA 14. DISTRIBUIÇÃO MARGINAL DE PROBABILIDADE DE X: ONDE X É O NÚMERO DE POÇOS ARTESIANOS.

X	0	1	2	Total
P(x)	5/10	2/10	3/10	1,00

e

TABELA 15. DISTRIBUIÇÃO MARGINAL DE PROBABILIDADE DE Y: ONDE Y É O NÚMERO DE RIACHOS OU RIOS PRESENTES NA SUB-REGIÃO.

Y	0	1	2	Total
P(y)	3/10	4/10	3/10	1,00

Exemplo 2. O acréscimo anual na área atingida por uma certa praga (mosca das frutas), numa região produtora de frutas, pode ser modelado por uma variável aleatória contínua, medida em hectares (10 mil m^2), com função densidade de probabilidade dada por:

$$f(x) = \begin{cases} \frac{2}{3}x, & \text{se o} < x < 1 \\ 1-\frac{x}{3}, & \text{se } 1 \leq x < 3 \\ 0, & \text{caso contrário} \end{cases}$$

a) Construa o gráfico dessa função densidade de probabilidade.
b) Qual seria a probabilidade da praga atingir entre 2 e 3 hectares esse ano?
c) Que área será atingida com 50% de certeza?
d) Determine o acréscimo médio anual na área atingida pela praga.

Exemplo 3. (X, Y) é uma variável aleatória bidimensional discreta com a seguinte distribuição conjunta: Pede-se calcular.

a. A $COV(X, Y)$ e
b. $r(X, Y)$

TABELA 16. DISTRIBUIÇÃO CONJUNTA BIDIMENSIONAL DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA (X, Y).

X \ Y	-3	2	4	TOTAL
1	0,1	0,2	0,2	0,5
3	0,3	0,1	0,1	0,5
TOTAL	0,4	0,3	0,3	1,00

Primeiro encontramos as distribuições marginais de x e y assim:

TABELA 17. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL DAS VARIÁVEIS ALEATÓRIAS X E Y.

X	1	3
$P(X)$	0,5	0,5

Y	-3	2	4
$P(Y)$	0,4	0,3	0,3

(a) $E(X) = \mu_x = 1 \cdot 0{,}5 + 3 \cdot 0{,}5 = 2{,}0$ $\quad E(Y) = -3 \cdot (0{,}4) + 2 \cdot (0{,}3) = 4 \cdot (0{,}3) = 0{,}6$

$E(X) = 1^2 \cdot (0{,}5) + 3^2 \cdot (0{,}5) = 5{,}0$ $\quad E(Y^2) = (-3)^2 \cdot (0{,}4) + 2^2 \cdot (0{,}3) + 4^2 \cdot (0{,}3) = 9{,}6$

$$\sigma^2_{(x)} = E(x^2) - [E(x)]^2 = 5 - (2)^2 = 1{,}0 \therefore \sigma_{(x)} = 1{,}00$$

$$\sigma^2_{(y)} = E(y^2) - [E(y)]^2 = 9{,}6 - (0{,}6)^2 = 9{,}24 \therefore \sigma_{(y)} = 3{,}04$$

$$E(\mathrm{XY}) = \sum_i \sum_j X_i Y_j \,.\, \mathrm{P}(X_i Y_j) =$$

$$= 1(-3)\,.\,0{,}1 + 1\,.\,2\,.\,0{,}2 + 1\,.\,4\,.\,0{,}2 + 3(-3)\,.\,0{,}3 + 3\,.\,2\,.\,0{,}1 + 3\,.\,4\,.\,0{,}1$$
$$= 0$$

$$COV(X,Y) = E(X,Y) - \mu x \,.\, \mu y = 0 - 2\,.\,0{,}6 = -1{,}2$$

(b) $r_{(xy)} = \dfrac{E(\mathrm{XY}) - \mu_x \cdot \mu_y}{\sqrt{\sigma^2_{(x)} \cdot \sigma^2_{(y)}}} = \dfrac{E(\mathrm{XY}) - \mathrm{E}(X)\,.\,\mathrm{E}(Y)}{\sqrt{\sigma^2_{(x)} \cdot \sigma^2_{(y)}}} = \dfrac{-1{,}2}{\sigma_x \cdot \sigma_y} = \dfrac{-1{,}2}{1\,.\,3{,}04} \cong -0{,}39$

Portanto um grau de correlação entre as variáveis x e y, fraco e inverso (ou negativo).

Exemplo 4. Considere a distribuição de probabilidade bidimensional discreta dada abaixo. Pede-se, calcular:

a) A covariância;
b) O coeficiente de correlação da distribuição discreta de probabilidade conjunta; e
c) Verificar se x e y são independentes.

Resolução:

(a)

TABELA 18. DISTRIBUIÇÃO DE PROBABILIDADE CONJUNTA BIDIMENSIONAL DA VARIÁVEL ALEATÓRIA (X, Y).

X \ Y	**1**	**2**	**3**	**TOTAL**
2	0,10	0,30	0,20	0,60
3	0,06	0,18	0,16	0,40
TOTAL	0,16	0,48	0,36	1,00

TABELA 19. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA X.

X	**2**	**3**	**Total**
$P(X)$	0,60	0,40	1,00

$$E(X) = 2\,.\,0{,}6 + 3\,.\,0{,}4 = 2{,}4$$

$$E(X^2) = 2^2 \cdot 0{,}6 + 3^2 \cdot 0{,}4 = 6{,}0$$

$$V(X) = 6{,}0 - (2{,}4)^2 = 0{,}24$$

TABELA 20. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA Y.

Y	**1**	**2**	**3**	**Total**
$P(Y)$	0,16	0,48	0,36	1,00

$$E(Y) = 1 \cdot 0{,}16 + 2 \cdot 0{,}48 + 3 \cdot 0{,}36 = 2{,}2$$

$$E(Y^2) = 1^2 \cdot 0{,}16 + 2^2 \cdot 0{,}48 + 3^2 \cdot 0{,}36 = 5{,}32$$

$$V(Y) = 5{,}32 - (2{,}2)^2 = 0{,}48$$

$$E(\mathrm{XY}) = \sum_{\mathrm{i}=1}^{m} \sum_{\mathrm{j}=1}^{n} X_i Y_j\,.\,\mathrm{P}(X_i Y_j) =$$

$$= 2\,.\,1\,.\,0{,}10 + 2\,.\,2\,.\,0{,}30 + 2\,.\,3\,.\,0{,}20 + 3\,.\,1\,.\,0{,}06 + 3\,.\,2\,.\,0{,}18$$
$$+ 3\,.\,3\,.\,0{,}16 = 5{,}3$$

$$COV(XY) = 5{,}3 - 2{,}4 \cdot 2{,}2 = 0{,}02$$

(b) $r_{\mathrm{xy}} = \frac{0{,}02}{\sqrt{0{,}24\,.\,0{,}48}} = 0{,}05893 \cong +0{,}06$

(c) X e Y são variáveis aleatórias independentes?

$$P(x = 2;\ y = 1) = 0{,}10$$

$$P(x = 2) = 0{,}6$$

$$P(y = 1) = 0{,}16$$

$$P(x = 2) \cdot P(y = 1) = 0{,}6 \cdot 0{,}16 = 0{,}096$$

Portanto, x e y não são variáveis aleatórias independentes, pois $P(x = 2;\ y = 1) \neq P(x = 2)P(y = 1)$.

Exemplo 5. Dada a distribuição conjunta a seguir, determinar.

a) As distribuições de probabilidades marginais de X e Y;

b) Calcular $E(X), E(Y), V(X), V(Y), COV(X,Y)$ e $r(X,Y)$;

c) Verificar se X e Y são independentes;

d) Calcular a função de distribuição para (função de distribuição conjunta discreta) $X = 0$ e $Y = +1$, $F\ (0{,}1) = P(x \leq 0;\ y \leq 1)$;

e) Calcular a probabilidade de $x \leq 0$ (Função de distribuição marginal de x) $F\ (x) = F(0) = P(x \leq 0;\ y = \text{qualquer})$.

Resolução:

a) e b)

TABELA 21. DISTRIBUIÇÃO DE PROBABILIDADE CONJUNTA BIDIMENSIONAL DA VARIÁVEL ALEATÓRIA (X,Y).

