Capítulo 1

Probabilidade

Objetivo: Definir um modelo estatístico que seja adequado à descrição e interpretação de fenômenos aleatórios.

1.0.1 Conceitos básicos

Experimentos ou fenômenos aleatórios (ε) : são os acontecimentos cujos resultados não podem ser previstos com certeza, sob condições idênticas.

Exemplos :

- Lançamento de um dado.
- Lançamento de uma moeda.
- Tempo de vida útil de um componente eletrônico.

Espaço Amostral (Ω): refere-se ao conjunto de todos os possíveis resultados de um experimento ou fenômeno aleatório.

Exemplos :

$\Omega_1 = \{1, 2, 3, 5, 6\}$

$\Omega_2 = \{c, k\}$ Sendo k cara, c coroa.

$\Omega_3 = [0, \infty\}$

Evento: qualquer subconjunto do espaço amostral Ω do experimento aleatório ε.

Notação: $A, B, C, D, \ldots, A_1, A_2, A_3, \ldots, B_1, B_2, \ldots$

Notamos que, como A é um evento então $A \subset \Omega$.

1.0.2 Tipos de Eventos

Evento Simples ou Elementar: é o evento formado por um único ponto do espaço amostral.

Exemplo: $A = \{w\}$.

Evento Composto: é o evento formado por dois ou mais pontos do espaço amostral.

Exemplo: $A = \{w_1, w_2, w_3\}$

Evento Certo: é o evento formado por todos os pontos amostrais.

Exemplo : $A = \Omega$

Evento Impossível : É o evento que não possuí elementos de Ω, isto é, evento vazio.

Notação : $A = \{\}$ ou $A = \emptyset$

Alguns Exemplos :

$A =$ Face do dado maior que 5.
$A = \{6\}$

$B =$ Face do dado ser par.
$B = \{2, 4, 6\}$

$C =$ Face do dado maior ou igual a 1.
$C = \{1, 2, 3, 4, 5, 6\} = \Omega$

$D =$ Face do dado maior que 6.
$D = \{\}$ ou $D = \emptyset$

1.0.3 Operação com Eventos

Para ilustrar graficamente eventos é costume utilizar-se dos mesmos diagramas de Venn utilizados na teoria de conjuntos.

Considere eventos definidos em um espaço amostral Ω de um experimento aleatório ε.

União de eventos $(A \cup B)$: é o evento formado por todos os elementos que pertencem a A, ou B, ou ambos.

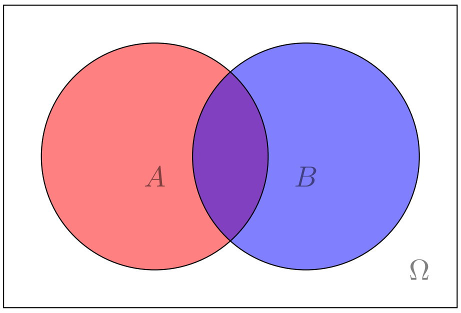


Figura 1.1:

Intersecção de eventos ($A \cap B$) : é o evento formado pelos elementos que pertencem a A e a B.

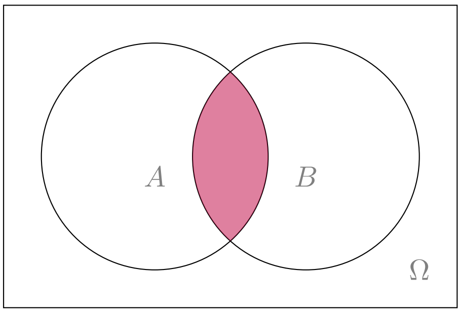


Figura 1.2:

Casos Particulares :

1. Se $B \subset A$, então $A \cap B = B$

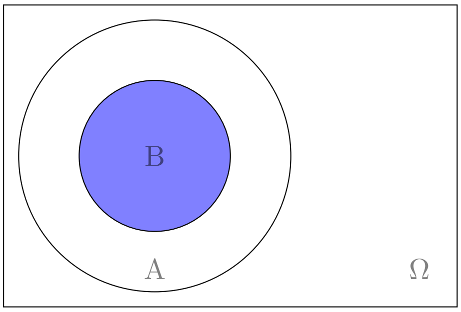


Figura 1.3:

2. Se A e B são eventos disjuntos ou mutuamente exclusivos (não possui elementos comuns), então $A \cap B = \emptyset$.

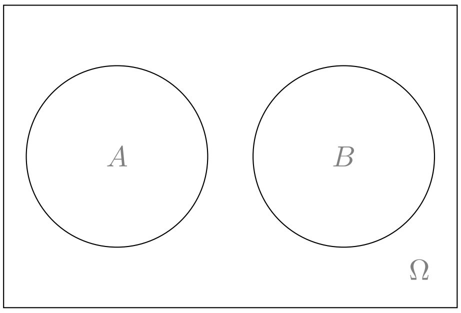


Figura 1.4:

Diferença de eventos ($A - B$) : é o evento formado pelos elementos que pertencem a A mas não pertencem a B.

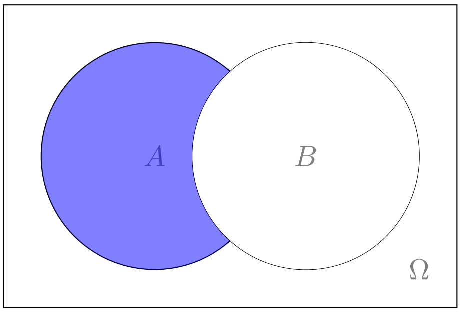


Figura 1.5:

Evento Complementar ($\bar{A}$ ou A^c) : é o evento formado por todos os elementos de Ω que não pertencem a A.

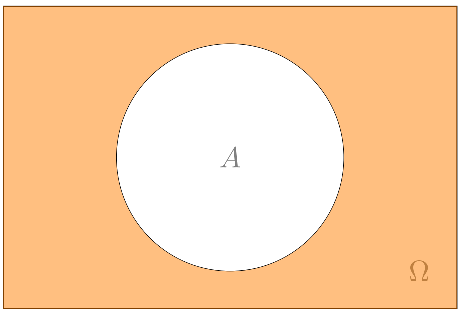


Figura 1.6:

Alguns exemplos de eventos complementares :

(a) $(A \cup B)^c = A^c \cap B^c$

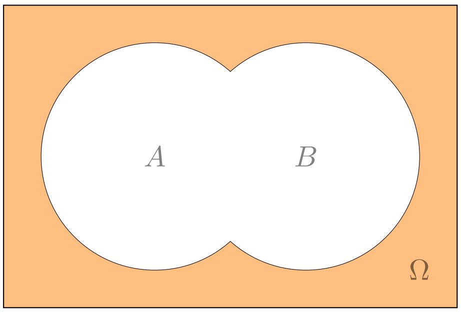


Figura 1.7:

(b) $(A \cap B)^c = A^c \cup B^c$

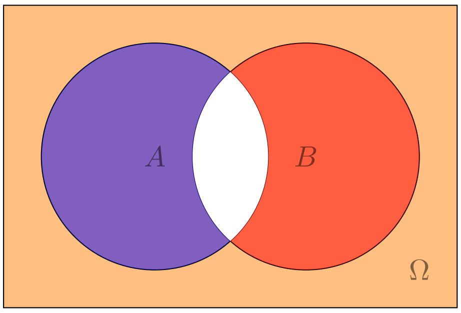


Figura 1.8:

Os itens (a) e (b) são conhecidos como Lei de Demorgan.

(c) $A \cup B^c = (A \cap B) \cup B^c = A \cup (A \cup B)^c$

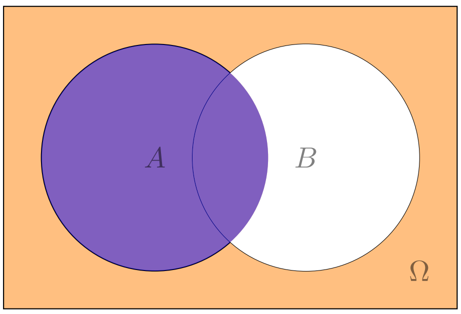


Figura 1.9:

(d) $A \cap B^c = B^c \cap (A \cup B)$

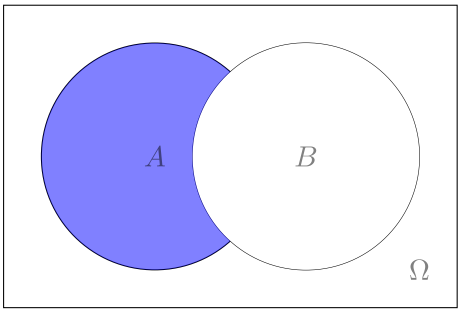


Figura 1.10:

Outras operações :

I. $A \cap \emptyset = \emptyset$
II. $A \cup \emptyset = A$
III. $\emptyset^c = \Omega$
IV. $\Omega^c = \emptyset$
V. $(A^c)^c = A$
VI. $B = (A \cap B) \cup (A^c \cap B)$
VII. $A = (A \cap B) \cup (A \cap B^c)$

Exemplo : Escrever $A \cup B$ como união de eventos disjuntos.

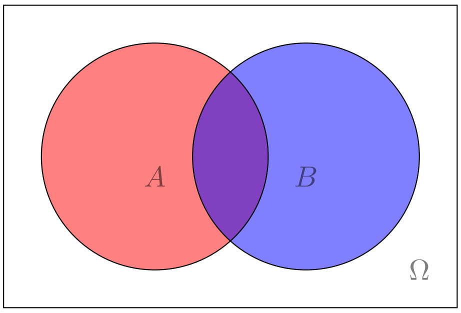


Figura 1.11:

Situação 1: $(A \cup B) = A \cup (A^c \cap B)$
Situação 2: $A \cup B = B \cup (A \cap B^c)$

Considerando a Situação 1 , vamos verificar se os eventos são disjuntos. Os eventos serão disjuntos se $A \cap (A^c \cap B) = \emptyset$. Vericando, temos:

$$(A \cap A^c) \cap B = \emptyset$$
$$\emptyset \cap B = \emptyset$$

1.1 Definições de Probabilidade

1.1.1 Probabilidade em Espaços Equiprováveis

Se um experimento aleatório tiver n resultados possíveis, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, mutuamente exclusivos e igualmente possíveis e se um evento A tiver n_A desses resultados, então a probabilidade do evento A, representado por $P(A)$, é dada por:

$$P(A) = \frac{N_a}{n}, \quad A \subset \Omega \tag{1.1}$$

Sendo que Ω é definido como todo o espaco amostral, n_A o número de casos favoráveis A e n o número de casos possíveis.

Exemplo: Dado o lançamento de duas moedas honestas, calcule a probabilidade de:

(a) Obter duas faces iguais.

(b) Obter pelo menos uma face diferente de cara

(c) Obter pelo menos uma face diferente.

	c	k
c	(c,c)	(c,k)
k	(k,c)	(k,k)

$$\Omega = \{(c,c); (c,k); (k,c); (k,k)\}$$

(a) $A =$ Faces iguais

$$A = \{(c,c); (k,k)\}$$
$$P(A) = \frac{2}{4} = \frac{1}{2} = 0.5$$

(b) $B =$ Pelo menos uma face diferente de cara.

$$B = \{(c,k); (k,c); (k,k)\}$$
$$P(B) = \frac{3}{4}$$

(c) $C =$ Obter pelo menos uma face diferente

$$C = \{(c,k); (k,c)\}$$
$$P(C) = \frac{2}{4} = \frac{1}{2}$$

1.1.2 Probabilidade Frequentista

Um experimento é realizado n vezes, sendo n um número grande. O evento A ocorre exatamente N_a vezes com: $0 \leq N_a \leq n$. A frequência relativa de vezes que ocorreu o evento A é uma forma de aproximar a probabilidade do evento A, ou seja:

$$f_r(A) = \frac{n_a}{n} \tag{1.2}$$

Quando $n \to \infty$, $f_r(A)$ aproxima-se de $P(A)$.

Exemplo :

Geração de n número inteiros entre 1 e 5, $\{1, 2, 3, 4, 5\}$, e o evento de interesse é a ocorrência do número 4.

1.1.3 Probabilidade axiomática

A probabilidade de um evento A é definida como sendo um número $P(A)$ que satisfaz os seguintes axiomas:

I. $P(A) > 0, \forall A \subset \Omega$

II. $P(\Omega) = 1$

III. Se $A_1, A_2, \ldots$ são eventos mutuamente exclusivos ($A_i \cap A_j = \emptyset, \forall i \neq j$), então:

$$P(\cup_{i=1}^{\infty} A_i) = P(A_1 \cup A_2 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i) \tag{1.3}$$

Propriedades :

(a) $0 \leq P(A) \leq 1$

(b) $P(\emptyset) = 0$

(c) Se $A \subset \Omega$ então $P(A) = 1 - P(A^c)$

(d) Se $A \subset B \subset \Omega$, então $P(A) \leq P(B)$

(e) Se $A, B \subset \Omega$, então vale:

$$P(B) = P(B \cap A) + P(B \cup \bar{A}) \tag{1.4}$$

(f) Se $A, B \subset \omega$, então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.5}$$

(g) Se $A, B, C \subset \omega$, então:

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ -P(B \cap C) + P(A \cap B \cap C) \end{aligned} \tag{1.6}$$

Exemplo: Mostre a propriedade (g).

Use o fato de que $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$:

$$\begin{aligned} P(A \cup B \cup C) \\ = P(A \cup (B \cup C)) \\ = P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ = P(A) + P(B) + P(C) - P(B \cap C) - P((A \cap B) \cup (A \cap C)) \\ = P(A) + P(B) + P(C) - P(B \cap C) \\ - (P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C))) \\ = P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C) \end{aligned}$$

Exercício: Considere um experimento aleátorio e os eventos A e B associados, tais que:

$$\begin{aligned} P(A) &= \frac{1}{2} \\ P(B) &= \frac{1}{3} \\ P(A \cap B) &= \frac{1}{4} \end{aligned}$$

Calcule as probabilidades:

(a) $P(\bar{A} \cap \bar{B})$

(b) $P(\bar{A} \cup \bar{B})$

(a)

$$\begin{aligned} P(A^c \cap B^c) = 1 - P(A \cup B) \\ = 1 - \{P(A) + P(B) - P(A \cap B)\} \\ = 1 - \{\frac{1}{2} + \frac{1}{3} - \frac{1}{4}\} = \frac{5}{12} \end{aligned}$$

(b)

$$\begin{aligned} P(A^c \cup B^c) = P((A \cap B)^c) \\ = 1 - P(A \cap B) = 1 - \frac{1}{4} = \frac{3}{4} \end{aligned}$$

Ou de maneira similar:

$$\begin{aligned} P(A^c \cup B^c) &= P(A^c) + P(B^c) - P(A^c \cap B^c) \\ &= (1 - P(A)) + (1 - P(B)) - \frac{5}{12} \\ &= (1 - \frac{1}{2}) + (1 - \frac{1}{3}) - \frac{5}{12} \end{aligned}$$

1.1.4 Probabilidade Condicional

Sejam A e B dois eventos definidos em um mesmo espaco amostral Ω. A probabilidade de A dado que ocorre o evento B, denotada por $P(A/B)$ é definida por:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0. \tag{1.7}$$

Consequentemente, podemos escrever:

Figura 1.12: Name

$$P(A \cap B) = P(A/B).P(B) \tag{1.8}$$

Conhecida como regra do produto.

Observação:

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{P\,(A \cap B)}{P(A)} \tag{1.9}$$

$$P(A \cap B) = P(B/A) \times P(A) \tag{1.10}$$

Logo:

$$\begin{aligned} P(A \cap B) &= P(B/A) \times P(A) &(1.11)\\ &= P(A/B) \times P(B) &(1.12) \end{aligned}$$

Exemplo : Suponha que um escritório possua 100 computadores de tipos Desktop (D) e Laptop (L) sendo alguns novos (N) e outro com um certo tempo de uso (U), distribuídos da seguinte forma:

	D	L	Total
N	40	30	70
U	20	10	30
Total	60	40	100

Um funcionário escolhe um laptop ao acaso. Qual a probabilidade de que seja novo?

Resolução:

$$P(N/L) = \frac{P(N \cap L)}{P(L)} = \frac{\frac{30}{100}}{\frac{40}{100}} = \frac{3}{4}$$

Obs: $P(A \cap B)$ e $P(A/B)$

1.2 Arvore de Probabilidades

Sejam $A, B \subset \Omega$. Uma representação bastante útil é a àrvore de probabilidades.

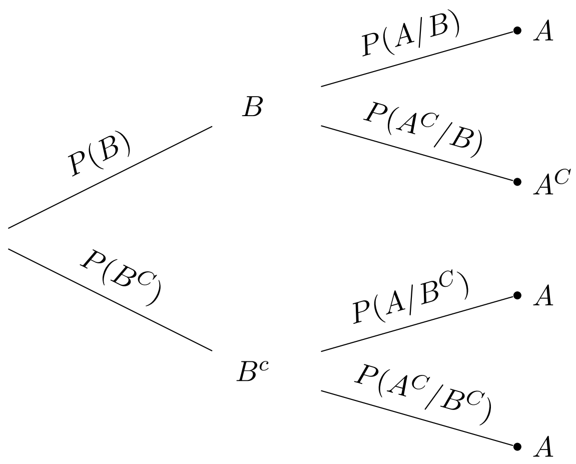


Figura 1.13: Um exemplo de uma àrvore de probabilidades

Exemplo : No exemplo anterior, qual a probabilidade de um funcionário selecionar um desktop usado?