X	Y		$P(x)$
	-1	+1	
-1	0,10	0,25	0,35
0	0,30	0,25	0,35
P(Y)	0,50	0,50	1,00

$$E(XY) = (-1)(-1)(0{,}10) + (-1) \cdot 1 \cdot 0{,}25 + 0 \cdot (-1)(0{,}30) + 0 \cdot 1 \cdot 0 + 1 \cdot (-1) \cdot 0{,}10 + 1 \cdot 1 \cdot 0{,}25 = 0$$

TABELA 22. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL DA VARIÁVEL ALEATÓRIA X.

X	$P(X)$	$XP(X)$	X^2	$X^2P(X)$
-1	0,35	-0,35	1	0,35
0	0,30	0,00	0	0,00
1	0,35	0,35	1	0,35

SOMA	1,00	0,00	-	0,70

$$E(X) = 0,00; V(X) = 0,70 - (0,00)2 = 0,70$$

TABELA 23. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL DA VARIÁVEL ALEATÓRIA Y.

Y	P(Y)	YP(Y)	Y^2	$Y^2P(Y)$
-1	0,50	-0,50	1	0,50
+1	0,50	+0,50	1	0,50
SOMA	1,00	0,00	-	1,00

$$E(Y) = 0,00;\ V(Y) = 1,00 - (0,00)2 = 1,00;\ Cov(xy) = 0 - (0).(0) = 0$$

$$r = \frac{0}{\sqrt{0,70\,.\,1,00}} = 0$$

Assim as variáveis X e Y possuem correlação nula.

$$P(X = 1, Y = 1) = 0,25;\ P(X = 1) = 0,35;\ P(Y = 1) = 0,50$$

c) $P(X = 1)P(Y = 1) = 0,35 \cdot 0,50 = 0,175$. Logo X e Y não são variáveis aleatórias independentes, pois $P(x = 1, y = 1) \neq P(x = 1)P(y = 1)$.

d)

$$F(0,1) = P(x \leq 0;\ y \leq 1)$$

$$F(0,1) = P(x = -1; y = -1) + P(x = 1; y = +1) = P(x = 0; y = -1) + P(x = 0; y = +1)$$

$$F(0,1) = 0,10 + 0,25 + 0,30 + 0$$

$$F(0,1) = 0,65$$

e)

$$F(x) = P(X \leq x;\ Y = qualquer)$$

$$F(0) = P(X \leq 0;\ Y = qualquer)$$

$$F(0) = 0,35 + 0,30$$

$$F(0) = 0,65$$

Exemplo: 6. Considere a seguinte distribuição de probabilidade discreta bidimensional.

TABELA 24. DISTRIBUIÇÃO DE PROBABILIDADE CONJUNTA BIDIMENSIONAL DA VARIÁVEL ALEATÓRIA (X, Y).

X \ Y	**1**	**2**	**3**	**TOTAL**
1	$\frac{1}{72}$	$\frac{1}{36}$	$\frac{5}{72}$	$\frac{8}{72}$
2	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{5}{36}$	$\frac{8}{36}$
3	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{5}{12}$	$\frac{8}{12}$
TOTAL	$\frac{9}{72}$	$\frac{9}{36}$	$\frac{45}{72}$	1,0

Pede-se:

a) A covariância entre X e Y

b) O coeficiente de correlação

Resolução:

a)

$$\mathrm{E(X)} = 1 \cdot \frac{9}{72} + 2 \cdot \frac{9}{36} + 3 \cdot \frac{45}{72} = \frac{5}{2} = 2{,}5$$

$$E(X^2) = 1^2 \cdot \frac{9}{72} + 2^2 \cdot \frac{9}{36} + 3^2 \cdot \frac{45}{72} = \frac{27}{4} = 6{,}75$$

$$\sigma_x^2 = \frac{27}{4} - \left(\frac{5}{2}\right)^2 = \frac{1}{2} = 0{,}5$$

$$\sigma_x = \sqrt{0{,}5} = 0{,}707$$

$$E(Y) = 1 \cdot \frac{8}{72} + 2 \cdot \frac{8}{36} + 3 \cdot \frac{8}{12} = \frac{23}{9} = 2{,}56$$

$$E(Y^2) = 1^2 \cdot \frac{8}{72} + 2^2 \cdot \frac{8}{36} + 3^2 \cdot \frac{8}{12} = 7$$

$$\sigma_y^2 = 7 - (2{,}56)^2 = 0{,}4464 \ ,$$

$$\sigma_y = \sqrt{0{,}4464} = 0{,}668$$

$$E(\mathrm{XY}) = \left(1 \cdot 1 \cdot \frac{1}{72}\right) + \left(1 \cdot 2 \cdot \frac{1}{36}\right) + \left(1 \cdot 3 \cdot \frac{1}{12}\right) + \left(2 \cdot 1 \cdot \frac{1}{36}\right) + \left(2 \cdot 2 \cdot \frac{1}{18}\right) + \left(2 \cdot 3 \cdot \frac{1}{6}\right) + \left(3 \cdot 1 \cdot \frac{5}{72}\right) +$$

$$+ \left(3 \cdot 2 \cdot \frac{5}{36}\right) + \left(3 \cdot 3 \cdot \frac{5}{12}\right) = 6 \cdot \frac{7}{18} = \frac{115}{18} = 6{,}39$$

$$COV(X, Y) = 6{,}39 - 2{,}5 \cdot 2{,}56 = 0$$

b)

$$r_{xy} = \frac{0}{0{,}707 \,.\, 0{,}668} = 0$$

Portanto a correlação entre as variáveis x e y é nula.

Exemplo 7. Dada a função $\mathrm{f(x)} = \&\frac{1}{6}\,\mathrm{x} + \mathrm{k}$ para $0 \leq \mathrm{x}$
$\leq 3\&0$ para x fora deste intervalo

Pede-se:

a) Calcular a constante k, para que $f(x)$ seja uma função de densidade de probabilidade.

$$\int_0^3 f(x)\mathrm{d}x = 1$$

$$\int_0^3 \left(\frac{1}{6}x + k\right)\mathrm{d}x = 1 \therefore \frac{1}{6}\int_0^3 x\mathrm{d}x + \mathrm{k}\int_0^3 \mathrm{d}x = 1$$

$$\frac{1}{6}\left[\frac{x^2}{2}\right]_0^3 + k\,[x]_0^3 = 1 \therefore \frac{1}{2}[9 - 0] + k\,[3 - 0] = 1$$

$$\frac{9}{3} + 3k = 1 \therefore \frac{3}{4} + 3k = 1 \therefore 3 + 12k = 4$$

$$12\,k = 1 \therefore k = \frac{1}{12}$$

b) O gráfico de $\mathrm{f}(x) = \frac{1}{6}x + \frac{1}{12}$ para $0 \leq x \leq 3$.

TABELA 25. DISTRIBUIÇÃO DE PROBABILIDADE DA VARIÁVEL ALEATÓRIA X.

X	$f(x)$
0	1/12
1	3/12
2	5/12
3	7/12

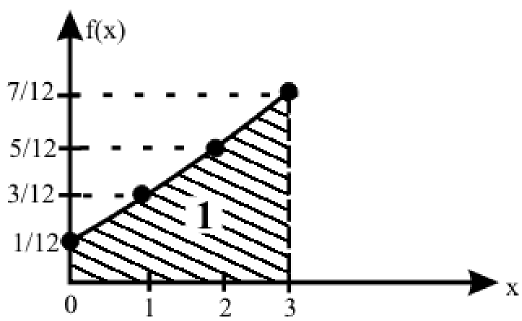


Figura 20. Área sob a curva da distribuição densidade de probabilidade da variável contínua X.

Exemplo 8. Uma variável aleatória contínua X tem a função de densidade:

$$f(x) = \begin{cases} 3x^2 & \text{para -1} \leq x \leq 0 \\ 0 & \text{para outros valores } x \end{cases}$$

Pede-se:

a) A função de distribuição

Para $x < 0$, tem-se $F(x) = \int_{-\infty}^{0} 0\mathrm{dt} = 0$

$$\mathrm{P}(x \leq x) = \mathrm{F}(x) = \int_{-\infty}^{\mathrm{x}} \mathrm{f}(t)\,\mathrm{dt} \therefore \mathrm{F}(x) = 3\int_{-1}^{\mathrm{x}} x^2 \mathrm{d}x$$

$$\mathrm{F}(x) = 3\left[\frac{x^3}{3}\right]_{-1}^{x} \therefore \mathrm{F}(x) = x^3 + 1$$

Para $x > 0$ tem-se.

$$F(x) = \int_{-1}^{0} f(x)\,\mathrm{d}x + \int_{0}^{x} \mathrm{f}(x)\,\mathrm{d}x = \int_{-1}^{0} 3x^2 \mathrm{d}x + \int_{0}^{x} 0 \;\; \mathrm{d}x$$

$$F(x) = \left.\frac{3x^3}{3}\right]_{-1}^{0} + 0 = (0^3 - (-1)^3) = -(-1) = +1$$

Então da função de distribuição é dada por:

$$F(x) = \begin{cases} 0 & \text{para } x < \text{-1} \\ x^3 + 1 & \text{para -1} \leq x \leq 0 \\ 1 & \text{para } x > 0 \end{cases}$$

b) O gráfico de $f(x) = 3x^2$ para $-1 \leq x \leq 0$ é dado por:

c)

TABELA 26. DISTRIBUIÇÃO DE VALORES DA VARIÁVEL ALEATÓRIA CONTÍNUA X, E OS RESPECTIVOS VALORES OBTIDOS PATRA A FUNÇÃO DENSIDADE DE PROBABILIDADE $f(X)$.

$\boldsymbol{X}$	$\boldsymbol{f(x)}$
-1,0	3,00
-0,5	0,75
00	0,00

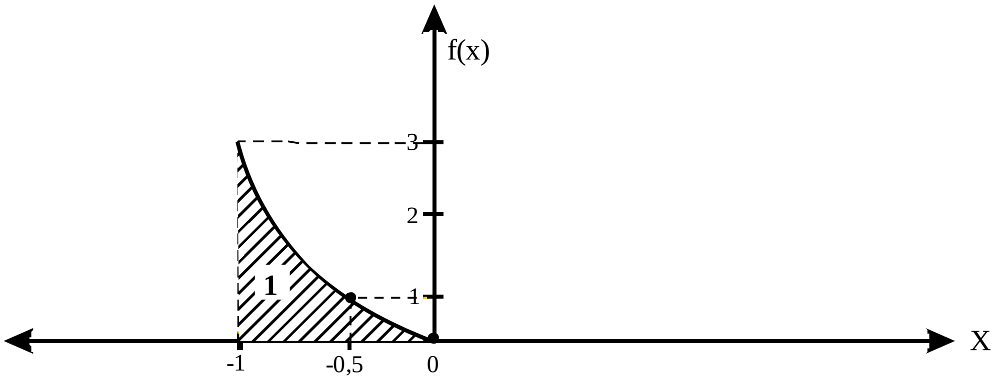


Figura 21. Área sob a curva da distribuição densidade de probabilidade da variável aleatória contínua X.

Exemplo 9. Suponha que a variável aleatória bidimensional (X, Y) tenha a função densidade de probabilidade conjunta.

$$f(x,y) = \begin{cases} kx\,(x-y), & 0 < x > 2, \quad -x < y < x \\ 0 & , \text{ para outros quaisquer valores} \end{cases}$$

a) Calcule a constante k.

[resposta $k = \frac{1}{8}$]

$$\int_0^2 \int_{-x}^{x} \mathrm{k}x\,(x - y)\ \mathrm{d}x\ \mathrm{d}y = 1\mathrm{k} \int_{-x}^{x} \int_0^2 (x^2 - x\mathrm{y})\mathrm{d}x\,\mathrm{d}y = \mathrm{k} \int_{-x}^{x} \left(\frac{x^3}{3} - \frac{x^2}{2} y\right)_0^2 \mathrm{d}x = 1$$

$$\int_{-x}^{x} \mathrm{k}\left[\frac{2^3}{3} - \frac{2^2}{2}\,y\right] \mathrm{d}y = 1 \ \int_{-x}^{x} \mathrm{k}\left[\frac{8}{3} - 2y\right]\ \mathrm{d}y = 1$$

$$\mathrm{k}\left[\frac{8}{3}\,y - \frac{2y^2}{2}\right]_{-x}^{x} = 1\mathrm{k}\left[\frac{8}{3}(x + x) - (x^2 - (-x)^2)\right] = 1$$

$$k\left[\frac{8}{3}\,(2x) - 2x^2\right] = 1 \ \therefore \ \left[-2x^2 + \frac{16x}{3}\right] \mathrm{k} = 1$$

$$-2x^2 + \frac{16x}{3} = \frac{1}{k} \ \therefore \ \mathrm{k} = \frac{1}{8}$$

Sendo assim temos que:

$$f(x,y) = \begin{cases} \frac{1}{8} x\,(x - y), & 0 < x > 2, -x < y < x \\ 0 & , \text{ para outros quaisquer valores} \end{cases}$$

Exemplo 10. A função de densidade conjunta de duas variáveis aleatórias x e y, contínuas é.

$$\mathrm{f}(x,y) = \begin{cases} cxy & 0 < x < 4, 1 < y < 5 \\ 0 & \text{em caso contínuo} \end{cases}$$

a) Determine o valor da constante c.

b) Determine $P(1 < x < 2, 2 < y < 3)$

c) Determine $P(x \geq 3, y \leq 2)$.

Resolução:

a) Devemos ter a possibilidade total igual a 1, isto é,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$$

Utilizando a definição de $f(x,y)$, a integral tem o valor

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\mathrm{d}x\,\mathrm{d}y = c\int_{x=0}^{4}\left[\int_{y=1}^{5} \mathrm{y}x\,\mathrm{d}y\right]\mathrm{d}x$$

$$\therefore\ c\int_{x=0}^{4}\frac{xy^2}{2}\int_{y=1}^{5}\mathrm{d}x = c\int_{x=0}^{4}\left(\frac{25x}{2}-\frac{x}{2}\right)\mathrm{d}x$$

$$\therefore\ c\int_{x=0}^{4} 12x\,\mathrm{d}x = c\,(12\frac{x^2}{2})\ \Big|_{x=0}^{x=4} = c[6\,[(4)^2-(0)^2]] = 1$$

$$\therefore\ \mathrm{C}(96) = 1$$

Então:

$$96c = 1 \text{ e } c = \frac{1}{96}$$

b) Utilizando o valor de c obtido em (a), temos

$$P(1 < x < 2, 2 < y < 3) = \int_{x=1}^{2}\int_{y=2}^{3}\frac{xy}{96}\mathrm{d}x\,\mathrm{d}y$$

$$\mathrm{P}(1 < x < 2, 2 < y < 3) = \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,\mathrm{d}y\right]\mathrm{d}x$$

$$= \frac{1}{96}\int_{x=1}^{2}\frac{xy^2}{2}\Big|_{y=2}^{3}\mathrm{d}x = \frac{1}{96}\int_{x=1}^{2}\frac{x}{2}[3^2-2^2]dx$$

$$\mathrm{P}(1 < x < 2, 2 < y < 3) = \frac{1}{96}\int_{x=1}^{2}\frac{5x}{2}\mathrm{d}x = \frac{5}{192}\left(\frac{x^2}{2}\right)\Big|_{x=1}^{2} = \frac{5}{128}$$

c)

$$\mathrm{P}(x \geq 3, y \leq 2) = \int_{x=3}^{4}\int_{y=1}^{2}\frac{xy}{96}\,\mathrm{d}x\mathrm{d}y$$

$$\mathrm{P}(x \geq 3, y \leq 2) = \frac{1}{96}\int_{x=3}^{4}\left[\int_{y=1}^{2} xy\,\mathrm{d}y\right]\mathrm{d}x = \frac{1}{96}\int_{x=3}^{4}\frac{xy^2}{2}\int_{y=1}^{2}\mathrm{d}x$$

$$\mathrm{P}(x \geq 3, y \leq 2) = \frac{1}{96}\int_{x=3}^{4}\frac{3x}{2}\,\mathrm{d}x$$

$$\mathrm{P}(x \geq 3, y \leq 2) = \frac{1}{96}\left[\frac{3x^2}{4}\right]\Big|_{x=3}^{4} = \frac{1}{96}\left[\frac{3\,.\,4^2}{4} - \frac{3\,.\,3^2}{4}\right] = \frac{1}{96}\left[\frac{48-27}{4}\right] = \frac{21}{384} = \frac{7}{128}$$

Exemplo 11. Verifique se (X, Y) uma variável aleatória bidimensional discreta são independentes.

$$P(x = 2;\ y = 1) = 0{,}10$$

$$P(x = 2) = 0{,}60$$

$$P(y = 1) = 0{,}16$$

$$P(x = 2)\,.\,P(y = 1) = 0{,}60\,.\,0{,}16 = 0{,}096$$

TABELA 27. DISTRIBUIÇÃO CONJUNTA BIDIMENSIONAL DE PROBABILIDADES DA VARIÁVEL ALEATÓRIA (X, Y).