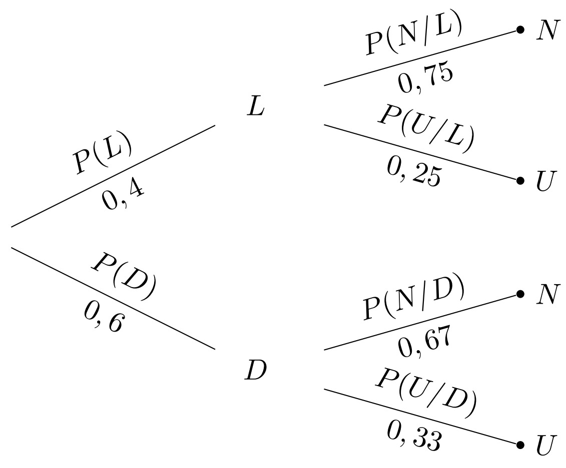


Figura 1.14: Arvore de probabilidade Desktop/Laptop/Usado/Novo

$$P(D \cap U) = P(D/U)P(U)$$

Ou:

$$P(D \cap U) = P(U/D)P(D) = \frac{20}{60} \times \frac{60}{100} = 0,2$$

Algumas propriedades :

(a) $P(\emptyset/B) = 0$

(b) Se $A \subset \Omega$, entao $P(A^c/B) = 1 - P(A/B)$

(c) Se $A, C \subset \Omega$, então:

$$P(A \cup C/B) = P(A/B) + P(C/B) - P(A \cap C/B) \quad (1.13)$$

1.3 Independência de Eventos

Definição : Dois eventos A e B definidos em Ω são independentes se a informação da ocorrência ou não de B não altera a probabilidade de ocorrência de A. Isto é:

$$P(A/B) = P(A), \quad P(B) > 0. \quad (1.14)$$

Logo, dois eventos A e B são independentes se, e somente se, $P(A \cap B) = P(A) \times P(B)$.

Observação :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

Exemplo : Um estudante se inscreve em dois processos seletivos com probabilidade 30% de ser aprovado na empresa I e 50% de ser aprovado na empresa II. Se as aprovações são independentes, qual a probabilidade de que ele seja aprovado em pelo menos uma?

Definindo os eventos:

A: O estudante ser aprovado na empresa I.

B: O esutdante ser aprovado na empresa II.

$$\begin{aligned} P(A) &= 0.30 \\ P(B) &= 0.50 \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A) \times P(B) \\ &= 0.3 + 0.5 - 0,3 \times 0,5 \\ &= 0.65 \end{aligned}$$

1.3.1 Independência de três eventos

Definição: Os eventos A,B,C em Ω são independentes se e somente se:

(a) $P(A \cap B) = P(A) \times P(B)$

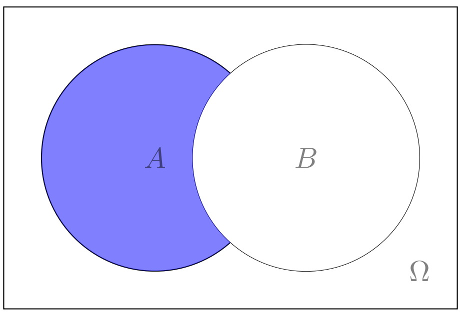
(b) $P(A \cap C) = P(A) \times P(C)$

(c) $P(B \cap C) = P(B) \times P(C)$

(d) $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$

Resultado : Se A,B são eventos independentes em Ω, então:

1. A e B^c são independentes.
2. A^c e B são independetes.
3. A^c e B^c são independentes.



Prova: Resultado 1

$$\begin{aligned}P(A \cap B^c) &= P(A) - P(A \cap B)\\ &= P(A) - P(A) \times P(B)\\ &= P(A)(1 - P(B))\\ &= P(A) \times P(B^c)\end{aligned}$$

Observação : Não confundir eventos mutuamente exclusivos com eventos independentes. Ou seja, não confunda $P(A \cap B) = 0$ com $P(A \cap B) = P(A) \times P(B)$.

Exemplo : Um atirador acerta 80% dos disparos e outro acerta, nas mesmas condições acerta 70%. Qual a probabilidade de o alvo ser acertado se ambos os atiradores disparam simultaneamente?

A: Atirador 1 acerta o alvo

B: Atirador 2 acerta o alvo

Resolução : A intersecção de dois eventos independentes é dada pela multiplicação de suas probabilidades:

$$\begin{aligned}P(A \cap B) &= P(A)P(B)\\ &= 0.8 \times 0.7 = 0.56\end{aligned}$$

Logo a probabilidade de interesse é dada pela a união de dois eventos:

$$\begin{aligned}P(A \cup B) &= P(A) + P(B) - P(A \cap B)\\ &= P(A) + P(B - P(A) \times P(B)\\ &= 0.8 + 0.7 - 0.8 \times 0.7\\ &= 0.94\end{aligned}$$

1.4 O Teorema de Bayes

1.4.1 Partições do espaco amostral

Definição : Uma coleção de eventos $A_1, A_2, \ldots, A_k$ formam uma partição do espaço amostral Ω se:

I. $A_i \cap A_j = \emptyset, \forall i \neq j$, com $i, j = 1, \ldots, k$

II. $\cup_{i=1}^{k} A_i = \Omega$

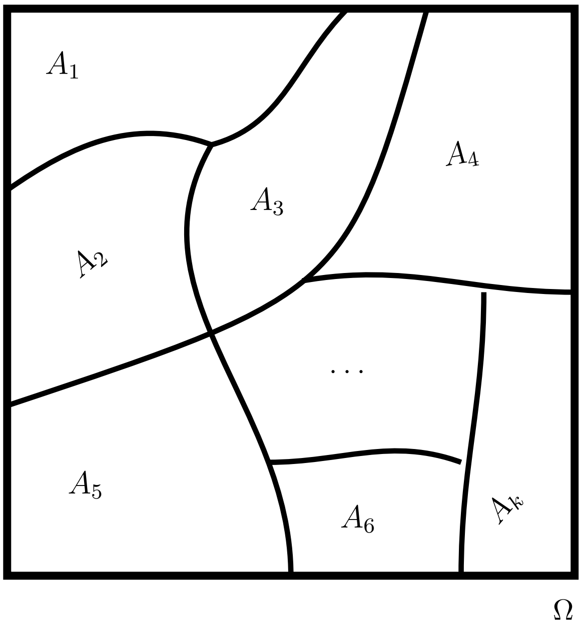


Figura 1.15: Espaco amostral com k particoes

1.4.2 Lema da probabilidade total

Definição: Se $A_1, \ldots, A_k$ é uma partição de Ω, então para qualquer evento B de Ω, vale:

$$P(B) = P\left(\sum_{i=1}^{k} B(B \cap A_i)\right) \tag{1.15}$$
$$= \sum_{i=1}^{k} P(B/A_i)P(A_i).$$

Vejamos:

$$B = \cup_{i=1}^{k} (A_i \cap B) \tag{1.16}$$

$$P(B) = P\left(\cup_{i=1}^{k} (A_i \cap B)\right) \tag{1.17}$$
$$= \sum_{i=1}^{k} P(B/A_i)P(A_i)$$

1.4.3 Fórmula de Bayes

Definição Se $A_1, A_2, \ldots, A_k$ formam uma partição de Ω e $B \subset \Omega$ com $P(B) > 0$, então:

$$\begin{aligned} P(A_i/B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(B/A_i)P(A_i)}{\sum\limits_{j=1}^{k} P(B/A_j)P(A_j)} \end{aligned} \tag{1.18}$$

Exemplo 1 :

Uma montadora trabalha com dois fornecedores A e B de uma determinada peça. Sabe-se que 10% e 5% das peças provenientes dos fornecedores A e B respectivamente, estão fora de especificação. A montadora recebe 30% das peças do fornecedor A e 70% do fornecedor B. Se uma peça do estoque inteiro é escolhido ao acaso, calcule:

(a) A probabilidade que ela esteja fora de especificação.

(b) Se uma peça é escolhida ao acaso está fora de especificação, qual é a píobabilidade de que tenha sido fornecido por A?

A: Peça é do fornecedor A.

B: Peça é do fornecedor B.

C: Peça está fora de especificação.

$$\begin{aligned} P(A) &= 0.3 \\ P(B) &= 0.7 \\ P(C/A) &= 0.10 \\ P(C/B) &= 0.05 \end{aligned}$$

(a) $P(C) =?$

Usando o lema da probabilidade total:

$$\begin{aligned} P(C) &= P(A \cap C) \cup P(B) \\ P(C) &= P(A \cap C) + P(B \cap C) \\ &= P(C/A)P(A) + P(C/B)P(B) \\ &= 0,1 \times 0,3 + 0.05 \times 0.7 \\ &= 0.065 \end{aligned}$$

(b) $P(A/C) =?$

$$
\begin{aligned}
P(A/C) &= \frac{P(A\cap C)}{P(C)} \\
&= \frac{P(C/A)P(A)}{P(C/A)P(A)+P(C/B)P(B)} \\
&= \frac{0.1\times 0.3}{0.065} = 0.4615
\end{aligned}
$$

Exemplo 2: Estudos reveleram que 40% dos estudantes universitários já experimentaram algum tipo de droga ilícita. Uma universidade resolve aplicar um teste com detector de mentira para descobrir se seus estudantes já usaram algum tipo de droga ilícita. Sabemos que se o estudante já usou algum tipo de droga o detector vai dar positivo com certeza. Porém, sabemos que o detector erra, ou seja, apresenta um falso positivo em 5% quando aplicado em estudantes que nunca usaram drogas.
Se um estudante é selecionado aleatoriamente e o teste aplicado nele deu positivo, qual a probabilidade de ele já ter usado algum tipo de droga?

A: O estudante já usou droga.

B: O detector deu positivo.

$$
\begin{aligned}
P(A) &= 0.4 \\
P(B/A) &= 1 \\
P(B/A^c) &= 0.05 \\
P(A/B) &=?
\end{aligned}
$$

$$
\begin{aligned}
P(A/B) &= \frac{P(A\cap B)}{P(B)} \\
&= \frac{P(B/A)P(A)}{P(B/A)\times P(A)+P(B/A^c)\times P(A^c)} \\
&= \frac{1\times 0.4}{1\times 0.4+0.05\times 0.6} \\
&= 0.93
\end{aligned}
$$

Exercício : Para selecionar seus funcionários uma empresa oferece aos candidatos um curso de treinamento durante uma semana. No final do curso, eles são classificados em uma prova; 25% são classificados como bons (B), 50% como médios (M) e os 25% restantes como fracos (F).

A empresa pretende substituir o treinamento por um teste contendo questões de conhecimentos gerais. Para isso gostaria de conhecer qual a probabilidade de um indíviduo aprovado no teste ser considerado fraco (F), se fizesse o curso. Assim, antes do início do curso, os candidatos do curso, foram submetidos ao teste e receberam o conceito aprovado (A) ou reprovado (R). No final do curso, obtiveram-se as seguintes probabilidades condicionais:

$$P(A/B) = 0,8$$
$$P(A/M) = 0,5$$
$$P(A/F) = 0,2$$

Resposta : 0.1

1.5 Variáveis Aleatórias

Definição : Seja um experimento aleatório e Ω o espaço amostral associado a esse experimento. Uma função $X(\omega)$ que associa cada elemento $\omega \in \Omega$ a um número real $x = x(\omega)$ é denominada variável aleatória (v.a.).

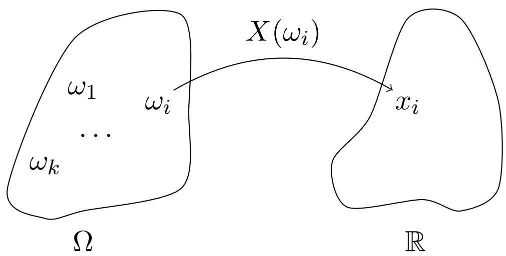


Figura 1.16: Representacao da variavel aleatoria que associa ω_i a um x_i

Notação : $X : \Omega \to \mathbb{R}$

Exemplo : Lançamento de uma moeda duas vezes. A v.a. , x, é o n^o de caras.

	c	k
c	(c,c)	(c,k)
k	(k,c)	(k,k)

$$\Omega : \{\underbrace{cc}_{\omega_1};\underbrace{kc}_{\omega_2};\underbrace{kc}_{\omega_3};\underbrace{kk}_{\omega_4}\}$$

$$X : \quad n^o \text{ de caras}$$

$$X(\omega_1) = X((c,c)) = x(\omega_1) = 2$$

$$X(\omega_2) = X((k,c)) = x(\omega_2) = 1$$

$$X(\omega_3) = X((c,k)) = x(\omega_3) = 1$$

$$X(\omega_4) = X((k,k)) = x(\omega_4) = 0$$

$$R_x : \{0,1,2\}$$

$$P(x=0) =?$$

$$P(x=1) =?$$

$$P(x=2) =?$$

$$X = \begin{cases} 0, & \text{se ocorrer } (k,k) \\ 1, & \text{se ocorrer } (c,k) \text{ ou } (k,c) \\ 2, & \text{se ocorrer } (c,c) \end{cases}$$

Exemplo : Em uma linha de produção, peças são classificadas em defeituosas ou não defeituosas. Podemos definir a v.a. X como:

$$\begin{cases} x = 1, & \text{se a peça é defeituosa} \\ 0, & \text{a peça não é defeituosa} \end{cases}$$

Observação : Uma v.a. X desse tipo é chamada de v.a. de Bernoulli. Nesse caso, $\Omega = \{$peça defeituosa, peça não defeituosa$\}$

1.5.1 Classificação de Variáveis Aleatórias

Definição : Se a v.a. X assume valores em um conjunto finito ou infinito e numerável é chamado de variável aleatória discreta. Se X assume valores, em um conjunto infinito não-enumerável, é chamada de v.a. contínua.

Exemplos :

(a) x indica o n^o de residentes em um domicílio. X pode assumir valores em $\mathbb{N}$ e assim é chamada de v.a. discreta.

(b) Y indica o tempo de vida (em horas) de um equipamento eletrônico. Y pode assumir valores em $\mathbb{R}^+$ e assim é chamado de v.a. contínua.

1.5.2 Função Massa de Probabilidade

Definição : Seja X uma v.a. discreta que assume valores em $R_x = \{x_1, x_2, \ldots, x_k, \ldots\}$. A cada possível x_i, associamos um número,

$$P_i = p(x_i) = P(X = x_i) = P(X(\omega_i) = x_i),$$
$$\omega_i \in \Omega \quad , x_i \in R_x$$

dito probabilidade de x_i. A função $p(x)$ é definida como função massa de probabilidade (f.m.p) de X.

As probabilidades $p(x_i)$ devem satisfazer as seguintes condições:

(i) $p(x_i) > 0, \forall x_i \in R_x$

(ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

Interpretação da f.m.p :

Seja x uma v.a. discreta com $R_x = \{x_1, x_2, \ldots, x_k\}$ e $p(x_i) = p_i$.

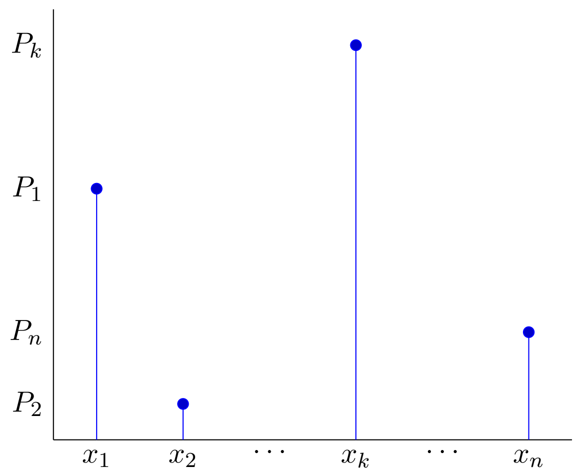


Figura 1.17: Exemplo de uma função massa probabilidade

Exemplo: Lançamento de uma moeda duas vezes e x é o número de caras.

$$\Omega = \{\underbrace{cc}_{2}; \underbrace{ck}_{1}; \underbrace{kc}_{1}; \underbrace{kk}_{0}\}$$
$$R_x = \{0, 1, 2\}$$

A f.m.p de x é dada por:

x	0	1	2
$P(x) = P(X = x)$	1/4	1/2	1/4

$$p(0) = P(X = 0) = P((k, k)) = \frac{1}{4}$$
$$p(1) = P(X = 1) = P((c, k)) + P((k, c)) = \frac{2}{4} = \frac{1}{2}$$
$$p(2) = P(X = 2) = P((c, c)) = \frac{1}{4}$$
$$P(X = x) = \begin{cases} 1/4, & \text{se sair } kk \text{ ou } cc \text{ ou } x = 0 \text{ ou } x = 2 \\ 1/2, & \text{se sair } ck \text{ ou } kk \text{ ou } x = 1 \end{cases}$$

Exemplo: Um carregamento de 8 computadores contém 3 defeituosos. Se uma empresa de uma compra aleatória de dois computadores, apresente a f.m.p para o número de computadores com defeitos adquiridos.

$$X : \text{número de computadores defeituosos}$$
$$R_x = \{0, 1, 2\}$$
$$P(X = 0) = \frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} = 5/14$$

Obs:

$$\binom{8}{2} = \frac{8!}{2!\,(8-2)!}$$

$$P(X = 1) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} = 15/28$$
$$P(X = 2) = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} = 3/28$$

Então a f.m.p. é:

x	0	1	2	Total
$P(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Exemplo: A demanda diária de um item é uma v.a. discreta com f.m.p. dada por:

$$P(D = d) = \frac{2^d k}{d!}, \quad d = 1, 2, 3, 4 \tag{1.19}$$

(a) Determine a constante k

(b) Calcule $P(D > 2)$

(a) Sabemos que:

$$\begin{aligned}
\sum_{i=1}^{n} p(x_i) &= 1, \forall x_i \in R_x \\
P(D=1) + P(D=2) + P(D=3) + P(D=4) &= 1 \\
\frac{2^1 k}{1!} + \frac{2^2 k}{2!} + \frac{2^3 k}{3!} + \frac{2^4 k}{4!} &= 1 \\
2K + 2K + \frac{4K}{3} + \frac{2K}{3} &= 1 \\
4K + \frac{6K}{3} &= 1 \\
6K &= 1 \\
K &= \frac{1}{6}
\end{aligned}$$