X \ ***Y***	**1**	**2**	**3**	**SOMA**
2	0,10	0,30	0,20	0,60
3	0,06	0,18	0,16	0,40
SOMA	0,16	0,48	0,36	1,00

Logo, X e Y não são variáveis aleatórias independentes, pois, $P(x = 2;\ y = 1) \neq P(x = 2)P(y = 1)$.

Exemplo 12. Exercícios de aplicação resolvidos.

1) E e Y são variáveis aleatórias independentes, ambas com valor esperado 0 e variância σ^2. Consideramos as variáveis $U = aX + bY\ e\ V = cX + dY,\ [E(X) = E(Y) = 0;\ V(X) = V(Y) = \sigma 2]$. Determinar:

i) $E(U)$

ii) $E(V)$

iii) $V(U)$

iv) $V(V)$

v) $COV(U, V)$

vi) $\rho_{UV} = r_{UV}$

Solução:

i) $E(U) = aE(X) + bE(Y) = a(0) + b(0) = 0$

ii) $E(V) = cE(X) + dE(y) = c(0) + d(0) = 0$

iii) $V(U) = V(aX + bY) = a^2V(X) + b^2V(X) = a^2\sigma^2 + b^2\sigma^2(a^2 + b^2)\sigma^2$

iv) $V(cX + dY) = c^2V(X) + d^2V(Y) = c^2\sigma^2 + d^2\sigma^2 = (c^2 + d^2)\sigma^2$

v) $COV(U,V) = E(UV) - E(U)E(V)$

$E(UV) = E[(aX + bY)(cX + dY)]$

$E(UV) = E[acX^2 + adXY + bcXY + bdY^2]$

$E(UV) = acE(X^2) + adE(XY) + bcE(XY) + bdE(Y^2)$

$E(UV) = acE(X^2) + \underbrace{adE(Y)E(E)}_{=0} + \underbrace{bcE(X)E(Y)}_{=0} + bdE(Y^2)$

$E(UV) = acE(X^2) + bdE(Y^2)$

Sabemos que:

$$V(X) = E(X^2) - [E(X)]^2$$
$$V(Y) = E(Y^2) - [E(Y)]^2$$

No caso presente temos:

$$\sigma^2 = E(X^2) - [0]^2 \rightarrow \sigma^2 = E(X^2)$$
$$\sigma^2 = E(Y^2) - [0]^2 \rightarrow \sigma^2 = E(Y^2)$$

Voltando a expressão (2) temos: $E(UV) = acE(x^2) + bdE(y^2)$ temos:

$$E(UV) = ac\sigma^2 + bd\sigma^2 = (ac + bd)\sigma^2$$

Voltando a expressão $COV(U,V) = E(UV) - E(U)\,E(V)$, temos $COV(U,V) = (ac + bd)\sigma^2 - 0 = (ac + bd)\sigma^2$

vi)

$$\rho_{UV} = \frac{\mathrm{COV(U,V)}}{\sqrt{V(U)V(V)}}$$

$$\rho_{UV} = \frac{(ac + bd)\,\sigma^2}{\sqrt{(a^2 + b^2)\sigma^2(c^2 + d^2)\sigma^2}}$$

$$\rho_{UV} = \frac{(ac + bd)\,\sigma^2}{\sigma^2\sqrt{(a^2 + b^2)(c^2 + d^2)}}$$

$$\rho_{UV} = \frac{ac + bd}{\sqrt{(a^2 + b^2)(c^2 + d^2)}}$$

2) X e Y são variáveis aleatórias independentes, para as quais existem os momentos até a ordem 2; determinar uma condição necessária e suficiente para $U = X + Y$ e $V = X - Y$ serem não correlacionadas.

Solução:

Se existem os momentos até de ordem 2, consequentemente existem:

$$E(X), E(Y), V(X)\ e\ V(Y)$$

$$E(U) = E(X) + E(Y)$$

$$E(V) = E(X) - E(Y)$$

$$V(U) = V(X) + V(Y)$$

$$V(V) = V(X) + V(Y)$$

Sabe-se que:

$$\rho_{UV} = \frac{\mathrm{COV(U, V)}}{\sqrt{V(U)V(V)}}$$

$$COV(U,V) = E(UV) - E(U)E(V)$$

Para U e V serem não correlacionados devemos ter $\rho_{UV} = 0$ na expressão $COV(U,V) = E(UV) - E(U)\,E(V)$ $COV(U,V) = 0$ e $V(X) > 0, V(Y) > 0$.

Logo, para a expressão $COV(U,V) = E(UV) - E(U)E(V)$ temos:

$$0 = E(UV) - E(U)E(V) \therefore E(UV) = E(U)E(V)$$

$$E(UV) = E[(X + Y)(X - Y)] = E[X^2 - Y^2) = E(X^2) - E(Y^2)$$

$$E(U)E(V) = [E(X) + E(Y)][E(X) - E(Y)] = [E(X)]^2 - [E(Y)]^2$$

Voltando a expressão $E(UV) = E(U)E(V)$ temos:

$$E(X^2) - E(Y^2) = [E(X)]^2 - [E(Y)]^2$$

$$E(X^2) - [E(X)]^2 = E(Y^2) - [E(Y)]^2$$

$V(X) = V(Y)$ que é uma condição necessária e suficiente para U e V serem não correlacionados.

Exemplo 13. Determine a função de distribuição ou de repartição de uma variável aleatória discreta. Considere X: soma dos pontos no lançamento de dois dados.

i) A distribuição

ii)

TABELA 28. FUNÇÃO DE DISTRIBUIÇÃO OU DE REPARTIÇÃO DE UMA VARIÁVEL ALEATÓRIA DISCRETA X: SOMA DOS PONTOS NO LANÇAMENTO DE DOIS DADOS HONESTOS.

X_i	2	3	4	5	6	7	8	9	10	11	12
$F(X_i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

iii) O gráfico de $F(x)$

iv)

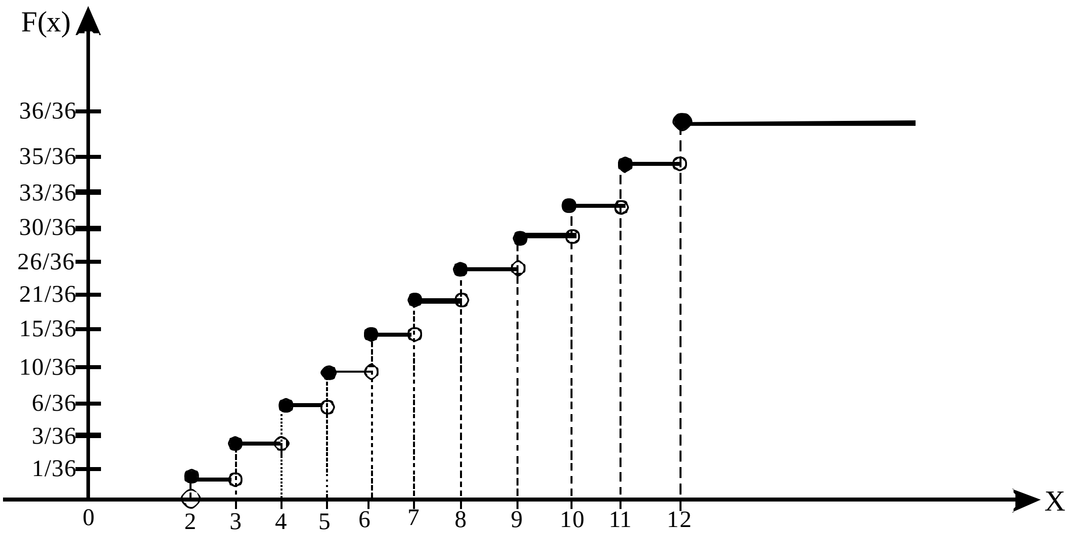


Figura 22. Gráfico em escada da variável aleatória discreta X: Soma dos pontos obtidos no lançamento de dois dados honestos.

v) A função $F(x)$

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \leq x < 3 \\ 3/36 & 3 \leq x < 4 \\ 6/36 & 4 \leq x < 5 \\ 10/36 & 5 \leq x < 6 \\ 15/36 & 6 \leq x < 7 \\ 21/36 & 7 \leq x < 8 \\ 26/36 & 8 \leq x < 9 \\ 30/36 & 9 \leq x < 10 \\ 33/36 & 10 \leq x < 11 \\ 35/36 & 11 \leq x < 12 \\ 1 & 12 \leq x < \infty \end{cases}$$

Exemplo 14. Exercícios de aplicação sobre função densidade 1. Seja X uma variável aleatória contínua com função densidade:

$$f(x) = \begin{cases} \frac{1}{8}, & \text{se } 0 \leq x \leq 8 \\ 0, & \text{em contrário} \end{cases}$$

i) Encontre $P(2 \leq x \leq 5)$

ii) $P(3 \leq x \leq 7)$

iii) $P(x \geq 6)$

iv) Determine e trace o gráfico da função de distribuição $F(x)$.

1. Seja X uma variável aleatória contínua com função densidade:

$$f(x) = \begin{cases} kx, & \text{se } 0 \leq x \leq 5 \\ 0, & \text{em caso contrário} \end{cases}$$

i) Calcule k;

ii) Encontre $P(1 \leq x \leq 3)$

iii) $P(2 \leq x \leq 4)$

iv) $P(x \geq 3)$.

2. Seja X uma variável aleatória com função densidade:

$$f(x) = \begin{cases} \frac{1}{2}x, & \text{se } 0 \leq x \leq 2 \\ 0, & \text{em contrário} \end{cases}$$

Encontre:

i) a função de distribuição acumulada

ii) $F(\mathrm{x})$ e o seu gráfico.

3. Considere a função dada por:

$$f(x) = \begin{cases} 0, & \text{se } x < 1{,}5 \\ x, & \text{se } 1{,}5 \leq x \leq 2 \\ 0{,}25 & \text{se } 2 \leq x \leq 2{,}5 \\ 0, & \text{se } x > 2{,}5 \end{cases}$$

i) Mostre que $\mathrm{f(x)}$ é uma função densidade de probabilidade;

ii) Escolhido um valor ao acaso para x, qual é a probabilidade de x pertencer ao intervalo $[1{,}5;\ 2{,}0]$?

4. Considere a função dada por:

$$f(x) = \begin{cases} 0, & \text{se } x < 0 \\ 0{,}2x & \text{se } 0 < x < 1 \\ 0{,}2, & \text{se } 1 \leq x \leq 5 \\ 1{,}2 - 0{,}2x & \text{se } 5 < x < 6 \\ 0, & \text{se } x > 6 \end{cases}$$

i) Mostre que $\mathrm{f(x)}$ é uma função densidade de probabilidade;

ii) Represente graficamente a função $\mathrm{f(x)}$;

iii) Calcule $P(X > 1)$

iv) Calcule $P(1 \leq \mathrm{x} \leq 5)$;

v) Calcule $P(0{,}5 \leq \mathrm{x} \leq 5{,}2)$.

5. Dada a função de densidade de probabilidade (fdp) abaixo:

$$f(x) = \begin{cases} cx, & \text{se } 0 \leq x \leq 2 \\ 0, & \text{em caso contrário} \end{cases}$$

i) Encontre o valor da constante c;

ii) $E(X)$;

iii) $Var(X)$.

Respostas:

1- i) $\frac{3}{8}; \frac{1}{2}; \frac{1}{4}$

ii) $F(\mathrm{x}) = \begin{cases} 0, & \text{se x} < 0 \\ \frac{\mathrm{x}}{8}, & \text{se } 0 \le \mathrm{x} \le 8 \\ 1 & \text{se x} > 8 \end{cases}$

2- i) $\mathrm{k} = \frac{2}{25};$

ii) $\frac{8}{5}; \frac{12}{25}; \frac{9}{25}$

3- $\mathrm{F(x)} = \begin{cases} 0, & \text{se x} < 0 \\ \frac{1}{4}\mathrm{x}^2, & \text{se } 0 \le \mathrm{x} \le 2 \\ 1, & \text{se x} > 2 \end{cases}$

4- i) $\mathrm{f(x)}$ é uma legítima fdp;

ii) 0,875

5- i) $\mathrm{f(x)}$ é uma legítima fdp;

ii) 0,9; d) 0,8;

iii) 0,911

6-

i) $\mathrm{C} = \frac{1}{2};$

ii) $\mathrm{E}(X) = \frac{4}{3};$

iii) $\mathrm{Var}(X) = \frac{2}{9}$

Exemplos 15.

Exercício 1. Para cada 5 m de perfuração, na construção das fundações de um prédio, observa-se a ocorrência ou não de zonas de fratura. Admita que a cada 5 m a ocorrência de fratura tem probabilidade igual a 0.1. O custo de perfuração a cada 5m é de 100 (unidades monetárias convenientes) se não houver fratura. Será perfurado 15m. Há um acréscimo total de 50K (unidades monetárias convenientes), onde K é o número de zonas de fratura encontradas nos 15 m, K = 0,1, 2 ou 3. Admita que só pode ocorrer uma zona de fratura a cada 5m. Estude a variável custo das perfurações.

i) Faça um diagrama de árvore que indique os diversos valores possíveis do custo.

ii) Considere X, custo, como uma variável aleatória. Construa o espaço amostral e a distribuição de probabilidade de X.

iii) Calcule o custo médio.

iv) Calcule o desvio padrão do custo.

Exercício 2. Devido a introdução de um novo imposto o custo do exercício anterior passa a ser $Y = 1{,}02X + 0{,}3$.

i) Calcule a nova média.

ii) E o novo desvio padrão.

Exercício 3. Dez poços artesianos serão furados até o limite de 40m. A probabilidade de se encontrar água naquela região é de 0,3. Admita as hipóteses necessárias e calcule:

i) Qual a probabilidade de se encontrar água em pelo menos 2 poços.

ii) Qual o número médio de poços em que se encontrará água.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMOWITZ, M.; I. A. STEGUN, **Pocketbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables**, Verlag Harri, Frankfurt, 468 p., Natl. Inst. of Stand. and Technol., Gaithersburg, Md., 1984.

ASSIS, F.N. DE ARRUDA, H. V., PEREIRA, A. R. **Aplicações de estatística a climatologia: Teoria e prática**, Pelotas: Ed. Universitária/ UFPel, 1996. 161 p.

BATSCHELET, E. **Introdução à matemática para biocientistas.** São Paulo: EDUSP, 1978. 596 p.

BOTELHO, V.A.; MORAIS, A.R. de. Estimativas dos parâmetros da distribuição gama aos dados pluviométricos do município de Lavras, Estado de Minas Gerais. **Ciência e Agrotecnologia**, Lavras, v. 23, n. 3, p. 697-706 julho/setembro de 1999.

BRESSAN, G. Modelagem e simulação de sistemas computacionais. São Paulo; Larc–PCS/EPUSP, IME/USP, 2002. 12p.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. 5. ed. São Paulo: Saraiva, 2002.

CASELLA, G.; BERGER, R. L. Statistical Inference (2nd edition) Wadsworth Pub Co.2001. 660 P.

CATALUNHA, M.J. et al. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v. 10, n.1, p.153-162; Santa Maria, 2002.

CHOW, V. T. **Handbook of Applied Hydrology**. New York, McGraw-Hill, 1964. Não paginado.

CHOW, V. T. Maidment, D. R.; Mays, L. W. **Appied hydrology**. New York:Mcgraw-Hill, 1988.572 p.

CLARK, A. B.; DISNEY, R. L. **Probabilidade e processos estocásticos.** Livros Técnicos e Científicos Editora S. A. Rio de janeiro. 1979. 352 p.

COUTO, H.T.Z. **Distribuições de diâmetro em plantações de *Pinus caribaea morelet***, Piracicaba: Escola Superior de Agricultura Luiz de Queiroz/USP, 79 P. Tese (Livre Docência), ESALQ, 1980.

CRAMER, H. **Mathematical methods of statistics**, Princeton University Press, New Jersey 1966.

DAGNELLI, P. **Estatística: Teoria e Métodos.** Publicações Europa-America, Portugal. Vol. 1. 1975.440 p.

DAGNELLI, P. **Estatística: Teoria e Métodos.** Publicações Europa-America, Portugal. Vol. 2. 1975. 534 p.