Então o f.m.p de x é:

$$P(D = d) = \frac{2^d}{6d!}, \quad d = 1, 2, 3, 4$$

(b)

$$\begin{aligned}
P(D > 2) &= P(D \geq 3) \\
&= P(D = 3) + P(D = 4) \\
&= \frac{2^3}{6 \times 3!} + \frac{2^4}{6 \times 4!} \\
&= \frac{1}{3}
\end{aligned}$$

ou

$$\begin{aligned}
P(D > 2) &= 1 - P(D \leq 2) \\
&= 1 - \{P(D = 1) + P(D = 2)\} \\
&= 1 - P(D = 1) - P(D = 2)
\end{aligned}$$

Obs: $R_x : \{0, 1, 2, 3, 4, 5\}$

$$P(X > 1) = P(x = 2) + P(x = 3) + \cdots + P(x = 5)$$

ou

$$\begin{aligned}
P(x > 1) &= 1 - P(x \leq 1) \\
&= 1 - \{P(x = 0) + P(x = 1)\} \\
&= P(x > 1) + P(x \leq 1) = 1
\end{aligned}$$

1.5.3 Densidade de Probabilidade

Definição: Seja X uma v.a. contínua que assume valores em $R_x, R_x \in \mathbb{R}$. A função $f(x)$ é a função densidade de probabilidade (f.d.p) para x, se satisfaz as seguintes propriedades:

(i) $f(x) \geq 0, \forall x \in R_x$

(ii) $\int\limits_{R_x} f(x)dx = \int\limits_{-\infty}^{\infty} f(x)dx = 1$

(iii) $P(a < x < b) = \int\limits_{a}^{b} f(x)dx$

Uma ilustração de f.d.p:

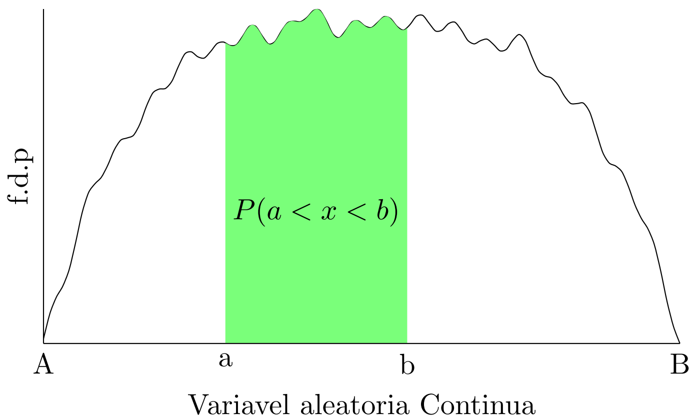


Figura 1.18:

Obs: Se X é uma v.a. contínua assumindo valores em R_x, então para toda $a \in R_x$, temos:

(a) $P(x = a) = 0$

(b) $P(x > a) = P(x \geq a)$

(c) $P(x < a) = P(x \leq a)$

(d) $P(x > a) = 1 - P(x \leq a)$
$= 1 - P(x < a)$

(e) $P(x < a) = 1 - P(x \geq a)$
$1 - P(x > a)$

Outros exemplos:

$$P(a \leq x \leq b) = P(a \leq x < b)$$
$$= P(a < x \leq b) = P(a < x < b)$$

Obs: Só vale para v.a. contínua.

Exemplo: O tempo de produção de um componente (em minutos) é uma v.a. com função densidade de probabilidade dada por:

$$f(x) = \begin{cases} \frac{5-x}{4}, & \text{se } 2 < x < 4 \\ 0, & \text{caso contrário} \end{cases} \tag{1.20}$$

(a) Mostre que $f(x)$ é uma f.d.p

(b) Calcule a probabilidade de que o tempo de produção de um componente seja menor do que 3 minutos.

(a) Devemos verificar:

i. $f(x) \geq 0, \forall x \in R_x$

ii. $\int_{R_x} f(x)dx = 1$

Verificando:

i. $f(x) \geq 0, \forall x \in 2 < x < 4$

ii.

$$\int_2^4 \frac{5-x}{4} dx = 1$$

$$\int_2^4 5dx - \int_2^4 xdx$$

$$= \frac{1}{4}\left(5x|_4^2 - \frac{x^2}{2}\bigg|_4^2\right)$$

$$= \frac{1}{4}\left(5(4-2) - \frac{1}{2}(16-4)\right) = 1$$

Portanto, $f(x)$ é uma função densidade de probabilidade

(b)

$$P(X < 3) = \int_{-\infty}^{(3)} f(x)dx$$

$$\int_{-\infty}^{2} 0dx + \int_{2}^{3} \frac{5-x}{4} dx$$

$$
\begin{aligned}
P(X < 3) &= \int\limits_{2}^{3} \frac{5-x}{4} dx \\
&= \frac{1}{4}\left(\int\limits_{2}^{3} 5dx - \int\limits_{2}^{3} xdx\right) \\
&= \frac{1}{4}\left(5 \times 1 - \frac{1}{2} \times 5\right) = \frac{5}{8} \\
P(A/B) &= \frac{P(A \cap B)}{P(B)}
\end{aligned}
$$

Exemplo: Seja x uma v.a. contínua com f.d.p dada por:

$$
f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{Caso contrário} \end{cases}
$$

(a) Verifique se $f(x)$ é uma f.d.p

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

(c) $P\left(X \leq \frac{1}{2} / \frac{1}{3} \leq x \leq \frac{2}{3}\right)$

(a) Devemos verificar:

i. $f(x) \geq 0, \forall x \in R_x$

ii. $\int_{R_x} f(x)dx = 1$

Verificando:

i.

$$
f(x) \geq 0, \forall x \in 0 < x < 1
$$

ii.

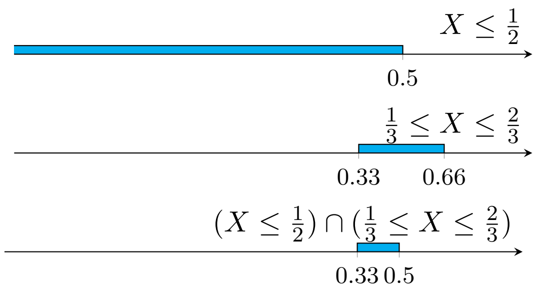
$$
\begin{aligned}
\int\limits_{R_x} 2xdx &= \int\limits_{0}^{1} 2xdx \\
&= \left.\frac{2x^2}{2}\right|_{1}^{0} = 1
\end{aligned}
$$

Portanto, $f_x(x)$ é uma f.d.p.

(b)

$$
\begin{aligned}
P(X \leq 1/2) &= \int_{-\infty}^{1/2} f_x(x)dx \\
&= \int_{-\infty}^{0} 0dx + \int_{0}^{1/2} 2xdx \\
&= x^2\Big|_{1/2}^{0} \\
&= \frac{1}{4}
\end{aligned}
$$

(c) Probabilidade Condicional



$$
\begin{aligned}
P(A/B) &= \frac{P(A \cap B)}{P(B)} \\
\frac{P\left(X \leq \frac{1}{2}\right) \cap P\left(\frac{1}{3} \leq X \leq \frac{2}{3}\right)}{P\left(\frac{1}{3} \leq X \leq \frac{2}{3}\right)} & \\
&= \frac{P(\frac{1}{3} \leq X \leq \frac{1}{2})}{P(\frac{1}{3} \leq X \leq \frac{2}{3})} \\
&= \frac{\int_{\frac{1}{3}}^{\frac{1}{2}} 2xdx}{\int_{\frac{1}{3}}^{\frac{2}{3}} 2xdx} \\
&= \frac{x^2\big|_{\frac{1}{2}}^{\frac{1}{3}}}{x^2\big|_{\frac{2}{3}}^{\frac{1}{3}}} \\
&= \frac{5}{12}
\end{aligned}
$$

1.5.4 Função distribuição acumulada discreta

Definição: Seja x uma v.a. discreta que assume valores em R_x e com f.m.p. $p(x) = P(X = x)$. Para qualquer $x \in \mathbb{R}$, a função de distribuição acumulada(f.d.a) de x, denotada por $F_x(x)$, é definida como:

$$F_x(x) = P(X \leq x) = \sum_{x_i \in R_x} P(X = x_i) = \sum_{x_i \in R_x} p(x_i), \quad \forall x_i \leq x$$

De forma mais clara, temos como exemplo:

$$\begin{aligned} R_x &= \{0, 1, 2, 3, 4\} \\ F_x(2,3) &= P(X \leq 2,3) \\ &= P(x = 0) + P(x = 1) + P(x = 2) \end{aligned}$$

Exemplo: Conidere o lançamento de uma moeda duas vezes e x é o número de caras. Já sabemos que a f.m.p de x é dada por:

x	0	1	2
$P(X = x)$	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{1}{4}$

Com $R_x = \{0, 1, 2\}$
Ou, de forma equivalente podemos escrever:

$$p(x) = P(X = x) = \begin{cases} \frac{1}{4}, & \text{se} \quad x = 0, 2 \\ \frac{1}{2}, & \text{se} \quad x = 1 \\ 0, & \text{caso contrário} \end{cases}$$

A f.d.a de x é dada por:

$$F_x = P(X \leq x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

A representação gráfica de $F_x(x)$ é:

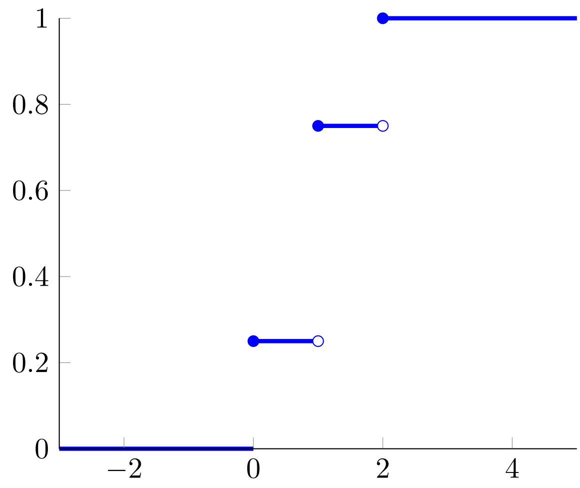


Figura 1.19:

1.5.5 Função distribuição acumulada contínua

Definição: Seja x uma v.a. contínua que assume valores em R_x e com f.d.p $f(x)$. Para qualquer $x \in \mathbb{R}$, a f.d.p de x, denotada por $F_x(x)$ é definida como:

$$F_x(x) = P(X \leq x) = \int\limits_{-\infty}^{x} f(t)dt \tag{1.21}$$

Observação: $F_x(x) = P(X \leq a) = \int\limits_{-\infty}^{a} f(t)dt, \quad a \in \mathbb{R}$

Como consequência imediata podemos escrever dois resultados:

$$P(a < x < b) = \int\limits_{a}^{b} f(x)dx \tag{1.22}$$
$$= F_x(b) - F_x(a) \tag{1.23}$$

$$f(x) = \frac{\partial F_x(x)}{\partial x}, \quad \text{se a derivada existir} \tag{1.24}$$

Exemplo: Considere a f.d.p dada por:

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

Determine $F_x(x)$ e use-a para avaliar $P(0 \leq x < 1)$.

A f.d.a de x é dada por:

$$F_x(x) = P(X \leq x) = \int\limits_{-\infty}^{x} f_x(t)dt$$

$$= \int\limits_{-\infty}^{x} \frac{t^2}{3}dt, \quad -1 < t < 2$$

$$\int\limits_{-1}^{x} \frac{t^2}{3}dt = \frac{1}{9}\, t^3\Big|_{-1}^{x}$$

$$F_x(x) = \frac{1}{9}\left(x^3 + 1\right)$$

$$F_x(x) = P(X \leq x) = \begin{cases} 0, x < -1 \\ \frac{1}{9}(x^3+1), -1 \leq x < 2 \\ 1, x \geq 2 \end{cases}$$

$$P(0 \leq x < 1) = ?$$

$$P(0 \leq x < 1) = F_x(1) - F_x(0)$$

$$P(x \leq 1) - P(x \leq 0) = \frac{1}{9}(1^3+1) - \frac{1}{9}(0^3+1)$$

$$P(0 \leq x < 1) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

Ou

$$P(0 \leq X < 1) = \int\limits_{0}^{1} \frac{x^2}{3}dx$$

1.5.6 Propriedades de f.d.a

(a) $F_x(x)$ é uma função contínua.

(b) $F_x(x)$ é uma função monótona não decrescente.

(c) $0 \leq F_x(x) \leq 1, \quad \forall x \in \mathbb{R}$

(d) $\lim\limits_{x \to -\infty} F_x(x) = 0$ e $\lim\limits_{x \to +\infty} F_x(x) = 1$

(e) $P(x \leq a) = F_x(a)$

(f) $P(x > a) = 1 - P(x \leq a) = 1 - F_x(a)$

(g) $P(a < x \leq b) = F_x(b) - F_x(a)$

Exemplo: Seja x uma v.a. contínua com f.d.p dada por:

$$f_x(x) = \begin{cases} kx^2, 0 < x < 1 \\ 0, \quad \text{c.c.} \end{cases}$$

(a) Achar k.

(b) Determine $F_x(x)$.

(c) $P(\frac{1}{3} < x < \frac{1}{2})$

(a)

$$\int\limits_{R_x} f(x)dx = 1$$

$$k\int\limits_0^1 x^2dx = 1$$

$$\frac{k}{3}\, x^3\Big|_0^1 = 1$$

$$\frac{k}{3} = 1 \therefore k = 3$$

Logo, a f.d.p de X é:

$$f_x(x) = \begin{cases} 3x^2, \quad 0 < x < 1 \\ 0, \quad \text{c.c.} \end{cases}$$

(b) $F_x(x) = ?$

$$F_x = P(X \leq x) = \int\limits_{-\infty}^{x} f_x(t)dt$$

$$= \int\limits_0^x 3t^2dt = t^3\Big|_x^0 = x^3$$

Assim, a f.d.a. é dada por:

$$F_x(x) = \begin{cases} 0, \quad x < 0 \\ x^3, \quad 0 \leq x < 1 \\ 1, \quad x \geq 1 \end{cases}$$

(c)

$$
\begin{aligned}
P\left(\frac{1}{3} < X < \frac{1}{2}\right) &= F_x\left(\frac{1}{2}\right) - F_x\left(\frac{1}{3}\right) \\
&= \left(\frac{1}{2}\right)^3 - \left(\frac{1}{3}\right)^3 \\
&= \frac{1}{8} - \frac{1}{27} \\
&= \frac{19}{216}
\end{aligned}
$$

Ou

$$P\left(\frac{1}{3} < x < \frac{1}{2}\right) = \int_{\frac{1}{3}}^{\frac{1}{2}} 3x^2 dx = \frac{19}{216}$$

Exemplo:

Seja F_x dada por:

$$F_x = \begin{cases} 0, & \text{se } x < 0 \\ \frac{1}{8}, & \text{se } 0 \leq x < 1 \\ \frac{1}{2}, & \text{se } 1 \leq x < 2 \\ \frac{5}{8}, & \text{se } 2 \leq x < 3 \\ 1, & \text{se } x \geq 3 \end{cases}$$

Determine:

(a) $P(1 < x \leq 3)$

(b) $P(x > 2)$

(c) Encontre a $P(x)$

(a)

$$P(1 \leq X \leq 3) = F_x(3) - F_x(1) = \frac{1}{2}$$

(b)

$$
\begin{aligned}
P(X > 2) &= 1 - F_x(2) \\
= 1 - F_x(2) &= 1 - \frac{5}{8} = \frac{3}{8}
\end{aligned}
$$

(c)

$$F_x(0) = P(X \leq 0) = \frac{1}{8}$$
$$P(X = 0) = \frac{1}{8}$$

$$F_x(1) = P(X \leq 1) = \frac{1}{2}$$
$$P(X = 0) + P(X = 1) = \frac{1}{2}$$
$$P(X = 1) = \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$$

$$F_x(2) = P(X \leq 1) = \frac{5}{8}$$
$$P(X = 0) + P(X = 1) + P(X = 2) = \frac{5}{8}$$
$$F_x(1) + P(X \leq 2) = \frac{5}{8}$$
$$P(X = 2) = \frac{5}{8} - \frac{1}{2} = \frac{1}{8}$$

$$F_x(3) = P(X \leq 3) = 1$$
$$\underbrace{P(X = 0) + P(X = 1) + P(X = 2)}_{F_x(2)} + P(X = 3) = 1$$
$$P(X = 3) = 1 - \frac{5}{8} = \frac{3}{8}$$

Assim temos:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$

1.6 Esperança Matemática de uma Variável Aleatória

Definição: Seja X uma v.a. com f.m.p. $p_x(x)$ (no caso discreto) ou f.d.p $f_x(x)$ (no caso contínuo). Chamamos de esperança matemática ou valor médio de X ao valor:

$$\mu_x = E(X) = \sum_{x \in R_x} x p_x(x), \quad \text{no caso discreto} \tag{1.25}$$

$$\mu_x = E(X) = \int_{x \in R_x} x f_x(x) dx, \quad \text{no caso contínuo} \tag{1.26}$$

Considerando $a, b \in \mathbb{R}$, constantes, temos algumas propriedades:

(a)

$$E(aX) = aE(x) \tag{1.27}$$

(b) Se $X = a$,

$$E(X) = E(a) = a \tag{1.28}$$

(c)

$$E(E(X)) = E(X) \tag{1.29}$$

(d)

$$E(\pm a) = E(X) \pm a \tag{1.30}$$

(e)

$$E(ax + b) = aE(X) + b \tag{1.31}$$

(f)

$$E(ax + by) = aE(X) + bE(Y) \tag{1.32}$$

Exemplo: Considere a v.a. x com f.d.p dada por:

$$f_x(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Calcular o $E(X)$:

$$E(X) = \int_{R_x} x f_x(x) dx$$

$$E(X) = \int_0^1 x \times 2x dx = 2 \int_0^1 x^2 dx$$
$$= \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3}$$

Resultado: Seja X uma v.a. com f.m.p $p_x(x)$ ou f.d.p $f_x(x)$. Uma função de X, dita $g(x)$, é também uma v.a. e:

$$\mu = E\left(g(x)\right) = \sum_{x \in R_x} g(x) p_x(x) = \sum_{x \in R_x} g(x) P(X = x), \quad \text{no caso discreto}$$
$$\mu = E\left(g\left(x\right)\right) = \int_{R_x} g(x) f_x(x) dx, \quad \text{no caso contínuo} \tag{1.33}$$

1.7 Variância

Definição: Seja x uma v.a. com f.m.p $p_x(x)$ ou f.d.p $f_x(x)$, com média $\mu_x = E(x)$. Chamamos de variância da v.a. X o valor:

$$\sigma^2 = var(x) = E\left((X - E(X))^2\right) \tag{1.34}$$
$$= E\left((x-\mu)^2\right), \tag{1.35}$$

Ou seja,

$$\sigma^2 = var(x) = \sum_{x \in R_x} (x-\mu_x)^2 P(X = x) \quad \text{no caso discreto)} \tag{1.36}$$
$$\sigma^2 = var(x) = \int_{R_x} (x-\mu_x)^2 f_x(x)dx \quad \text{no caso contínuo} \tag{1.37}$$

A raíz quadrada da variância($\sqrt{\sigma^2} = \sqrt{var(x)}$) é o desvio padrão, denotado por σ.

Resultado: Podemos escrever a variância v.a. x por:

$$\sigma^2 = var(x) = E(X^2) - (E(X))^2 \tag{1.38}$$
$$\sigma^2 = E(X^2) - (\mu_x)^2 \tag{1.39}$$

Demonstração:

$$var(x) = E(x^2 - 2\mu_x x + \mu_x^2) \tag{1.40}$$
$$= E(x^2) - 2\mu_x E(x) + \mu_x^2 \tag{1.41}$$
$$= E(x^2) - 2\mu_x \mu_x + \mu_x^2 \tag{1.42}$$
$$= E(x^2) - \mu_x \tag{1.43}$$
$$= E(x^2) - (E(x))^2 \tag{1.44}$$

Obs: $E(x^2) \neq (E(x))^2$

Propriedades Considerando $a, b \in \mathbb{R}$, constantes:

(a)

$$Var(ax) = a^2 Var(x)$$

Obs: $Var(-x) = Var(x)$

(b) Se $x = a$, então:

$$Var(x) = Var(a) = 0$$

(c)

$$Var(x \pm a) = Var(x)Var(a) = Var(x)$$

(d) Se a, b são constantes,

$$\begin{aligned} Var(ax + \pm + b) &= a^2 Var(x) + Var(b) \\ &= a^2 var(x) \end{aligned}$$

(e) Se x e y são duas v.a.'s independetes, então:

$$Var(ax \pm by) = a^2 Var(x) + b^2 Var(y)$$

Exemplo: Seja x uma v.a. com f.d.p dada por:

$$f_x = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{caso contrário} \end{cases} \tag{1.45}$$

Encontre :

(a) $E(X)$
(b) $Var(X)$
(c) $E(4X + 3)$
(d) $Var(4X + 3)$

Resolução :

(a)

$$\begin{aligned} \mu = E(X) &= \int_{-1}^{2} \frac{x \times x^2}{3)} dx \\ &= \frac{1}{3} \int_{-1}^{2} x^3 dx = \frac{1}{12} \left. x^4 \right|_{-1}^{2} \\ &= \frac{1}{12}(16 - 1) = \frac{15}{12} \end{aligned}$$

(b)

$$\begin{aligned} Var(x) &= E(x^2) - (E(x))^2 \\ E(x^2) &= \int_{-1}^{2} x^2 \frac{x^2}{3} dx = \frac{1}{3} \int -12x^4 dx \\ &= \frac{1}{15} \left. x^5 \right|_{-1}^{2} = \frac{1}{15}(32 + 1) = \frac{33}{15} \\ Var(x) &= \frac{33}{15} - \left(\frac{15}{12} \right)^2 = 0,6375 \end{aligned}$$

(c)

$$E(4x+3) = 4E(x) + 3 = 8$$

Obs: $E(ax) = a^2 var(x)$

(d)

$$Var(4x+3) = 16Var(x) = \frac{51}{80} \times 16 = 10,2$$

Capítulo 2

Principais Modelos Probabilísticos

2.1 Modelos Discretos

2.1.1 Distribuição Uniforme Discreta

A v.a. assume cada um de seus valores com igual probabilidade.

Definição: A v.a. discreta X, assumindo valores $x_1, x_2, \ldots, x_k$ tem distribuição uniforme se, e somente se, sua f.m.p. é definida por:

$$p(x) = P(X = x) = \begin{cases} \frac{1}{k}, & x = x_1, x_2, \ldots, x_k \\ 0, & \text{caso contrário} \end{cases} \tag{2.1}$$

Notação: $X \sim U_d(k)$ A média e variância de x é:

$$\mu_x = E(x) = \sum_{x \in R_x} x_i P(X = x_i) \tag{2.2}$$

$$= \sum_{x \in R_x} x_i \frac{1}{k} = \frac{1}{k} \sum_{x_i \in R_x} x_i$$

$$Var(x) = E(X^2) - (E(X))^2 \tag{2.3}$$

$$= \sum_{x_i \in R_x} (x_i - \mu_x)^2 \times p\,(X = x_i)$$

$$= \frac{1}{k} \left(\sum_{i=1}^{k} x_i^2 - \frac{(\sum_{i=1}^{k})^2}{k} \right)$$

Exemplo: Seja x uma v.a. que indica o n^o de pontos marcados na face superior de um dado quando ele é lançado. Portanto, temos uma distribuição uniforme discreta, cuja f.m.p. é dada por:

$$p(x) = P(X = x) - \begin{cases} \frac{1}{6}, & x = 1, 2, 3, 4, 5, 6 \\ 0, & \text{caso contrário} \end{cases} \tag{2.4}$$

E o $E(x)$ é:

$$\mu_x = E(x) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) \tag{2.5}$$

$$= \frac{1}{6}21 = \frac{7}{2} = 3,5 \tag{2.6}$$

E a $Var(x)$ é:

$$E(x^2) = \frac{1}{6}(1^2 + 2^2 + \cdots + 6^2) \tag{2.7}$$

$$= \frac{91}{6} = 15,17 \tag{2.8}$$

$$var(x) = \frac{91}{6} - \{\frac{7}{2}\}^2 = \frac{105}{36} = 2,92 \tag{2.9}$$

2.1.2 Distribuição Bernoulli

Considere uma v.a. x que assume apenas dois valores 1, se ocorrer sucesso,0, se ocorrer fracasso. Indicaremos por p a probabilidade de sucesso, isto é $P(\text{Sucesso}) = P(x = 1) = p$.

Definição: A v.a. discreta x que assume apenas valores 0 ou 1 tem Distribuição Bernoulli se, somente se, sua f.m.p. é definida por:

$$p(x) = P(X = x) = \begin{cases} p^x(1-p)^{1-x}, & x = 0, 1, \quad 0 < p < 1 \\ 0, & \text{caso contrário} \end{cases} \tag{2.10}$$

Notação: $X \sim Ber(p)$

Exemplos:

$$p(x) = \left(\frac{2}{5}\right)^x \left(1 - \frac{2}{5}\right)^{1-x} \quad , x = 0, 1$$

$$p(x) = \left(\frac{7}{8}\right)^x \left(1 - \frac{7}{8}\right)^{1-x} \quad , x = 0, 1$$

De forma similar, podemos apresentar $p(x)$ por:

x	0	1
$P(X = x)$	$1 - p$	p

x	0	1
$P(X = x)$	$\frac{5}{6}$	$\frac{1}{6}$

A média de x é:

$$\mu_x = E(x) = p \tag{2.11}$$

e a variância de x é:

$$\sigma^2 = var(x) = p(1-p) \tag{2.12}$$

Exemplo: Suponha o lançamento de um dado perfeito e o interesse é ocorrer a face 3.

$$X = \begin{cases} 1, & \text{se ocorrer a face 3} \\ 0, & \text{caso contrário}\{1,2,4,5,6\}] \end{cases}$$

Cuja f.m.p. é dada por:

$$p(x) = P(X = x) = \begin{cases} \frac{1}{6}(1-\frac{1}{6})^{1-x}, & x = 0, 1 \\ 0, & \text{caso contrário} \end{cases} \tag{2.13}$$

$$X \sim Ber(\frac{1}{6})$$

A média de x é:

$$\mu_x = (X) = p = \frac{1}{6}$$

E a variância de x é:

$$\sigma^2 = var(x) = p(1-p) = \frac{1}{6} \times \frac{5}{6} = \frac{5}{36}$$

2.1.3 Distribuição Binominal

Considere M ensaios de Bernoulli independentes, todos com a mesma probabilidade de sucesso p. A v.a. que conta o número de sucessos nos M ensaios de Bernoulli é denominada v.a. binomial com parâmetro M e p.

Definição: A v.a. discreta x, correspondente ao n^o de sucesso em M ensaios independentes de Bernoulli, tem distribuição binomial se, e somente se, sua f.m.p. é definida por:

$$p(x) = P(X = x) = \begin{cases} \binom{M}{x} p^x (1-p)^{M-x}, & x = 0, 1, \ldots, M \quad ,0 < p < 1 \\ 0, & \text{Caso contrário} \end{cases} \tag{2.14}$$

Notação:$X \sim Bin(M, p)$

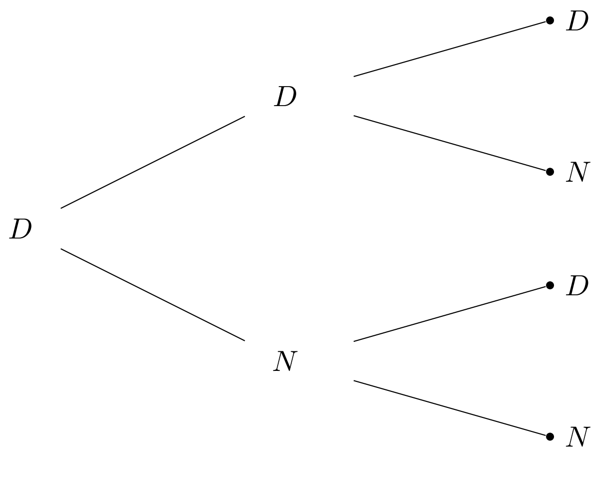
Obs: $\binom{M}{x} = \frac{M!}{x!(M-x)!}$

Temos que:

$$\mu_x = E(X) = Mp \tag{2.15}$$
$$\sigma^2 = var(x) = Mp(1-p) \tag{2.16}$$

Exemplo : Considere uma linha de produção, onde 3 peças são selecionadas aleatoriamente e são classificadas como defeituosas (D) ou não-defeituosas(N).
X_1, X_2, X_3 são variáveis aleatórias que assumem 1 se a peça for não-defeituosa e 0 caso contrário. A probabilidade da peça ser não-defeituosa é p e , consequentemente, a probabilidade da peça ser defeituosa é $1-p$.
Estamos interessados na distribuição de:$Y = X_1 + X_2 + X_3$

$$X_1 = \begin{cases} 1, & \text{Peça não deituosa} \\ 0, & \text{Peça deituosa} \end{cases}$$



	X_1	X_2	X_3	$Y = X_1 + X_2 + X_3$	$P(Y = y)$
DNN	0	1	1	2	$p^2(1-p)$
DND	0	1	0	1	$p(1-p)^2$
DDN	0	0	1	1	$p(1-p)^2$
DDD	0	0	0	0	$(1-p)^3$
NNN	1	1	1	3	p^3
NND	1	1	0	2	$p^2(1-p)$
NDN	1	0	1	2	$p^2(1-p)$
NDD	1	0	0	1	$p(1-p)^2$

Abaixo estão os possíveis resultados do experimento:

y	0	1	2	3
$P(Y = y)$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3

Ou seja,

$$p(y) = P(Y = y) = \begin{cases} \binom{3}{y} p^y (1-p)^{3-y}, & y = 0, 1, 2, 3 \\ 0, & \text{Caso contrário} \end{cases}$$

Exemplo :

Uma rede varejista compra certo tipo de equipamento eletrônico. O fabricante indica que a taxa de quipamentos em perfeito estado é 97%.

(a) Seleciona-se ao acaso 20 itens. Qual a probabilidade de haver pelo menos um item defeituoso nesses 20?

(b) Seleciona-se aleatoriamente 20 itens em cada 10 carregamentos, qual a probabilidade de haver 3 carregamentos com pelo menos um item defeituoso?

p taxa de equipamentos em perfeito estado

$$p = 0,97$$

y: n^o de equipamentos em perfeito estado.

$$y \sim Binom(20; 0,97)$$

Resolução :

(a)

$$
\begin{aligned}
M = & \quad 20 \text{ itens} \\
p: & \quad \text{Probabilidade de pelo menos um item defeituoso.} \\
& p = 0,03
\end{aligned}
$$

$$X_i = \begin{cases} 1, & \text{se o equipamente e defeituoso} \\ 0, & \text{c.c.} \end{cases}$$

$$X \sim Binom(20 : 0,03) \quad x = 1, 2, 3, 4, 5, \ldots, 20$$

$$Y = X_1 + X_2 + \ldots + X_20$$

$$
\begin{aligned}
P(Y \geq 1) &= 1 - P(Y < 1) \\
&= 1 - P(Y = 0)
\end{aligned}
$$

$$P(Y = 0) = \binom{20}{0}(0,03)^0(1 - 0,03)^{20-0} = 0,5438$$

$$
\begin{aligned}
P(Y \geq 1) &= 1 - 0,5438 \\
&= 0,4562
\end{aligned}
$$

Ou, resolvendo pela v.a. Y, temos:

$$X_i = \begin{cases} 1, & \text{se o equipamento for perfeito} \\ 0, & \text{se o equipamente for defeituoso} \end{cases}$$

$$
\begin{aligned}
P(Y \leq 19) &= 1 - P(Y > 19) \\
&= 1 - P(Y = 20)
\end{aligned}
$$

$$
\begin{aligned}
P(y = 20) = \binom{20}{20} 0,97^{20}(1 - 0,97)^{20-20} \\
= 0,5438
\end{aligned}
$$

$$
\begin{aligned}
P(y \leq 19) &= 1 - 0,5339 \\
&= 0,4562
\end{aligned}
$$

(b)

10 Carregamentos

20 itens são selecionados

Y : n^o de carregamento com pelo menos um item defeituoso

$M = 10$

$Y = 0, 1, 2, \ldots, 10$

p : Proporção de ter pelo menos um item defeituoso em um carregamento

$p = 0,4562$

$Y \sim B(10 : 0,4562)$

$$P(Y = 3) = \binom{10}{3}(0,4562)^3(1 - 0,4562)^{10-3} = 0,1602$$

2.1.4 Distribuição Geométrica

Considere uma sequência de ensaios Bernoulli independentes em probabilidade de sucesso p $(0 < p < 1)$. Seja a v.a. x o n^o de fracassos até a ocorrência do 1^o sucesso. Similarmente, a v.a. x pode ser vista como o n^o de ensaios que precedem o primeiro sucesso.

Definição : A v.a. x tem distribuição geométrica se, e somente se, sua f.m.p. é definida por:

$$p_x(x) = P(X = x) = \begin{cases} p(1-p)^x, & x = 0, 1, 2, \ldots \quad 0 < p < 1 \\ 0, & \text{caso contrário} \end{cases} \tag{2.17}$$

Notação: $X \sim Geo(p)$

Temos que:

$$\mu_x = E(x) = \frac{1-p}{p} \quad \text{e a} \tag{2.18}$$

$$\sigma^2 = var(x) = \frac{1-p}{p^2} \tag{2.19}$$

Similarmente, a variável aleatória Y pode ser vista como o n^o de ensaios que precedem o primeiro sucesso assim, Y tem a distribuição geométrica com f.m.p. dada por:

$$p_y(y) = P(Y > y) = \begin{cases} (1-p)^{y-1} & p, y = 1, 2, \ldots \\ 0, & \text{Caso contrário} \end{cases} \tag{2.20}$$

Neste caso, a v.a. Y pode ser vista como $Y = X + 1$ e, consequentemente, o valor esperado de y é:

$$mu_y = E(Y) = E(X + 1) = E(X) = E(X) + 1 = \frac{1}{p} \tag{2.21}$$

E a variância de Y e:

$$\sigma_y^2 = var(Y) = var(X + 1) = var(X) = \frac{1 - p^2}{p} \tag{2.22}$$

Exemplo : Um pesquisador está realizando um experimentos químicos independetes e sabe que a probabilidade de que cada cada experimento apresnete uma reação positiva é $0,3$. QUal é a probabilidade de que menos de 3 reações negativas ocorram antes da primeira positiva?

x:n^o de ocorrência positiva

$$x = \begin{cases} 1, & \text{se a reação positiva} \\ 0, & \text{caso contrário} \end{cases}$$

$$P(x = 1) = p = 0,3$$

X : n^o de reações negativas até a ocorrência de uma positiva

$$Y \sim Geo(0,3)$$

$$\begin{aligned} P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) \\ P(X = 0) = (1 - 0,3)^0 \times 0,3 = 0,3 \\ P(X = 1) = (1 - 0,3)^1 \times 0,3 = 0,21 \\ P(X = 2) = (1 - 0,3)^2 \times 0,3 = 0,147 \\ P(X < 3) = 0,3 + 0,21 + 0147 = 0,657 \end{aligned}$$

Ou ainda,

$$\begin{aligned} Y : n^o \quad \text{de ensaios até a ocorrência de reação positiva} \\ R_y = \{1, 2, 3, \ldots\} \\ P(Y < 4) = P(Y \leq 3) = P(Y = 1) + P(Y = 2) + P(Y = 3) \\ P(Y = y) = p(1 - p)^{y-1} \quad y = 1, 2, \ldots \end{aligned}$$

2.1.5 Distribuição Binomial Negativa

Considere ensaios independetes de Bernoulli (p) e definimos X como o n^o de fracassos anteriores ao r-ésimos sucesso.