DAGNELLI, P. **Estatística: teoria e métodos.** Publicações Europa-América, Portugal. Vol. 1 e 2. 1985. 536 p.

DANTAS, CARLOS A. B. **Probabilidade- um curso introdutório.** Coleção: Acadêmica Editora: EDUSP. 2ª Edição - 2004 – 256 p.

DEVORE, J. L. Probability and Statistics for Engineering and the Sciences With Infotrac. Hardcover: Thomson learning.2003, 764 p.

DHARMADHIKARI, S. Bounds on quantiles: A comment on O`Cinneide. **The American Statistician**, Alexandria, v. 45, n. 3, p. 257-258, Aug. 1991.

FELLER, W. **An Introduction to probability theory and its application** 3 ed. New york, John Wiley & Sons, 1967, vol. 1.236 p.

FERNANDEZ, P., **Introdução à Teoria das Probabilidades**. Rio de Janeiro, LTC, Coleção Elementos da Matemática, 1975.

FERREIRA, D. F;**Estatística Básica,** Lavras: Editora UFLA, 1ª Edição, 2005.664 p.

FISZ,M., **Probability Theory and Mathematical Statistics,** Jonh Wiley & Sons, Inc., New York, 1963.

FONSECA, J., **Introdução à Estatística Matemática – Aplicações**, Edição SPB, 1994.

GAMA, S.M. A.; PEDROSA, A. C.. **Introdução computacional a probabilidade e estatística.** Coleção: Universitária: Porto Editora. Portugal, 1ª edição. 2004.590 p.

GRAÇA, M. E., **Introdução às Probabilidades e Estatística**, DEIO, FCUL, Sociedade Portuguesa de Estatística,1998.

GUIMARÃES, R.C.; SARSFIELD C. J.A., **Estatística**, McGraw-Hill, 1997.

HANN, C. T. **Statistical methods in hydrology**, Ames: Iowa State University Press, 1977. 378 p.

HASTINGS, N.A.J.; PEACOCK, J.B. **Statistical distributions**: A handbook for students and practitioners, Longon Butterworths, England, 1975. 129p.

HAYTER, A. J. Probability and Statistics for Engineers and Scientists (2nd edition.) Duxbury P.2002 916 P.

HAZAN, S. **Fundamentos da Matemática Elementar: combinatória e probabilidade**. 6. ed. São Paulo: Atual Editora, 1993.

HOEL P. G. **Introduction to mathematical statisitcs.** Wiley, New York, 1965, 427 p.

HOEL, P. G. ; PORT, S. C. ; STONE, C. J. **Introdução á teoria da probabilidade**. Rio de janeiro, Interciência, 1978.269 p.

HOEL, P.G. **Estatística matemática,** 4 ed. Rio de Janeiro:Editora Guanabara, 1980.373 p.

JAMES, B. R., **Probabilidade: Um Curso em Nível Intermediário**. Rio de Janeiro, IMPA, Projeto Euclides, 1981.

JOHNSON, N. L., KOTZ, S. **Continous univariate distributions**. Boston: Houghton Millin, 1970, v.2 p.112.

JOHNSON, N. L., KOTZ, S. **Distribution in statistics, continue univariate distribution**, New York: Hougton Mifflin, v. 2, 1970. 328 p.

KVANLI, A. H. **Statistics**: A Computer Integrated Approach Publisher: West Publishing Company. 1988.935 P.

LANNA, A. E. **Elementos de estatística e probabilidade**. In: TUCCI, C. E. M. Hidrologia. Porto Alegre: EDUSO; ABRH, 1993. P. 79-164.

LIPSCHUTZ, S. **Probabilidade**. São Paulo: Mc Graw-Hill, 1974. 228p.

LOÉVE, M;. **Probability Theory**. 3 ed.. New York, D. Van Nostrand, 1963.

MAGALHÃES, M. N.; LIMA, ª C. P. **Noções de Probabilidade e Estatística.** São Paulo, EDUSP, 2002, 398 p.

MARTINS, M. J. **Simulação** (1ª Parte). Lisboa: Instituto Superior de Agronomia, 2003. 60 p. (notas de Aula).

MELLO, F., G. **Probabilidades e Estatística,** Volumes 1 e 2, Escolar Editora, 1993.

MELLO, F. , **Introdução aos Métodos Estatísticos**, vol. I e II, Cadernos do Instituto de Orientação profissional, Lisboa, 1973.

MEYER, P. L. **Probabilidade**: **Aplicações à Estatística**. 1ª ed. Rio de Janeiro; Livro Técnico, 1969, 391 p.

MEYER, P. L. **Probabilidade**: **Aplicações à Estatística**. Livros Técnicos Científicos Editora S.A. 1ª ed. Rio de Janeiro; Livro Técnico, 1978, 405 p.

MONTGOMERY, D. C. e RUNGER, G. C.. **Applied Statistics and Probability for Engineers**, 3ª Edição.John Wiley & Sons, New York. 2003.

MONTGOMERY, D. C. **Estatística aplicada e probabilidade para engenheiros** Editora**:** LTC. 2ª Edição – 2003. 478p

MOOD, A. M.; GRAYBILL, F. A. **Introdution to the theory of statistics.** Mc Graw Hill, New York, 1963, 443 p.

MOOD, A. M., GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. McGraw-Hill.1974.

MORETTIN, P. A; BUSSAB, W. O. **Estatística básica**. 5ª edição., São Paulo, Saraiva, 2003, 526 p.

MORGADO, A. C. O.; CARVALHO, J. B. P.; CARVALHO, P. C. P.; FERNANDEZ, P. **Análise Combinatória e Probabilidade**. 6. edição. Rio de Janeiro: SBM, 2004.

MURRAY R. SPIEGEL; JOHN SCHILLER; E. R. ALU SRUNIVASAN. **Probabilidade e Estatística**. Porto Alegre: 2. edição. Bookman, 2004. 398p.

MURTEIRA, B. J., RIBEIRO, C. S., ANDRADE E SILVA, J. e PIMENTA, C.. **Introdução à Estatística**.. McGraw Hill, Lisboa. 2002.

MURTEIRA, B., **Probabilidades e Estatística**, Volumes 1 e 2, McGraw-Hill, 1997.

NAGHETTINI, M.; PINTO, E. J. A. **Hidrologia estatística.** Serviço geológico do Brasil – CPRM. Belo Horizonte, MG. 2007. 552 p.

O`CINNEIDE, C. A. The mean is within one standard deviation of any median. **The American statistician**, Alexandria. V. 44, n. 4 p. 292-293, nov.. 1990.

OLIVEIRA, P. L.; COSTA NETO; CYMBALISTA, M. **Probablidades: resumos teóricaos, exercícios resolvidos; exercícios propostos**. Editora Edgard Blucher ltda. São Paulo. 1988.145 p.

PACITTI, T. **Fortran**. Rio de janeiro: Livro Técnico e Científico, 1974. 377 p.

PAPOULIS, A.Probability Random Variables, and Stochastic Processes. McGraw-Hill College.2001.576 P.

PARZEN, E. **Modern probality Theory**, Van Nostrand, New York, 1955.

PAULINO E BRANCO. **Exercícios de Probabilidade e Estatística**. Escolar Editora.2005.

PAULINO, C. D.; Branco, J. **Exercícios de Probabilidade e Estatística**.Escolar Editora, Lisboa..2004.

PEDROSA, A. C. ; GAMA, S. M. A. **Introdução Computacional à Probabilidade e Estatística**.Porto Editora.Porto-Portugal.2004, 607 p.

PIMENTEL-GOMES, F. **Iniciação à Estatística**. 5ª ed. São Paulo: Livraria Nobel S.A., 1976. 236p.

ROBALO, A., **Estatística- Exercícios,** Volumes 1 e 2, Edições Sílabo, 1995.

ROHATGI **An Introduction to Probability Theory and Mathematical Statistics**. John Wiley.1976.

ROSS, S. M. **Introduction to Probability and Statistics for Engineers and Scientists**, 3ª Edição. Elsevier/Academic Press, Burlington, MA.2004.

ROSS, S. M. **A First course in probability**. 4 ed. London/ New York, mac-millan, 1994.

ROSS, S. M. First Course in Probability. Prentice Hall. 2005.576 p.

SÀNCHEZ, J. A. P. Um dia inesquecível na vida de Gauss. **Revista do professor de matemática.** São Paulo: SBM, Nº 37: 11 – 13, 1998.