Definição : A v.a. x tem distribuição Binomial negativa se, e somente se, sua f.m.p. é definida por:

$$p(x) = P(X = x) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{x=0,1,2,...} \\ 0, & \text{caso contrário} \end{cases} \tag{2.23}$$

Notação: $X \sim BN(r, p)$

Temos que:

$$\mu_x = E(x) = \frac{r(1-p)}{p} \tag{2.24}$$

$$\sigma^2 = var(x) = \frac{1-p}{p^2} \tag{2.25}$$

Observações :

1. Note que se $r = 1$, temos o modelo geométrico.
2. $$\binom{x+r-1}{r-1} = \binom{x+r-1}{x}$$

A Binomial Negativa pode ser definida como o n^o de ensaios necessários para a obtenção do r-ésimo sucesso. Formando $y = x + r$ temos a quantidade desejada e seus valores variam de r em diante. Assim sendo Y o n^o de ensaios até a obtenção de r sucessos, temos:

$$p(y) = p(Y = y) = \begin{cases} \binom{y-1}{m-1} p^m (1-p)^{y-m}, & y = m, m+1, \ldots \\ 0, & \text{caso contrário} \end{cases} \tag{2.26}$$

Exemplo : Em uma serie do campeonato de basquete, o time que ganhar quatro em sete jogos sera o vencedor. Suponha que o time A tenha probabilidade de 55% de ganhar de B e que A e B se enfrentarao em uma serie de sete jogos. Qual a probabilidade de que A venca a serie em 6 jogos?

X : O número de derrotas de A até que A ganhe 4 partidas

$$X \sim BN(r = 4, p = 0,55)$$

$$P(X=2)=\binom{2+4-1}{4-1}0,55^4\times 0,45^2=0,1853$$

Ou, de maniera similar, temos:

$$Y:\ \text{número de partidas até que } A \text{ venca o campeonato}$$

$$P(Y=6)=\binom{6-1}{4-1}0,55^4\times 0,45^{6-4}=0,1853$$

2.1.6 Distribuição Hipergeometrica

Considere um conjunto de n objetos dos quais m são do tipo I e $(n-m)$ são do tipo II. Para um sorteio de r objetos $(r<n)$, feitos ao acaso e sem repeticao, defina X como o número de objetos do tipo I selecionados.

Definição: A v.a. X tem distribuição Hipergeometrica se sua f.m.p. é dada por:

$$P_x(x)=P(X=x)=\frac{\binom{m}{xs}\binom{n-m}{r-x}}{\binom{n}{r}}$$

Em que x (inteiro) é tal que:

$$max\{0,r-(n-m)\}\leq x\leq min\{r,m\}$$

Observação : Os limites de x garantem que situaçoes absurdas ocorram.

Temos que:

$$E(X)=r\frac{m}{n}$$
$$Var(x)=\frac{r\times m(n-m)(n-r)}{n^2(n-1)}$$

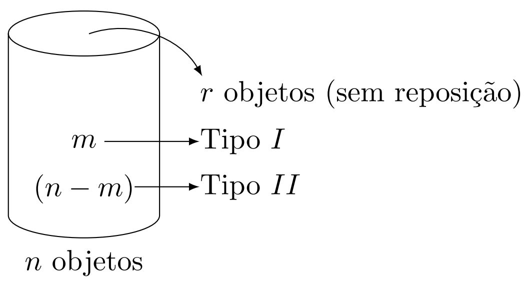
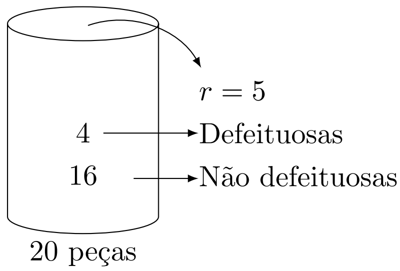


Figura 2.1:

Notacao : $X \sim Hgeo(m, n, r)$

Exemplo : Considere que em um lote de 20 peças, existam 4 defeituosas. Seleciona-se 5 dessas peças, sem resposicao, qual seria a probabilidade de duas defeituosas terem sido escolhidas?



$$X : N^o \text{ de peças defeituosas em 5 retiradas}$$

$$P(x = 2) = \frac{\binom{4}{2}\binom{16}{3}}{\binom{20}{5}}$$

$$R_x : \{0, 1, 2, 3, 4\}$$

2.1.7 Distribuição de Poisson

É largamente empregada quando se deseja contar o número de eventos de certo tipo que ocorrem em um certo período de tempo ou superficie ou volume.

Definicao Uma v.a. tem Distribuição Poisson com parâmetro λ, $\lambda > 0$,se sua f.m.p. é dada por:

$$P_x(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots \tag{2.27}$$

Notacao: $X \sim Poi(\lambda)$

Temos que:

$$E(X) = \lambda \tag{2.28}$$
$$Var(x) = \lambda \tag{2.29}$$

Exemplo : Uma central telefonica recebe, em média, 5 chamadas por minuto. Supondo que a distribuição de Poisson seja adequada nessa situação, obtenha a probabilidade de que a central receba no maximo 2 chamadas durante um intervalo de um minuto.

X : n^ode chamadas recebidas em 1 minuto em uma central telefonica

$E(X) = \lambda = 5$ chamadas/minuto

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$\begin{aligned}
P(X = 0) &= \frac{e^{-5}5^0}{0!} = 0,0067 \\
P(X = 1) &= \frac{e^{-5}5^1}{1!} = 0,0334 \\
P(X = 2) &= \frac{e^{-5}5^2}{2!} = 0,0842
\end{aligned}$$

Logo,

$$P(X \leq 2) = 0,0067 + 0,0334 + 0,0842 = 0,1243$$

O processo de Poisson

Suponha que μ seja a média de ocorrência do evento de interesse em t unidades de medida(por exemplo, tempo). Denotamos por λ, a taxa média de ocorrência em uma unidade de medida, como $\mu = \lambda t$. Podemos reescrever a f.m.p. por:

$$p_x = P(X = x) = \frac{e^{-\lambda t}\,(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Exemplo : Considere o exemplo anterior e calcule a probabilidade de que a central telefonica receba no maximo duas chamadas em 4 minutos.

$$P(X \leq 2t = 4) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$\begin{aligned}
P(X = 0) &= \frac{e^{-5\times 4}(5 \times 4)^0}{0!} = 2,06 \times 10^{-9} \\
P(X = 1) &= \frac{e^{-5\times 4}(5 \times 4)^1}{1!} = 4,12 \times 10^{-8} \\
P(X = 2) &= \frac{e^{-5\times 4}(5 \times 4)^2}{2!} = 4,12 \times 10^{-7}
\end{aligned}$$

$$P(X \leq 2t = 4) = 2,06 \times 10^{-9} + 4,12 \times 10^{-8} + 4,12 \times 10^{-7} = 4,56 \times 10^{-7} \approx 0$$

Resultados Importantes **Resultado 1** : Se $X_1, X_2, \ldots, X_n$ são independetes e $X_i \sim Poisson(\lambda_i)$, então:

$$Y = X_1 + X_2 + \ldots + X_n \sim Poisson\left(\lambda_1 + \ldots + \lambda_n\right) \qquad (2.30)$$

Resultado 2 : Se $X \sim Bin(M, P)$. com m grande e P pequeno, pode-se aproximar a distribuição de X pela distribuição de Poisson, cujo parâmetro sera $\lambda = M \times P$.

Exemplo : número de gols marcados em $M = 100$ tentativas.

$$P = 0,05 \quad \therefore \quad X \sim Bin(100, 0, 05)$$
$$P(X \geq 50) = P(X = 50), \ldots + P(X = 100) = 1 - \{P(X = 0) + \ldots + P(X = 49)\}$$
$$E(X) = M \times p = \mu$$
$$X \sim Poi(\lambda = M \times P)$$

Exercício : Em certa instalacao industrial, acidentes ocorrem com baixa frequência. Sabe-se que a probabilidade de um acidente ocorrer em um certo dia é $0,005$ e que os acidentes são independentes.

(a) Qual a probabilidade de que em um período de 400 dias haja no maximo 3 dias com acidente? Utilize a aproximacao pela distribuição de Poisson.

(b) Calcule a probabilidade exata do item anterior e compare os resultados.

2.2 Modelos Contínuos

2.2.1 Distribuição Uniforme

Definição : Uma v.a. contínua X tem distribuição uniforme no intervalo (a, b), $\quad a, b \in \mathbb{R}$, se sua f.d.p. é dada por:

$$f_x(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases} \qquad (2.31)$$

Notação : $X \sim U(a, b)$

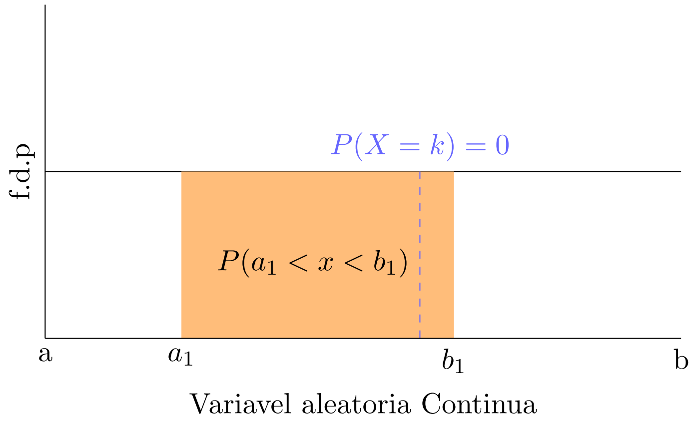


Figura 2.2: ilustração da f.d.p de X

A f.d.a. é dada por:

$$F_x x = \begin{cases} 0, & \text{se} x < a \\ \frac{x-a}{b-a}, & \text{se} a \leq < b \\ 1, x \geq b \end{cases} \tag{2.32}$$

$$\begin{aligned} P(X \leq x) &= \int_{(-\infty}^{x} f_x(t)dt = \int_{a)}^{(x} \frac{1}{b-a)} dt \\ &= \frac{1}{b-a} t|_a^x = \frac{x^a - a}{b-a} \end{aligned} \tag{2.33}$$

Exemplo :

$$\begin{aligned} P(a_1 \leq X \leq b_1) &= F_x(b_1) - F_x(a_1) \\ P(X > a_1) &= 1 - P(X \leq a_1) \\ &= 1 - F_x(a_1) \\ P(X > a_1) &= F_x(a_1) \end{aligned}$$

A média de x é:

$$E(X) = \frac{a+b}{2}$$

E a variância de x é:

$$Var(x) = \frac{(b-a)^2}{12}$$

Exemplo : Seja X uma v.a. com distribuição uniforme, $U(-1/2, 1/2)$. Calcule:

(a) $F_x(x)$

(b) $P(\frac{-1}{4} \leq x \leq \frac{1}{4})$

(c) $E(x)$ e $Var(x)$

Sabe-se que:

$$X \sim U(\frac{-1}{2}; \frac{1}{2})$$

Assim, a f.d.p. de x é:

$$f_x(x) = \frac{1}{\frac{1}{2} - (\frac{-1}{2})} = 1 \quad \frac{-1}{2} < x < \frac{1}{2}$$

(a) $F_x(x) = ?$

$$F_x(x) = P(X \leq x) = \int_{(-\infty)}^{(x)} f(t)dt$$

$$= \int_{(\frac{-1}{2})}^{(x} 1dt = t|_{\frac{-1}{2}} x)$$

$$= x + \frac{1}{2}$$

$$F_x(x) = P(X \leq x) = \begin{cases} 0, & \text{se} x < \frac{-1}{2} \\ x + \frac{1}{2}, & \text{se} \frac{-1}{2} \leq x < \frac{1}{2} \\ 1, & x \geq \frac{1}{2} \end{cases}$$

(b)

$$P(\frac{-1}{4} \leq X \leq \frac{1}{4}) = \int_{(\frac{-1}{4})}^{(\frac{1}{4})}$$

$$= x|_{\frac{-1}{4}}^{\frac{1}{4}} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Ou

$$P(\frac{-1}{4} \leq X \leq \frac{1}{4}) = F_x(\frac{1}{4}) - F_x(\frac{-1}{4})$$
$$\frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

(c)

$$E(X) = \frac{a+b}{2} = \frac{\frac{-1}{2}+\frac{1}{2}}{2} = 0$$

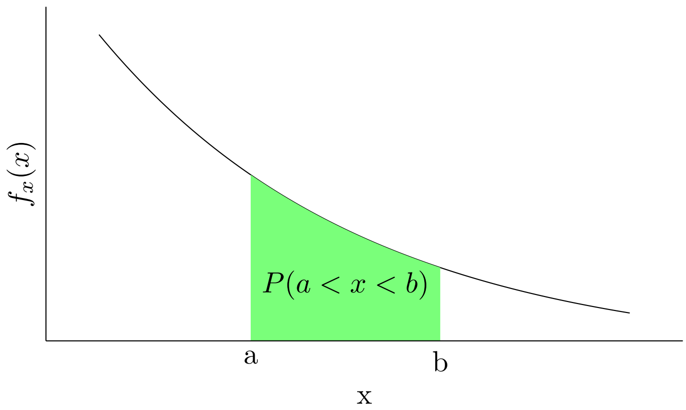
$$Var(x) = \frac{(b-a)^2}{12} = \frac{\left(\frac{1}{2}+\frac{1}{2}\right)^2}{12} = \frac{1}{12}$$

2.2.2 Distribuição Exponencial

Definição : Uma v.a. contínua X tem distribuição exponencial com parâmetro $\lambda, \lambda > 0$, se sua f.d.p. é dada por:

$$f_x(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{caso contrário} \end{cases}$$

Notação : $X \sim Exp(\lambda)$



A f.d.a. de x é:

$$P(X \leq x) = \int_{-\infty}^{x} f_x(t)dt = \int_{0}^{x} \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = -e^{-\lambda x} + 1 = 1 - e^{-\lambda x}$$

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

A média e variância de x são:

$$E(x) = \frac{1}{\lambda}$$

$$Var(x) = \frac{1}{\lambda^2}$$

Propriedade : Se $X \sim Exp(\lambda)$, então:

$$P(X > a + bx > b) = P(X > a)$$

Esta propriedade é conhecida por falta de memoria e é a única distribuição contínua que tem essa propriedade.

Observação :

Outra parametrização para a distribuição exponencial é:

$$f_x(x) = \frac{1}{2}e^{-\frac{x}{2}}, \quad x \geq 0$$

Ou seja, $\lambda = \frac{1}{\alpha}$ A média de x sera:

$$E(x) = \alpha$$

E a variância é dada por:

$$Var(x) = \alpha^2$$

Exemplo: Seja X o tempo de vida util de um fusivel que tem distribuição Exponencial com vida média de 100 horas. Qual a probabilidade de um fusivel durar mais de 150 horas?

$$X \sim Exp(\frac{1}{100})$$

$$\mu = E(x) = \frac{1}{\lambda} = 100 \left(\lambda = \frac{1}{100}\right)$$

$$P(X > 150) = \int_{150}^{\infty} \frac{1}{100}e^{-\frac{1}{100}x}dx = 0,22313$$

Ou

$$P(X > 150) = 1 - P(X \leq 150) =$$

$$1 - \int_{0}^{150} \frac{1}{100}e^{-\frac{1}{100}x}dx = 0,22313$$

2.2.3 Distribuição Normal

E a distribuição mais importante dos modelos probabilisticos. é tambem conhecida como distribuição gaussiana. Sua representação gráfica é conhecida por curva em forma de sino.

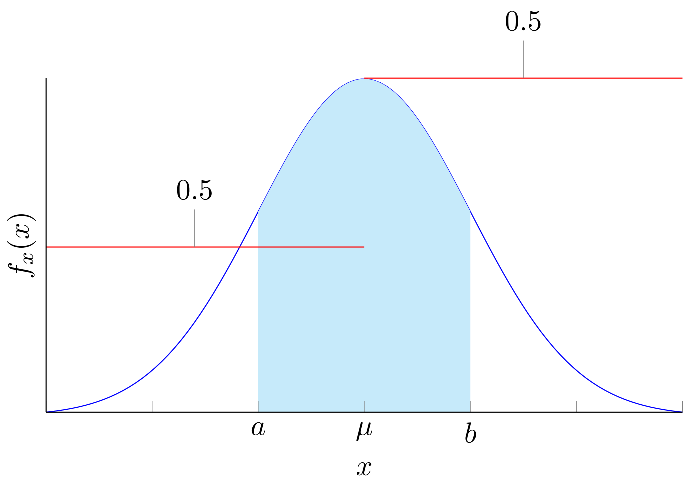
Definição: Uma v.a. contínua X tem Distribuição normal com parâmetro μ (média) e σ^2 se sua f.d.p. é dada por:

$$f_x(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} \quad -\infty < x < \infty \tag{2.34}$$

$$\text{Com} \quad -\infty < \mu < \infty, \quad \sigma^2 > 0 \tag{2.35}$$

Notação: $X \sim N(\mu; \sigma^2)$

Uma ilustração gráfica da sua f.d.p. é:



A média e variância de x são:

$$E(x) = \mu$$
$$Var(x) = \sigma^2$$

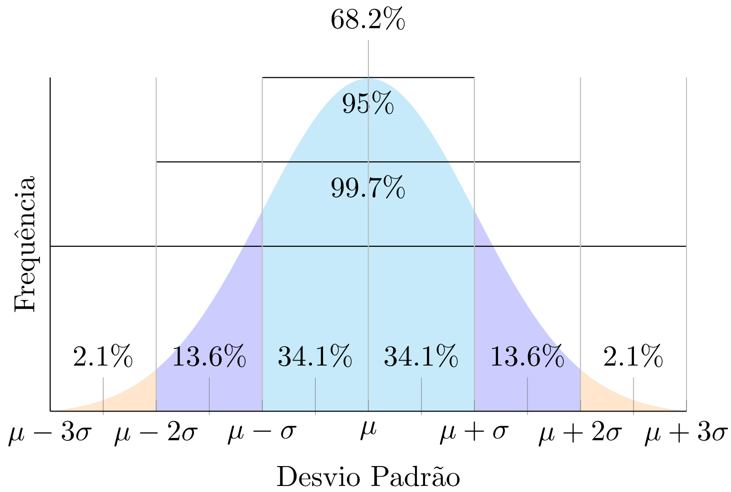
Propriedades: A distribuição é simetrica em relação à média. Isto é:

$$f(\mu - x) = f(\mu + x), \quad \forall x \in \mathbb{R} \tag{2.36}$$

Como a área total sob a curva é igual a 1, à esquerda e à direita de μ, a área é igual a $0,5$.

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0,6896$$
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,9546$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0,9973$$

Ilustrando:



A f.d.a. de uma v.a. $X \sim N(\mu, \sigma^2)$ é:

$$F_x(X) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2}(t-\mu)^2} dt$$

Cuja integral nao tem solução analítica. Assim, calculamos suas probabilidades com auxilio de tabelas.

Definição : Se $X \sim N(\mu; \sigma^2)$, entao a v.a. Z é definida por:

$$Z = \frac{X - \mu}{\sigma} \tag{2.37}$$

Terá a distribuição normal com média 0 e variância 1. Ou seja, $Z \sim N(0, 1)$.

Essa distribuição é conhecida como distribuição normal-padrão ou reduzida.

Cálculo de probabilidades

Exemplo Uso da tabela normal padrão:

Seja $Z \sim N(0, 1)$ calcule:

(a) $P(0 \leq Z \leq 1,65)$

(b) $P(Z \leq 0,5)$

(c) $P(Z < -1,57)$

(d) $P(-0,65 \leq Z \leq 0,5)$

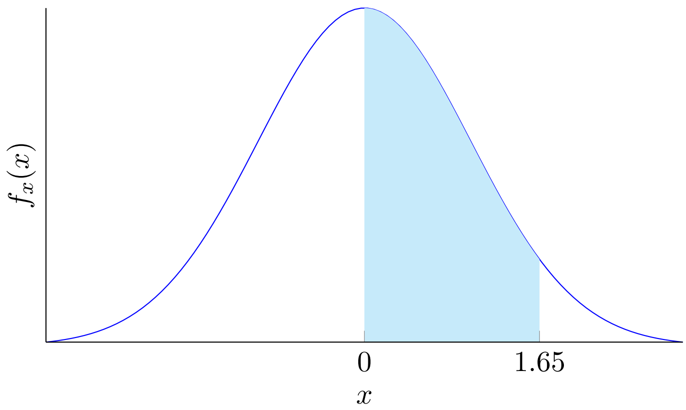
(e) $P(0,8 \leq Z < 1,4)$

(f) $P(0 \leq Z \leq z) = 0,4753$

(g) $P(Z < z) = 0,05$

Resolva:

(a) $P(0 \leq Z \leq 1,65)$

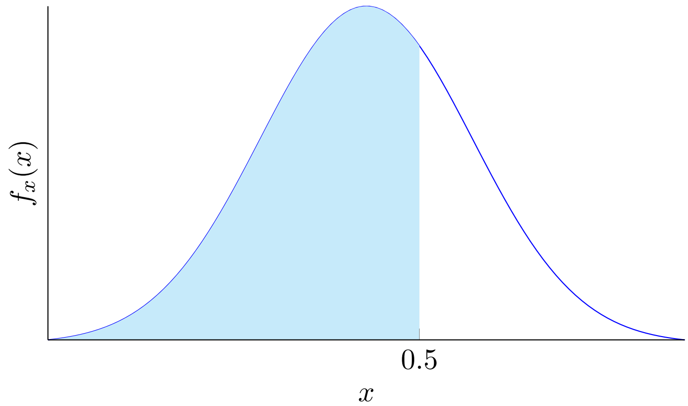


$$P(0 \leq Z \leq 1,65) = 0,45053$$

Ou:

$$\begin{aligned} P(0 \leq Z \leq 1,65) &= P(Z < 1,65) - P(Z \leq 0) \\ &= P(z < 1,65) - 0,5 \\ &= 0,9505 - 0,5 = 0,4505 \end{aligned}$$

(b) $P(Z < 0,5)$

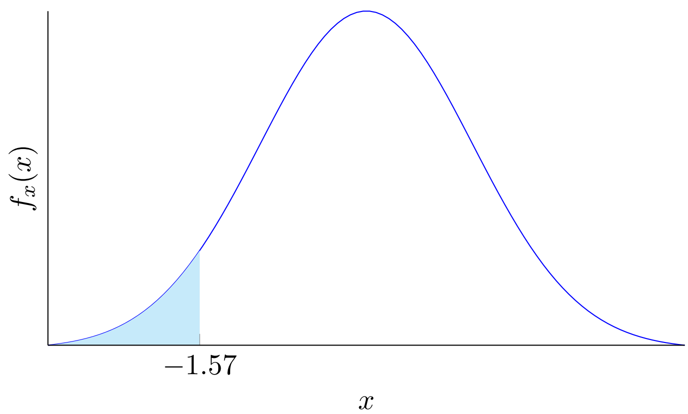


$$
\begin{aligned}
P(Z < 0,5) &= 0,5 + P(0 \leq Z \leq 0,5) \\
&= 0,5 + 0,19146 \\
&= 0,69146
\end{aligned}
$$

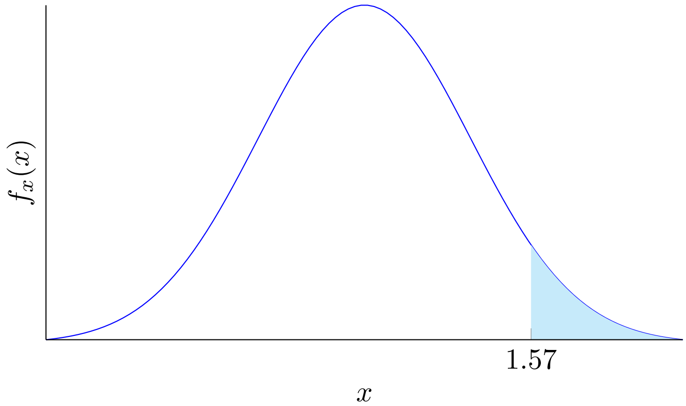
Ou

$$P(Z < 0,5) = 0,6915$$

(c) $P(Z < -1,57)$



Ou equivalentemente:

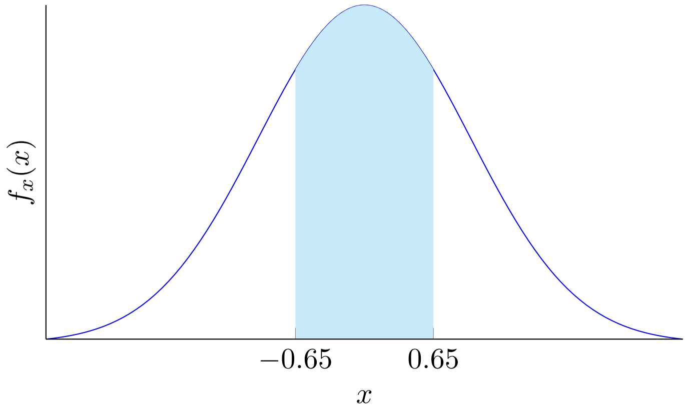


$$
\begin{aligned}
P(Z < -1,57) &= P(Z > 1,57) \\
&= 0,5 - P(0 \leq Z \leq 1,57) \\
&= 0,5 - 0,44179 \\
&= 0,0582
\end{aligned}
$$

Ou

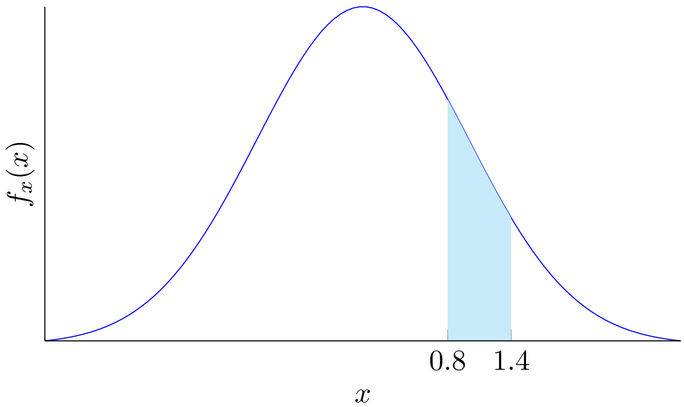
$$P(Z < -1,57) = 0,0582$$

(d) $P(-0,65 \leq Z \leq 0,65)$



$$
\begin{aligned}
P(-0,65 \leq Z \leq 0) &= P(-0,65 \leq Z \leq 0) \\
&= 2P(0 \leq Z \leq 0,65) \\
&= 2 \times 0,24215 \\
&= 0,4843
\end{aligned}
$$

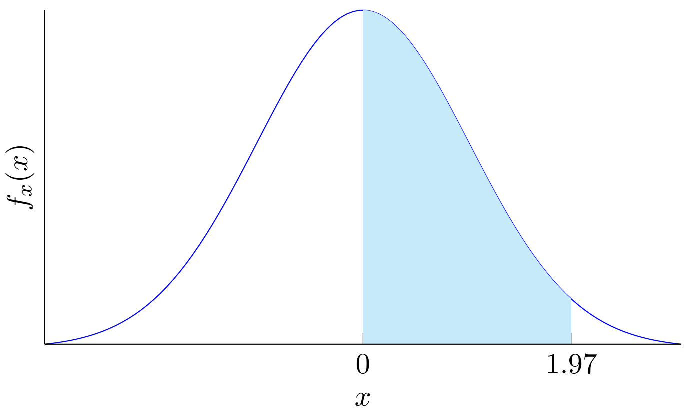
$$P(-0,65 \leq Z \leq -0,65) = P(Z < 0,65) - P(Z < -0,65)$$



(e)

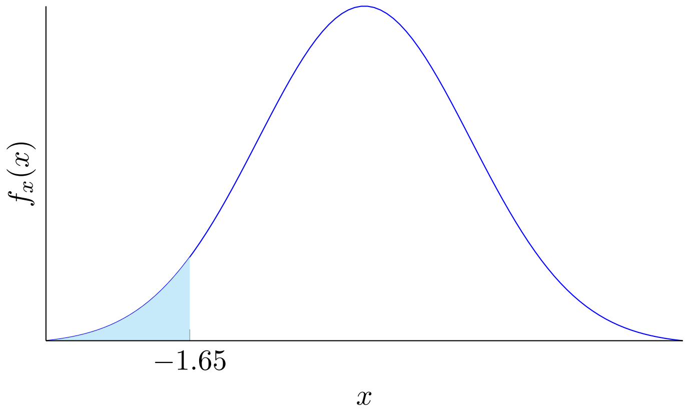
$$\begin{aligned} P(0,8 < z < 1,4) &= P(0 \leq Z \leq 1,4) - P(0 \leq Z \leq 0,8) \\ &= 0,41924 - 0,28814 \end{aligned}$$

(f) $P(0 \leq Z \leq z) = 0,4753$



$$z = 1,965$$

(g) $P(Z < z) = 0,05$



$$z = -1,645$$

Exemplo: Seja $X \sim N(90, 100)$

Determine :

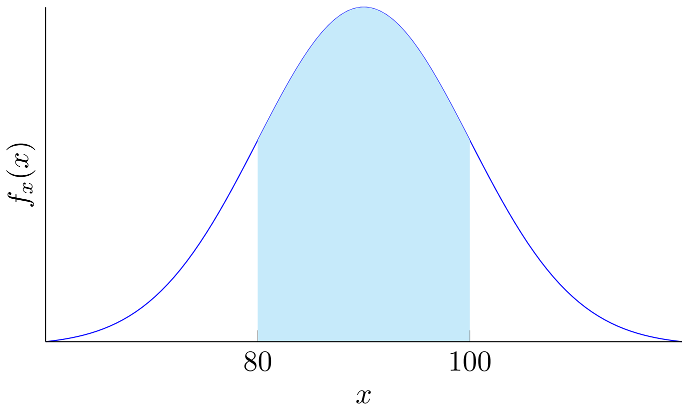
(a) $P(80 < x < 100)$

(b) $P(x \geq 90)$

(c) $P(60 \leq x \leq 75)$

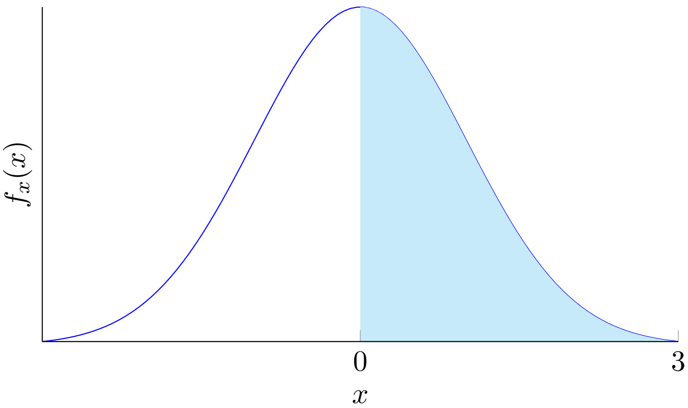
Resolução :

(a) $P(80 < x < 100)$



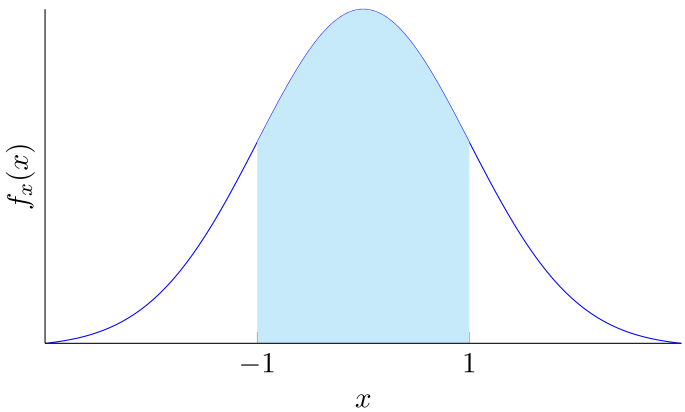
$$Z = \frac{X-\mu}{\sigma}$$
$$z_1 = \frac{80-90}{10} = -1$$
$$z_2 = \frac{100-90}{10} = 1$$

$$P(80 < X < 100) = P(-1 < Z < 1)$$



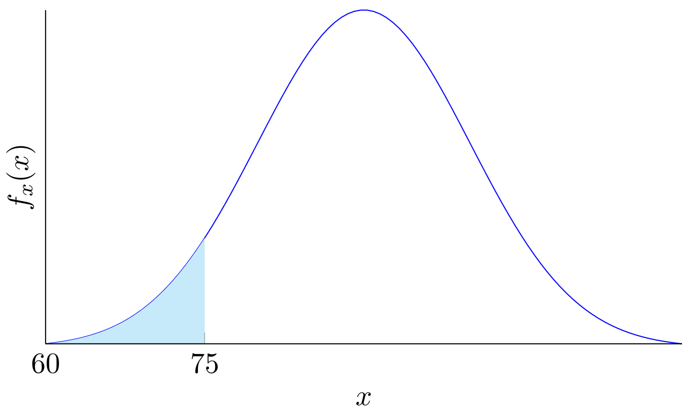
$$\begin{aligned} P(-1 < Z < 1) &= 2 \times P(0 < Z < 1) \\ &= 2 \times 0,34134 \\ &= 0,6826 \end{aligned}$$

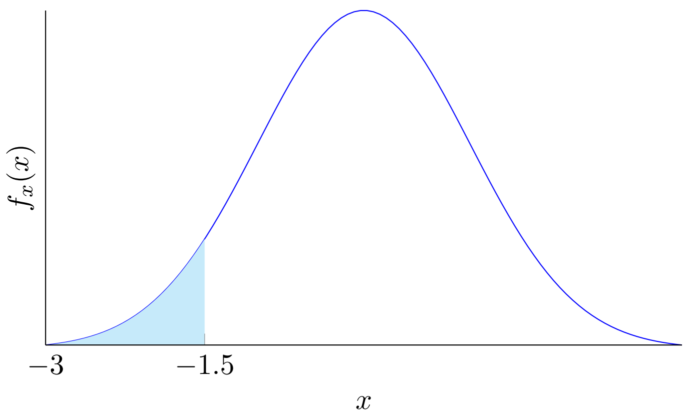
(b) $P(x \geq 90)$



$$
\begin{aligned}
P(x \geq 90) = P(Z \geq \frac{90-90}{10}) \\
= P(Z \geq 0) = 0,5
\end{aligned}
$$

(c) $P(60 \leq x \leq 75)$





$$
\begin{aligned}
P(60 \leq X \leq 75) \\
= P(\frac{60-90}{10} \leq Z \leq \frac{75-90}{10}) \\
= P(-3 \leq Z \leq 1,5)
\end{aligned}
$$

$$
\begin{aligned}
P(-3 < Z < -1,5) \\
= P(Z < 1,5) - P(Z < -3) \\
= 0,0668 - 0,0013 \\
= 0,0655
\end{aligned}
$$

Exemplo: Seja $X \sim \mathtt{N}(50; 10^2)$

Determine: (a) $P(|X-50|<10)$

(b) $P(\mu - a \leq X \leq \mu + a) = 0,9$

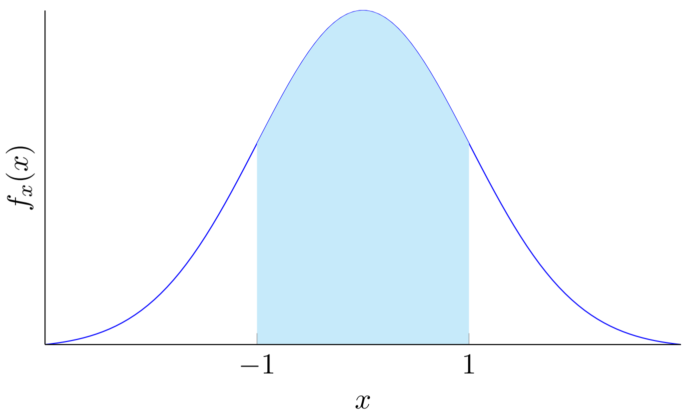
(c) $P(K < X < 70) = 0,8185$

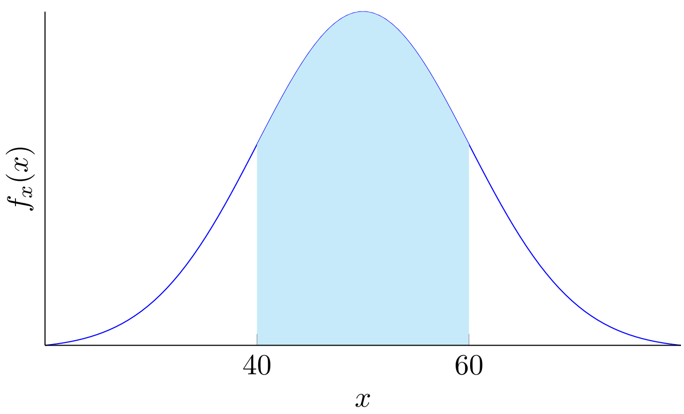
(d) $P(K < x < 75) = 0,3031$

Resolução: (a) $P(|X-50|<10)$

$$\begin{cases} X - 50 < 10 \Rightarrow X < 60 \\ -(X-50) < 10 \Rightarrow X > 40 \end{cases}$$

$$\begin{aligned} P(|X-50|<10) &= P(40 < X < 60) \\ &= P(\frac{40-50}{10} < Z < \frac{60-50}{10}) \\ &= P(-1 < Z < 1) \end{aligned}$$