SCOLFORO, J. R. S. **Modelos para expressar o crescimento e a produção florestal: Parte 1. Lavras**, ESAL/FAEPE. 188p. 1994.

SOONG, T. T. **Modelos Probabilísticos em Engenharia e Ciências**; Tradução Alfredo Alves de Farias. Rio de Janeiro: LTC-Livros Técnicos e Científicos Editora S. A.., 1986

SPIEGEL, M., **Probabilidade e Estatística**, Colecção Schaum, McGraw-Hill, 1978.

THOM, H. C. S. **A note on the gamma distribution. Monthly Weather Review**, Washington, v. 86, p. 117-122. 1958.

THOM, H. C. S. **Some methods of climatological analysis**. Roma, FAO, 1966. 50p. (FAO, Technical Notes. 81).

TIAGO DE OLIVEIRA. **Probabilidades e Conceitos, Métodos e Aplicações estatísticas.**vol. I, II. McGraw-Hill. 1990.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL: Departamento de Matemática e Estatística:em http://athena.mat.ufrgs.br/~portosil/histo2.html.Acesso em 25 de março de 2007.

WILKS S. S. **Mathematical statistics.** Wiley, New York, 1962, 644 p.

APÊNDICE 1

ALFABETO GREGO

NOME DA LETRA	SÍMBOLOS	
	MAIÚSCULAS	MINÚSCULAS
Alfa	Α	α
Beta	Β	β
Gama	Γ	γ
Delta	Δ	δ
Épsilon	Ε	ε
Zeta	Ζ	ζ
Eta	Η	η
Téta	Θ	θ
Iota	Ι	ι
Capa	Κ	κ
Lambda	Λ	λ
Um(mi)	Μ	μ
Nu(ni)	Ν	ν
Csi	Ξ	ξ
Omicron	Ο	ο
Pi	Π	π
Ró	Ρ	ρ
Sigma	Σ	σ
Tau	Τ	τ
Úpsilon(ipsilon)	Υ	υ
Fi	Φ	φ
Chi(qui)	Χ	χ
Psi	Ψ	ψ
Omega	Ω	ω

APÊNDICE 2

VALORES $e^{-\lambda}$

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
0,00	1,0000000	2,55	0,0780818	5,10	0,0060968	7,65	0,0004760	10,20	0,0000372
0,05	0,9512295	2,60	0,0742737	5,15	0,0057994	7,70	0,0004528	10,25	0,0000354
0,10	0,9048375	2,65	0,0706513	5,20	0,0055166	7,75	0,0004307	10,30	0,0000336
0,15	0,8607081	2,70	0,0672056	5,25	0,0052475	7,80	0,0004097	10,35	0,0000320
0,20	0,8187309	2,75	0,0639280	5,30	0,0049916	7,85	0,0003898	10,40	0,0000304
0,25	0,7788009	2,80	0,0608102	5,35	0,0047482	7,90	0,0003707	10,45	0,0000289
0,30	0,7408184	2,85	0,0578444	5,40	0,0045166	7,95	0,0003527	10,50	0,0000275
0,35	0,7046883	2,90	0,0550233	5,45	0,0042963	8,00	0,0003355	10,55	0,0000262
0,40	0,6703202	2,95	0,0523398	5,50	0,0040868	8,05	0,0003191	10,60	0,0000249
0,45	0,6376283	3,00	0,0497872	5,55	0,0038875	8,10	0,0003035	10,65	0,0000237
0,50	0,6065309	3,05	0,0473590	5,60	0,0036979	8,15	0,0002887	10,70	0,0000225
0,55	0,5769500	3,10	0,0450493	5,65	0,0035175	8,20	0,0002747	10,75	0,0000214
0,60	0,5488119	3,15	0,0428522	5,70	0,0033460	8,25	0,0002613	10,80	0,0000204
0,65	0,5220460	3,20	0,0407623	5,75	0,0031828	8,30	0,0002485	10,85	0,0000194
0,70	0,4965855	3,25	0,0387743	5,80	0,0030276	8,35	0,0002364	10,90	0,0000185
0,75	0,4723668	3,30	0,0368832	5,85	0,0028799	8,40	0,0002249	10,95	0,0000176
0,80	0,4493292	3,35	0,0350844	5,90	0,0027395	8,45	0,0002139	11,00	0,0000167
0,85	0,4274152	3,40	0,0333733	5,95	0,0026059	8,50	0,0002035	11,05	0,0000159
0,90	0,4065699	3,45	0,0317457	6,00	0,0024788	8,55	0,0001935	12,00	0,0000061
0,95	0,3867413	3,50	0,0301975	6,05	0,0023579	8,60	0,0001841	12,05	0,0000058
1,00	0,3678797	3,55	0,0287247	6,10	0,0022429	8,65	0,0001751	13,00	0,0000023
1,05	0,3499380	3,60	0,0273238	6,15	0,0021335	8,70	0,0001666	13,05	0,0000022
1,10	0,3328713	3,65	0,0259912	6,20	0,0020294	8,75	0,0001585	14,00	0,0000008
1,15	0,3166370	3,70	0,0247236	6,25	0,0019305	8,80	0,0001507	14,05	0,0000008
1,20	0,3011945	3,75	0,0235178	6,30	0,0018363	8,85	0,0001434	15,00	0,0000003
1,25	0,2865050	3,80	0,0223708	6,35	0,0017468	8,90	0,0001364	15,05	0,0000003
1,30	0,2725320	3,85	0,0212798	6,40	0,0016616	8,95	0,0001297	16,00	0,0000001

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
1,35	0,2592405	3,90	0,0202420	6,45	0,0015805	9,00	0,0001234	16,05	0,0000001
1,40	0,2465972	3,95	0,0192548	6,50	0,0015034	9,05	0,0001174	17,00	0,0000000
1,45	0,2345705	4,00	0,0183157	6,55	0,0014301	9,10	0,0001117	17,05	0,0000000
1,50	0,2231304	4,05	0,0174224	6,60	0,0013604	9,15	0,0001062	18,00	0,0000000
1,55	0,2122482	4,10	0,0165727	6,65	0,0012940	9,20	0,0001010	18,05	0,0000000
1,60	0,2018967	4,15	0,0157645	6,70	0,0012309	9,25	0,0000961	19,00	0,0000000
1,65	0,1920501	4,20	0,0149956	6,75	0,0011709	9,30	0,0000914	19,05	0,0000000
1,70	0,1826837	4,25	0,0142643	6,80	0,0011138	9,35	0,0000870	20,00	0,0000000
1,75	0,1737741	4,30	0,0135686	6,85	0,0010595	9,40	0,0000827	20,05	0,0000000
1,80	0,1652991	4,35	0,0129069	6,90	0,0010078	9,45	0,0000787	21,05	0,0000000
1,85	0,1572374	4,40	0,0122774	6,95	0,0009586	9,50	0,0000749	22,00	0,0000000
1,90	0,1495688	4,45	0,0116786	7,00	0,0009119	9,55	0,0000712	22,05	0,0000000
1,95	0,1422743	4,50	0,0111090	7,05	0,0008674	9,60	0,0000677	23,00	0,0000000
2,00	0,1353355	4,55	0,0105672	7,10	0,0008251	9,65	0,0000644	23,05	0,0000000
2,05	0,1287351	4,60	0,0100519	7,15	0,0007849	9,70	0,0000613	24,00	0,0000000
2,10	0,1224566	4,65	0,0095616	7,20	0,0007466	9,75	0,0000583	24,05	0,0000000
2,15	0,1164843	4,70	0,0090953	7,25	0,0007102	9,80	0,0000555	25,00	0,0000000
2,20	0,1108033	4,75	0,0086517	7,30	0,0006755	9,85	0,0000527	25,05	0,0000000
2,25	0,1053994	4,80	0,0082298	7,35	0,0006426	9,90	0,0000502	26,00	0,0000000
2,30	0,1002590	4,85	0,0078284	7,40	0,0006113	9,95	0,0000477	26,05	0,0000000
2,35	0,0953693	4,90	0,0074466	7,45	0,0005814	10,00	0,0000454	27,00	0,0000000
2,40	0,0907181	4,95	0,0070834	7,50	0,0005531	10,05	0,0000432	27,05	0,0000000
2,45	0,0862937	5,00	0,0067380	7,55	0,0005261	10,10	0,0000411	28,00	0,0000000
2,50	0,0820851	5,05	0,0064094	7,60	0,0005005	10,15	0,0000391	28,05	0,0000000