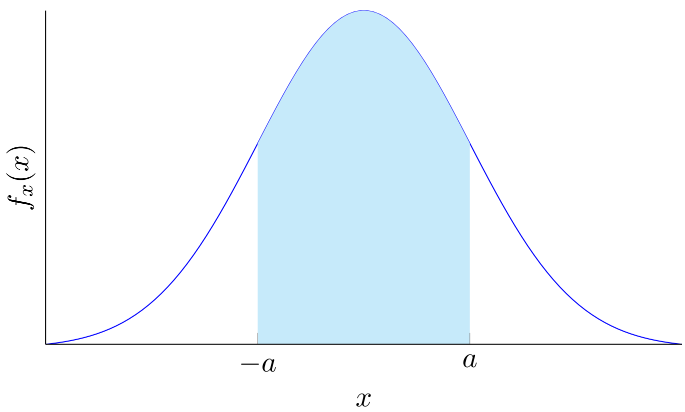




$$\begin{aligned} P(-1 < Z < 1) &= 2 \times P(0 < Z < 1) \\ &= 2 \times 0,34124 \\ &\simeq 0,68 \end{aligned}$$

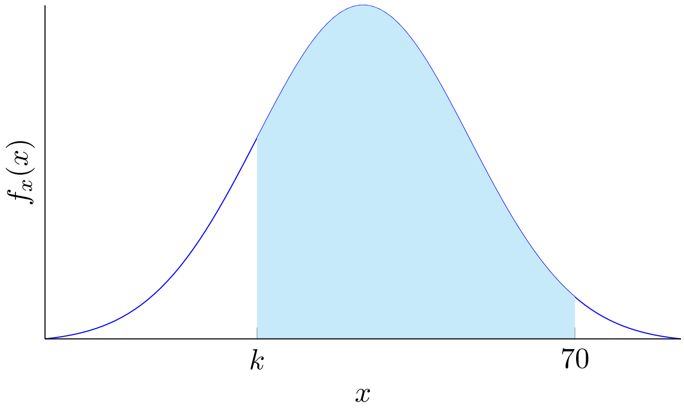
(b) $P\left(\mu - a \leq X \leq \mu + a\right) = 0,9$

$$\begin{aligned} P\left(50 - a \leq X \leq 50 + a\right) &= 0,9 \\ P\left(\frac{(50-a)-50}{10}\right) &\leq Z \leq \frac{(50+a)-50}{10} \\ P\left(\frac{-a}{10} \leq Z \leq \frac{a}{10}\right) &= 0,9 \end{aligned}$$



$$\begin{aligned} P(\frac{-a}{10} \leq Z \leq \frac{a}{10}) = 2 \times P(0 \leq Z \leq \frac{9}{10}) &= 0,9 \\ P(0 \leq Z \leq \frac{a}{10}) &= 0,45 \\ \frac{a}{10} &= 1,645 \\ a &= 16,5 \end{aligned}$$

(c) $P(K < X < 70) = 0,8185$



$$P(\frac{K-50}{10} < Z < \frac{70-50}{10}) = 0,8185$$
$$P(\frac{K-50}{10} < Z < 2) = 0,8185$$

$$P(\frac{K-50}{10} < Z < 2) = 0,8185$$
$$P(\frac{K-50}{10} < Z < 0) + P(0 < Z < 2) = 0,8185$$
$$P(\frac{K-50}{10} < Z < 0) = 0,8185 - 0,47725$$
$$P(\frac{K-50}{10} < Z < 0) = 0,3413$$

$$\frac{-(K-50)}{10} = 10$$
$$K = 40$$

(d) $P(K < x < 75) = 0,3031$
Temos o caso da tabela 1 caso $P(50 < X < 75) > 0,3031$. Do contrário, se $P(50 < X < 75) < 0,3031$, temos o caso do tabela 2.

$$P(50 < X < 75)$$
$$= P(\frac{50-50}{10} < Z < \frac{75-50}{10})$$
$$= P(0 < Z < 2,5) = 0,49379$$

$$P(K < X < 75) = 0,3031$$
$$= P(\frac{K-50}{10} < Z < 2,5) = 0,3031$$

$$P(0 < Z < 2,5) - P(0 < Z < \frac{k-50}{k}) = 0,3031$$
$$0,49379 - 0,3031 = P(0 < Z < \frac{k-50}{10})$$
$$P(0 < Z < \frac{\overbrace{k-50}^{0,495}}{10}) = 0,19069$$
$$\frac{K-50}{10} = 0,495 \Rightarrow K = 4,95 + 50$$
$$K = 54,95$$

Distribuição da Combinacao linear de v.a.s̀ normais independentes

Sejam $X_1, X_2, \ldots, X_n$ e n v.a.s̀ tais que:

$$X_i \sim \mathtt{N}(\mu; \sigma^2) \quad \text{para} \quad i = 1, 2, \ldots, n$$

Considere $Y = \sum_{i=1}^{n} X_i = X_1 + \ldots + X_n$, temos que:

$$E(Y) = E(\sum_{(i)}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$$
$$= \sum_{i=1}^{n} \mu = n \times \mu$$

Ou

$$E(Y) = E(\sum_{i=1}^{n} X_i) = E(X_1 + \ldots + X_n)$$
$$= E(X_1) + E(X_2) + \ldots + E(X_n)$$
$$= \underbrace{\mu + \mu + \cdots + \mu}_{n} = n \times \mu$$

Assim:

$$Var(Y) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$
$$= \sum_{i=1}^{n} \sigma^2 = n \times \sigma^2$$

Logo, $Y \sim \mathtt{N}(n\mu; n\sigma^2)$.
Pelas propriedades da distribuição normal, temos que se $Y \sim \mathtt{N}(n\mu; n\sigma^2)$, a variável reduzida sera:

$$Z = \frac{Y - E(Y)}{\sqrt{Var(Y)}} = \frac{Y - n\mu}{\sqrt{n\sigma^2}} \sim \mathtt{N}(0, 1)$$

Analogamente, temos a média amostral:

$$\bar{X} = \frac{\sum\limits_{(i)}^{n}}{X_i}$$

Com isso, temos:

$$E(\bar{X}) = E(\frac{\sum_{i=1}^{n} X_i}{n}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} \times n \times \mu = \mu$$

E também:

$$Var(\bar{X}) = Var(\frac{\sum_{i=1}^{n}}{n})$$
$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)$$
$$= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} \times n \times \sigma^2$$
$$= \frac{\sigma^2}{n}$$

Logo $\bar{X} \sim \mathtt{N}(\mu; \frac{\sigma^2}{n})$
A variável reduzida é:

$$Z = \frac{\bar{X} - E(\bar{X})}{\sqrt{Var(\bar{X})}} = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$
$$= \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$
$$Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim \mathtt{N}(0, 1)$$

Exemplo: O peso de sacos de parafusos empacotados por uma maquina tem média de $50g$ e desvio-padrão de $2g$. Assumindo que o peso tem distribuição normal, qual a probabilidade de que 10 sacos pesem juntos de $490g$? Qual a probabilidade de que a média desses 10 sacos sejam no maximo $52g$?

(a)

$$X \sim \mathtt{N}(50, 2^2)$$
$$X_1, X_2, \ldots, X_{10}$$
$$Y = \sum_{i=1}^{n} X_i < 490$$
$$Y \sim \mathtt{N}(500; 10 \times 2^2)$$
$$P(Y < 490) = P(Z < \frac{490 - 500}{\sqrt{40}})$$
$$P(Z < -1,58) = 0,0571$$

(b) $X \bar{\leq} 52g$

$$\bar{X} = \frac{\sum X_i}{n} \sim \mathtt{N}(\mu; \frac{\sigma^2}{n})$$
$$\bar{X} \sim N(50; \frac{2^2}{10})$$

$$P(X \leq 52) = P(Z \leq \frac{52 - 50}{\sqrt{\frac{2}{5}}})$$
$$= P(Z \leq 3,16)$$

$$P(Z \leq 3,16) = 0,9992$$

2.3 Variaveis Aleatorias Bidimensionais

Seja X e Y duas variáveis aleatórias quando ha interesse na variacao conjunta de X e Y, estudamos o par (X, Y) como uma variável aleatória bidimensional. '1'

Definição: A variável aleatória bidimensional discreta (X, Y) tem funcao massa de probabilidade conjunta definida por:

$$P_{x,y} = P(X = x, Y = y) \tag{2.38}$$

A funcao de distribuição acumulada $F_{x,y} = P(X \leq x, Y \leq y)$ para (x, y) valores que (X, Y) pode assumir.

Propriedades :

(a) $p_{x,y}(x, y) \geq 0, \quad \forall (x, y).$

(b) $\sum\limits_{R_X} \sum\limits_{R_Y} p_{x,y}(x, y) = 1.$

Exemplo : Em um processo de inspecao de qualdiade, cada peca passa por dois testes e eh classificada de acordo com duas variáveis X e Y, tais que:

$$X = \begin{cases} 1, & \text{Se a peca passa no primeiro teste} \\ 0, & \text{Se a peca falha no primeiro teste} \end{cases}$$

$$Y = \begin{cases} 1, & \text{Se a peca passa no segundo teste} \\ 0, & \text{Se a peca falha no segundo teste} \end{cases}$$

Segundo a f.m.p. conjunta dada por:

$p_{x,y}$	Y	$p_x x$	
X	0	1	
0	0,1	0,2	0,3
1	0,2	0,5	0,7
$p_y(y)$	0,3	0,7	1

$P(X = 0, Y = 0) = 0,1$, indica a proabilidade de uma peca selecionada ao acaso falhar no primeiro teste e no segundo teste.

2.3.1 Distribuiçoes Marginais

Definição: Seja (X, Y) variávei aleatória bidimensional com f.m.p. conjunta $p_{x,y}(x, y)$. A probabilidade marginal de X eh $\sum\limits_{R_Y} p_{x,y}(x, y)$. A probabilidade marginal de Y eh $p_y(y) = \sum\limits_{R_x} p_{x,y}(x, y)$.

Exemplo: Considere o exemplo anterior, e obtenha as distribuiçoes marginais.

Marginal de X:

$$R_x = \{0, 1\}$$

$$p_x(0) = P(X = 0, Y = y)$$
$$= P(X = 0, Y = 0) + P(X = 0, Y = 1) = 0,1 + 0,2 = 0,3$$

$$p_x(0) = P(X = 1, Y = y)$$
$$= P(X = 1, Y = 0) + P(X = 1, Y = 1) = 0,2 + 0,5 = 0,7$$

Com isso, temos que a f.m.p. marginal de X eh:

x	0	1
p_x	0,3	0,7

A f.m.p. marginal de Y eh dada por:

y	0	1
p_x	0,3	0,7

Definicao: A v.a. bidimensional continua (X, Y) tem f.d.p. conjunta dada por $f_{x,y}(x, y)$ e função acumulada $F_{x,y}(x, y)$, satisfazendo:

i.

$$f_{x,y} \geq 0, \quad \forall(x, y) \tag{2.39}$$

ii.

$$\int\limits_{R_y}\int\limits_{R_x} f_{x,y}(x, y)dxdy = int_{R_x}\int\limits_{R_y} f_{x,y}(x, y)dydx = 1$$

iii.

$$\begin{aligned} F_{X,Y}(x, y) &= P(X \leq x, Y \leq y) \\ = F_{X,Y}(x, y) &= P(X < x, Y \leq y) \\ = F_{X,Y}(x, y) &= P(X \leq x, Y < y) \\ = F_{X,Y}(x, y) &= P(X < x, Y < y) \end{aligned}$$

Propriedades :

(a)

$$f_{x,y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{x,y}(x, y)$$

(b)

$$F_{x,y}(x, y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f_{x,y}(x, y)dydx$$

(c)

$$P(a \leq X \leq b; c \leq y \leq d) = \int\limits_{a}^{b}\int\limits_{c}^{d} f_{x,y}(x, y)dydx$$

Exemplo: Duas caracteristicas do desempenho do motor de um foguete sao o empuxo (X) e a taxa de mistura (Y). Suponha que (X, Y) seja uma v.a. continua bidimensional com f.d.p. conjunta dada por:

$$f_{x,y}(x, y) = \begin{cases} 2(x + y - 2xy), & 0 \leq x \leq 1, \quad 0 \leq y \leq 1 \\ 0, & \text{c.c.} \end{cases}$$

Calcular a probabilidade:

$$y \in 0, 1$$

$f_{x,y}(x, y) = 2(x + y - 2xy), \quad x \in 0, 1$
y $\in 0, 1$

A marginal de X:

$$f_x(x) = \int_0^1 2\,(x + y - 2xy)\,dy$$
$$= 2\left(xy|_0^1 + xy|_0^1 - xy^2|_0^1 \right)$$

$$f_x(x) = 2\left(x + \frac{1}{2} - x\right) = 1$$

Logo, a f.d.p de X eh:

$$f_x(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{c.c.} \end{cases}$$

A marginal de Y:

$$f_y(y) = \int_0^1 2\,(x + y - 2xy)\,dx$$
$$= 2\left(x^2|_0^1 - yx^2|_0^1 \right)$$
$$= 2\left(\frac{1}{2} + y - y\right) = 1$$

2.3.2 Independência Probabilística

Duas v.a.'s X e Y sao independentes se a sua f.m.p. conjunta (ou f.d.p. conjunta) de (X, Y) for fatoravel no produto das f.m.p. (ou f.d.p. marginais)de X e Y, isto eh:

$$p_{x,y} = p_x(x) \times p_y(y) \tag{2.40}$$

No caso discreto, e no caso continuo:

$$f_{x,y}(x, y) = f_x(x) \times f_y(y) \quad \forall (x, y) \tag{2.41}$$

Exemplo: Suponha que uma maquina seja ajustada par auma determinada tarefa de manha e para outra tarefa a tarde. SUponha que o numero de vezes que a maquina da problema de manha e a tarde sejam representadas por variaveis X e Y com f.m.p. conjunta dada por:

x / y	0	1	2	$p_x(x)$
0	0,1	0,2	0,2	0,5
1	0,04	0,08	0,08	0,2
2	0,06	0,12	0,12	0,3
$p_Y(y)$	0,2	0,4	0,4	1

Observem que:

$$P(X = 0, Y = 0) = P(x = 0) + P(y = 0) = 0,5 \times 0,2 = 0,1$$

$$P(X = 0, Y = 1) = P(x = 0) + P(y = 1) = 0,5 \times 0,4 = 0,2$$

E assim por diante, para todo x e y.

Como esses produtos de probabilidades sao satisfeitos, temos que X e Y sao independentes.

2.3.3 Esperanca, Covariancia, Correlacao

Definição: Sejam x e y v.a.'s com $E(X) = \mu_x$, $E(Y) = \mu_y$, $Var(X) = \sigma_x^2$ e $Var(Y) = \sigma_y^2$. Seja $p_{x,y}(x, y)$ a f.m.p. conjunta ou $f_{x,y}(x, y)$ a f.d.p. conjunta de x e y. Seja $h(x, y)$ uma função de (X, Y). Definimos o valor esperado de $h(x, y)$ no caso discreto por:

$$E\left(h\left(x, y\right)\right) = \sum limits_{R_x} \sum_{R_y} h(x, y) \times p_{x,y}(x, y) \tag{2.42}$$

E no caso contínuo:

$$E(h(x,y)) = \int_{R_x}\int_{R_y} h(x,y) f_{x,y}(x,y)dydx \tag{2.43}$$

$$= \int_{R_y}\int_{R_x} h(x,y) f_{x,y}(x,y)dxdy \tag{2.44}$$

Alem disso, definimos para ambos os casos, a covariancia entre x e y:

$$\sigma_{x,y} = Cov(X,Y) = E\left((X - E(X))\,(Y - E(Y))\right) \tag{2.45}$$

$$= E(XY) - E(X)E(Y) \tag{2.46}$$

$$= E(XY) - \mu_x \times \mu_y \tag{2.47}$$

A correlacao entre x e y eh dada por:

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sqrt{Var(x) \times \sqrt{Var(y)}}} = \frac{\sigma_{x,y}}{\sigma_x \times \sigma_y} \tag{2.48}$$

Exemplo: Vamos calcular a Correlacao entre X e Y no exemplo anterior (exemplo da maquina).
A f.m.p. conjunta eh:

x / y	0	1	2	$p_x(x)$
0	0,1	0,2	0,2	0,5
1	0,04	0,08	0,08	0,2
2	0,06	0,12	0,12	0,3
$p_Y(y)$	0,2	0,4	0,4	1

$$\rho_{x,y} = \frac{Cov(x,y)}{\sigma_x, \sigma_y}$$

$$Cov(X,Y) = E(XY) - E(X) \times E(Y)$$

$$\mu_x = E(X) = \sum_{R_x} x \times p_x$$

$$= 0 \times 0.5 + 1 \times 0.2 + 2 \times 0.3$$

$$= 0.8$$

$$\begin{aligned}\mu_y = E(Y) = \sum_{R_y} y \times p_y \\ = 0 \times 0.2 + 1 \times 0.4 + 2 \times 0.4 \\ = 1.2\end{aligned}$$

$$\begin{aligned}E(XY) = \sum_{R_X}\sum_{R_y} x \times y \times p_{x,y}(x,y) \\ = 0 \times 0 \times 0.1 + \ldots + 2 \times 2 \times 0.12 \\ = 0.96\end{aligned}$$

$$\begin{aligned}Cov(x,y) = 0.96 - 0.8 \times 1.2 \\ = 0\end{aligned}$$

$$\sigma_x^2 - (E(X))^2$$

Como temos:

$$E(X^2) = \sum_{R_x} x^2 \times p_x$$

Logo a $\rho_{x,y}$ eh:

$$\rho_{x,y} = \frac{0}{x \times \sigma_y} = 0$$

Resultado :

Independencia e correlacao : Se X e Y sao duas v.a.'s independentes, então a correlacao entre X e Y eh o 0 ($\rho_{x,y} = 0$).