ÍNDICE REMISSIVO

SOBRE OS AUTORES

Janilson Pinheiro de Assis, Natural de João Câmara, RN

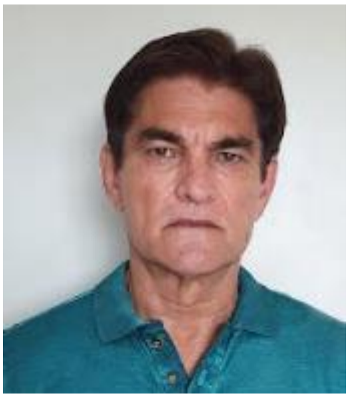
Engenheiro Agrônomo graduado em Engenharia Agronômica (1987) na ESCOLA SUPERIOR DE AGRICULTURA DE MOSSORÓ - ESAM. Concluiu o Curso de Pós-graduação-Mestrado (1990) em Engenharia Agronômica (Fitotecnia-Estatística experimental) na Universidade Federal do Ceará (UFC). Realizou o curso de Pós-graduação Doutorado (2014) em Produção Vegetal - Fitotecnia na Universidade de São Paulo (USP), com pesquisa direcionada para a área de estatística e modelagem aplicada à agricultura. Orientou e orienta estudantes bolsistas de monitoria, de residência acadêmica e iniciação científica. Atualmente, é Professor Titular da Universidade Federal Rural do Semi-Árido (UFERSA), Foi Professor da Escola Superior de Agricultura de Mossoró (ESAM) onde leciona a disciplina de Estatística há 33 anos, possui seis livros publicados, 11 apostilas e mais de 25 artigos completos publicados/aceitos em revistas nacionais e internacionais além de 20 resumos simples/expandido. É revisor de dez revistas nacionais e internacionais. Contato: (85) 99826636.

Isaac Reinaldo Pinheiro de Lima

Natural de Natal, RN, É estudante de graduação do curso de Bacharelado em Tecnologia da Informação da Universidade Federal do Rio Grande do Norte-UFRN (2019- 2023). Ingressou através do Processo de Seleção ENEM-Exame Nacional de Ensino Médio e do programa de seleção Unificado-SISU do Ministério da Educação. Já foi selecionado para ingresso (onde cursou parcialmente) curso de Graduação de Sistemas de Informação na Universidade de São Paulo-USP, EACH. Atualmente, é Bolsista do programa PSH – Painel de Segurança Hídrica na Universidade Federal do Rio Grande do Norte, Além de Estagiário de tecnologia na Empresa VTEX, possui um livro publicado. Já Concluiu cursos de Robótica e programação Computacional, possui um artigo científico publicado em revista internacional com Qualis da Capes B1. Contato: (84) 98810 9278.

Joelma de Assis França, Natural de João Câmara, RN

Professora do ensino fundamental do estado do Rio Grande do Norte, e Da Prefeitura da Cidade de João Câmara , RN, Com atuação durante mais de vinte anos, É Graduada em Licenciatura em Ciências Biológicas Pela Universidade federal do Rio Grande do Norte-UFRN, EM Natal, RN, Atualmente é Mestranda do Programa de Pós Graduação CIÊNCIA DA EDUCAÇÃO na Universidade Federal do Rio Grande do Norte (UFRN) em Natal , RN, com Longa experiência na área de Docência, E Administrativa como Diretora de Escola, Possui um Livro Publicado Durante o Ano de 2023 Sobre Transformação de Dados Estatísticos. Contato: (84) 988539347.



Roberto Pequeno de Sousa

Engenheiro Agrícola, graduado em Engenharia Agrícola (1981) na Universidade Federal da Paraíba (UFPB). Mestre (1985) em Engenharia Civil (Recursos Hídricos - Irrigação) na Universidade Federal da Paraíba (UFPB). Doutor (2013) em Agronomia - Fitotecnia na Universidade Federal Rural do Semi-Árido (UFERSA). Atualmente, é Professor Titular da Universidade Federal Rural do Semi-Árido (UFERSA), leciona a disciplina de Estatística Experimental, possui sete livros publicados, 62 artigos completos publicados/aceitos em revistas nacionais e internacionais, 42 resumos simples/expandido. É revisor de cinco revistas nacionais e internacionais. Contato: (84)99411-5032.

Robson Pequeno de Sousa

Doutor em Engenharia Elétrica pela Universidade Federal da Paraíba (2000), mestrado em Engenharia Elétrica pela universidade Federal de Pernambuco (1991) e graduação em Bacharelado em Matemática pela Universidade Federal da Paraíba (1985). Professor Doutor Associado D do Departamento de Computação da Universidade Estadual da Paraíba e membro efetivo do Programa de Pós-Graduação em Ciência e Tecnologia em Saúde – PPGCTS-UEPB. Coordena o Laboratório de Análises de Imagens e Sinais –LAIS do Núcleo de Tecnologias Estratégicas em Saúde – NUTES. Contatos: Fone: 083993129256.

Telde Natel Custódio

Natural de Lavras/MG. Possui graduação em Engenharia Agrícola pela Escola Superior de Agricultura de Lavras, atual Universidade Federal de Lavras (1988), mestrado em Agronomia com área de concentração em Estatística e Experimentação Agropecuária pela Universidade Federal de Lavras (1999), doutorado em Agronomia com área de concentração em Estatística e Experimentação Agronômica pela Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo (2004). Atualmente é professor na área de Estatística pela Universidade Federal de São João Del-Rei (UFSJ). Tem experiência docente na área de Probabilidade e Estatística. Atua principalmente nos seguintes temas: Estatística Experimental Superfície de resposta, Análise de regressão, Modelos lineares generalizados, Modelos lineares generalizados mistos, Meta-análise. Contato: natel@ufsj.edu.br. (35)99979-0218.

Walter Martins Rodrigues

Possui graduação em Licenciatura em Matemática pela USP (1987), onde também cursou o mestrado concluído em 2020 e doutorado em Matemática finalizado em 2005, na área de Representação de álgebras. Atualmente é professor Titular da Universidade Federal Rural do Semi-Árido (UFERSA). Exerceu a atividade de coordenador Pedagógico e pró-reitor adjunto de Graduação no período de agosto de 2013 a fevereiro de 2015. Coordenou o mestrado Profissional em Matemática da UFERSA de 2020 até março de 2023. Trabalha com pesquisa voltada para modelagem Matemática aplicada, Álgebra e Modelagem Estatística. Contato: Fone: 084988933633.

Joaquim Odilon Pereira

Possui Graduação em Engenharia Agrícola pela Universidade Federal da Paraíba – Campina Grande (1980), Mestrado em Engenharia Agrícola pela Universidade Federal de Viçosa (1986), Doutorado em Agronomia (Energia na Agricultura) pela Universidade Estadual Paulista Júlio de Mesquita Filho (1996). É Pós-doutorado em Engenharia Agrícola com domínio em Dinâmica do solo na interação Solo-máquina pelo Institut National de la Recherche Agronomique – França (2002). Foi Professor Associado da Universidade Estadual do Oeste do Paraná do Curso de Graduação e Pós-graduação em Engenharia Agrícola no período de 1988 a 2006. Atualmente é Professor Titular da Universidade Federal Rural do Semi Árido – Mossoró – RN no curso de Engenharia Agrícola e Ambiental. Avaliador Institucional do INEP/MEC. Tem experiência na área de Engenharia Agrícola, com ênfase em Mecanização e Máquinas Agrícolas, atuando, principalmente, nos seguintes temas: Ergonomia, Dinâmica do solo no sistema de interação solo-máquina, Compactação do solo e Sistema de manejo do solo. Participa como colaborador de trabalhos científicos em coautoria com pesquisadores de instituições de ensino e pesquisa internacional e nacional. É revisor para periódicos nacionais e internacionais, tais como Engenharia Ambiental, Soil Use and Management, Journal of Agricultural Science, Journal of Terramechanics, Engenharia Agrícola e Agricultural Engineering International: CIGR Journal. Contato: 084999310198.

Pantanal Editora
Rua Abaete, 83, Sala B, Centro. CEP: 78690-000
Nova Xavantina – Mato Grosso – Brasil
Telefone (66) 99682-4165 (Whatsapp)
https://www.editorapantanal.com.br
contato@editorapantanal.com.br