Obs: A volta nao vale. Se X e Y sao independetes $E(XY) = E(X) \times E(Y)$.

Resultados :

Para X e Y v.a. 's e $a, b \in \mathbb{R}$, constantes:

1.

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

2.

$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

3. Se X e Y sao independentes, entao:

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

4.

$$Var(aX \pm bY) - a^2Var(X) + b^2Var(Y) \pm 2abCov(X,Y)$$

5. Se X e Y sao independentes, entao:

$$Var(aX \pm bY) = a^2Var(X) + b^2Var(Y)$$

Exemplo: Seja (X,Y) uma v.a. bidimensional cuja f.d.p. conjunta eh dada por:

$$f_{x,y}(x,y) = \begin{cases} 9xy, & 0 < x < y < 1 \\ 0, & \text{c.c.} \end{cases}$$

Determine: (a) As marginais de X e Y

(b) $P(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$

Resolução: (a) $f_x(x) =?$ e $f_y(y) =?$

$$f_x(x) = \int_{R_y} f_{x,y}(x,y)dy$$

$$= \int_x^1 8xydy = \left.\frac{8xy^2}{2}\right|_x^1$$

$$= 4x(1-x^2)$$

Entao, a marginal de x eh:

$$f_x(x) = \begin{cases} 4x(1-x^2), & 0 < x < 1 \\ 0, & \text{c.c.} \end{cases}$$

$$f_y(y) = \int_{R_x} f_{xy}(x,y)$$

$$= \int_0^y 8xydx = \left.\frac{8yx^2}{2}\right|_0^y = 4y^3$$

Logo, a marginal de y eh:

$$f_y(y) = \begin{cases} 4y^3 = & 0 < y < 1 \\ 0, & \text{c.c.} \end{cases}$$

(b)

$$\begin{aligned}
P(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2}) &= \int_0^{\frac{1}{2}} 8xydydx \\
&= \int_0^{\frac{1}{2}} \int_0^{y} 8xydxdy \\
&= \frac{1}{16} = 0.0625
\end{aligned}$$

Capítulo 3

Inferencia Estatistica

3.1 Introducao

Objetivo: Produzir informacoes sobre dada caracteristica da populacao na qual estamos interessados, a partir de informacoes colhidas de uma parte dessa populacao (amostra).

Caracteristica: Variavel Aleatoria

$$X : \Omega \to \mathtt{R}$$
$$X : p_{x,\theta} ou f_x(x, \theta) \quad (3.1)$$

Situacao Pratica: θ eh desconhecido

Interesse: Estimar θ (Metodos de Estimacao Pontual)

Para isso, considere uma amostra aleatoria de X:

$$X = (X_1, X_2, \dots, X_n)$$

Cada X_i eh independente e tem a mesma distribuicao de X. ASsociada a uma a.a. X, temos o vetor de observacoes:

$$x = (x_1, x_2, \dots, x_n)$$

Com base na a.a. podemos obter uma estatistica ou estimador que eh funcao da a.a. :

$$T = T(X_1, X_2, \dots, X_n)$$

As estatisticas mais comuns:

Media amostral :

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Variancia Amostral :

$$s^2 = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{n-1}$$

Menor Valor da amostra :

$$X_{(1)} = min(X_1, X_2, \ldots, X_n)$$

Maior valor da amostra:

$$X_{(n)} = max(X_1, X_2, \ldots, X_n)$$

O problema da inferencia estatistica: faer afirmacoes sobre $\theta(s)$, parametros da populacao, atraves da amostra.

Iremos estudar distribuicoes amostrais de algumas estatisticas.

3.2 Distribuicoes Amostrais

3.2.1 Distribuicao Amostral da Media

Estatistica:

$$\tag{3.2}$$

Obs: Observe que se a populacao for normal, entao tera distribuicao exata normal.

Padronizando, obtemos:

$$Z = \frac{\bar{x} - E(\bar{X})}{\sqrt{Var(\bar{X})}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \tag{3.3}$$

$$= \frac{\sqrt{n}\left(\bar{X} - \mu\right)}{\sigma} \tilde{N}(0,1) \tag{3.4}$$

Exemplo : Uma v.a. X tem distribuicao nromal com media 10 e desvio-padrao 3.

(a) Se $\bar{X}$ for a media da amostra de 25 elementos retirados dessa populacao, calcule $P\left(a < \bar{X} < 1\right)$.

(b) Que tamanho deveria ter a amostra para que $P(a < \bar{X} < 11 = 0,95)$?

(a) $P(a < X \bar{<} 11) =)$?

$n = 25$

$$\mathbf{X} = (X_1, X_2, \ldots, \mathrm{X}_{25}$$
$$\bar{X}\tilde{N}(\mu; \frac{\sigma^2}{n})$$
$$\bar{X}\tilde{N}(10; \frac{3^2}{25})$$

$$P(a < \bar{X} < 11) = P\left(\frac{9-10}{\frac{3}{\sqrt{25}}} < Z < \frac{11-10}{\frac{3}{\sqrt{25}}}\right)$$
$$P(-\frac{5}{3} < Z < \frac{5}{3}) = P\left(-1,67 < Z < 1,67\right)$$
$$= 2P(0 < Z < 1,67)$$
$$= 2 \times 0,45254 = 0,9105$$

(b)

$$\bar{X}\tilde{N}(\mu; \frac{\sigma^2}{n})$$
$$\bar{X}\tilde{N}(10; \frac{3^2}{n})$$

$$P\left(9 < \bar{X} < 11\right) = 0,95$$
$$P\left(\frac{9-19}{\frac{3}{\sqrt{n}}} < Z < \frac{11-10}{\frac{3}{\sqrt{n}}}\right) = 0,95$$
$$P\left(\frac{-\sqrt{n}}{3} < Z < \frac{\sqrt{n}}{3}\right) = 0,95$$
$$P\left(0 < Z < \frac{\sqrt{n}}{3}\right) = \frac{0,95}{2} = 0,475$$

$$z = 1,96$$
$$\frac{\sqrt{n}}{3} = 1,96$$
$$n = (5,88)^2$$
$$n = 34,574$$
$$n = 35$$

3.2.2 Distribuicao Amostral para a Proporcao

Seja X uma v.a. definida da seguinte maneira:

$$X = \begin{cases} 1, & \text{se ocorre sucesso(interesse)} \\ 0, & \text{se ocorre frcasso} \end{cases} \tag{3.5}$$

X distribuicao de Bernoulli. Logo:

$$E(X) = \mu = p$$
$$Var(X) = \sigma^2 = p\,(1-p)$$

Considere uma a.a. $\mathbf{X} = (X_1, \ldots, X_n)$ da v.a. X e o interesse eh a proporcao de indivudos/elementos que possui a caracteristica de interesse (sucesso), $\tilde{p}$. Isto eh:

$$\tilde{p} = \frac{\sum\limits_{i=1}^{n}}{n} = \bar{X} \tag{3.6}$$

Observe que cada X_i,$i = 1, \ldots, n$, tem distribuicao Bernoulli e sao independentes. Pelo TCL, $\tilde{p}$ tera distribuicao aproximadamente normal com media $\mu = p$ e variancia:

$$\frac{\sigma^2}{n} = \frac{p(1-p)}{n}$$
$$\tilde{p}\tilde{N}(p, \frac{p(1-p)}{n})$$

Padronizando, temos:

$$Z = \frac{\tilde{p} - E(\tilde{p})}{\sqrt{var(\tilde{p})}} = \frac{p - \tilde{p}}{\sqrt{\frac{p(1-p)}{n}}} \quad \tilde{N}(0,1)$$

Exemplo :

Um procedimento de controle de qualdiade foi planejado para garantir um maximo de 10% de itens defeituosos na producao. Supondo que a producao esteja sob controle ($p = 0,1$) e que os itens sejam vendidos em caixas com 100 unidades, qual a probabilidade de que uma caixa:

(a) Tenha mais do que 10% de defeituosos.

(b) Nao tenha itens defeituosos

X: numero de pecas defeituosas a cada venda

$X\tilde{B}er(p) \quad p = 0,1$

$n = 100 \quad (X_1, X_2, \ldots, X_{100})$

$E(X) = \mu = p = 0,1$

$Var(X) = \sigma^2 = p\,(1-p) = 0,1 \times 0,9 = 0,09$

(a) 10% de 100 = 10

$$P\left(\sum_{i=1}^{100} X_i > 10\right) = P\left(\tilde{p} > 0,1\right)$$

$$\tilde{p}\tilde{N}(0,1; \frac{0,09}{100})$$

$$P\left(\tilde{p} > 0,1\right) = P\left(Z > \frac{0,1-0,1}{\frac{\sqrt{0,09}}{100}}\right)$$

$$= P(Z > 0) = 0,5$$

(b) $P\left(\sum_{i=1} X_i = 0\right) = P(\tilde{p} = 0) \approx 0$

3.3 Fix me

3.3.1 Eficiencia

Definicao Considere dois estimadores $T_1(x) = \tilde{\theta_1}$ e $T_2(x) = \tilde{\theta_2}$, nao-viciados para um parametro θ. Dizemos que $\tilde{\theta_1} = T_1(x)$ eh mais eficiente que $\tilde{\theta_2} = T_2(x)$ se $Var(T_1(x)) < Var(T_2(x))$.

3.3.2 Consistencia

Seja $T_n(x)_{n>1}$ uma sequencia de estimadores de θ:

$$T_1(x); T_2(x); T_3(x), \ldots$$

Definicao a sequencia $T_{n\,n\geq 1}$ eh definida ser uma sequencia consistente de estimadores de θ se, e somente se, para todo $\epsilon > 0$ satisfaz:

$$\lim_{n\to\infty} P(|T_n - \theta| \geq \epsilon) = 0$$

Obs: $\lim_{n\to\infty} (|T_n - \theta < \epsilon) = 1$ Em geral, para verficiar se um estimador eh consistente, utilizamos a desigualdade de Chebyshev que diz:

$$P(|T_n - \theta| \geq \epsilon) \leq \frac{Var(T_n(x))}{\epsilon^2}$$

Proposicao : Uma sequencia $T_n(x)_{n\geq 1}$ de estimadores de θ eh consistente se:

$$\lim_{n\to} E(T_n(x)) = \theta$$

E

$$\lim_{n\to\infty} Var(T_n(x)) = 0$$

Exemplo Seja $X = (X_1, X_2, \ldots, X_n)$ uma a.a. da v.a. X com distribuicao $N\left(\mu, \sigma^2\right)$. Considere os estimadoresde μ e σ^2, respectivamente, $\tilde{\mu} = \bar{X}$ e $\tilde{\sigma^2} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}$. Calcule os vicios para estes estimadores.

Para μ:

$$\beta_\mu\left(\tilde{\mu}\right) = \beta_\mu\left(\bar{X}\right) = E(\bar{X}) - \mu$$

Obtende o $E(\bar{X})$:

$$E(\bar{X}) = E\left(\frac{\sum\limits_{i=1}^{n} X_i}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$$

Logo,

$$\beta_\mu\left(\bar{X}\right) = \mu - \mu = 0$$

Portanto, $\bar{X}$ eh um estimador nao-viciado (nao-viesado).

Para σ^2

$$\beta_{\sigma^2} = E\left(\tilde{\sigma^2}\right) - \sigma^2$$

Obtendo o $E(\tilde{\sigma^2})$:

$$E\left(\tilde{\sigma^2}\right) = E\left(\frac{\sum\limits_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}\right)$$

$$= \frac{1}{n}E\left(\sum_{i=1}^{n}\left(X_i^2 - 2X_i\bar{X}_1 + \bar{X}^2\right)\right)$$

$$= \frac{1}{n}E\left(\sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} 2X_i\bar{X}_1 + \sum_{i=1}^{n}\bar{X}^2\right)$$

$$= \frac{1}{n}E\left(\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^2\right) - nE\left(\bar{X}^2\right)$$

Sabemos que:

$Var(X_i^2) = E\left(X_i^2\right) - E(X_i)^2 E(X_i^2) = Var\left(X_i\right) + E\left(X_i\right)^2 = \sigma^2 + \mu^2$

E alem disso:

$$\begin{aligned} Var(\bar{X}) &= E\left(\bar{X^2}\left\{E\left(\bar{X}\right)^2\right\}\right) \\ E\left(\bar{X^2}\right) &= Var(\bar{X}) + \left\{E\left(\bar{X}\right)^2\right\} \end{aligned}$$

$$\begin{aligned} Var(\bar{X}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n}\sigma^2 \\ &= \frac{\sigma^2}{n} \end{aligned}$$

$$\begin{aligned} E\left(\bar{X^2}\right) &= Var\left(\bar{X}\right) + \left\{E\left(\bar{X}\right)^2\right\} \\ &= \frac{\sigma^2}{n} + \mu^2 \end{aligned}$$

Com isso, temos que:

$$\begin{aligned} E\left(\sigma^2\right) &= \frac{1}{n}\left\{\sigma^2 + \mu^2\right\} \\ &= \sigma^2 + \mu^2 - \left(\frac{\sigma^2}{n} + \mu^2\right) \\ &= \sigma^2 - \frac{\sigma^2}{n} \\ &= \frac{\sigma^2\left(n-1\right)}{n} \end{aligned}$$

Logo,

$$\beta_{\sigma^2}\left(\bar{\sigma^2}\right) = \frac{-\sigma^2}{n}$$

Portanto, $\bar{\sigma^2}$ eh um estimador viciado. Por outro lado, temos:

$$\lim_{n\to\infty} \beta_{\sigma^2}\left(\sigma^2\right) = \lim_{n\to\infty}\left(\frac{-\sigma^2}{n}\right) = 0$$

E $\tilde{\sigma^2}$ sera assistonticamente nao-viciado.

Observacao O calculo do vicio de um estimador pdoe ser usado para sugerir novos estimadores. Por exemplo, considere $\tilde{\sigma^2}$ o estiamdor para σ^2, cujo vicio eh $\frac{-\sigma^2}{n}$. Um novo estiamdor nao-viciado pode ser obtido por:

$$E\left(\tilde{\sigma^2}\right) = \frac{\sigma^2(n-1)}{n}$$
$$\frac{n}{n-1}E\left(\tilde{\sigma^2}\right) = \sigma^2$$
$$E\left(\frac{n\tilde{\sigma^2}}{n-1}\right) = \sigma^2$$

Logo, $S^2 = \frac{n\tilde{\sigma^2}}{n-1}$ eh um estimador nao-viciado para σ^2, pois $E\left(E(S^2) = \sigma^2\right)$. Assim, o estimador nao-viciado para σ^2 eh:

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

Exemplo: Considere uma amostra de tamanho n da v.a. X com distribuicao $N\left(\mu;\sigma^2\right)$. Considere dois estimadores para μ:

$$\tilde{\mu_1} = X_1$$
$$\tilde{\mu_2} = \bar{X}$$

(a) Qual dos estimadores eh melhor do ponto de vista de eficiencia?

(b) E do ponto de ista de consistencia?

Resposta (a)

$$E\left(\tilde{\mu_1}\right) = E\left(X_1\right) = \mu$$
$$E\left(\mu_2\right) = E\left(\bar{X}\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)$$
$$= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n}\sum_{i=1}^{n} = \mu$$

$$Var\left(\tilde{\mu_1}\right) = Var\left(X_1\right) = \sigma^2$$
$$Var\left(\tilde{\mu_2}\right) = Var\left(\tilde{X}\right) = \frac{\sigma^2}{n}$$

Como $Var(\tilde{\mu_2} < Var(\tilde{\mu_1}))$, temos que $\tilde{\mu_2}$ eh melhor.

(b) Avaliando a consistencia pela proposicao.

Para $\tilde{\mu_1}$

$$\lim_{n\to\infty} E\left(\tilde{\mu_1}\right) = \lim_{n\to\infty} \mu = \mu$$

$$\lim_{n\to\infty} Var\left(\tilde{\mu_1}\right) = \lim_{n\to\infty} \sigma^2 = \sigma^2$$

Logo, $tilde\mu_1$ nao eh consistente.

Para $\tilde{\mu_2}$

$$\lim_{n\to\infty} E\left(\tilde{\mu_2}\right) = \lim_{n\to\infty} \mu = \mu$$

$$\lim_{n\to\infty} Var\left(\tilde{\mu_2}\right) = \lim_{n\to\infty} \sigma^2 = 0$$

Logo, $tilde\mu_2$ eh consistente. Do ponto de vista de consistencia $\tilde{\mu_2}$ eh melhor.

Exemplo :

Considere os estimadores $\tilde{\sigma^2}$ e $\tilde{s^2}$ para σ^2. Compare os seus EQM's e verifique se estes estimadores sao consistentes. Utilize o seguinte resultado:

$$Y = \frac{\sum\limits_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{2} \tilde{X}^2_{(N+1)}$$

$$E(Y) = n - 1 \quad \text{e} \quad Var(y) = 2\left(n-1\right)$$

$$\tilde{\sigma^2} = \frac{\sum\limits_{i=1}^{n} \left(X_i - \bar{X}\right)}{n}$$

$$s^2 = \frac{\sum\limits_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{n-1}$$

$$EQM_{\sigma^2} = (T(X)) = Var\left(T(X)\right) + \left\{\beta_{\sigma^2}(T(x))\right\}^2$$

$$EQM_{\sigma^2}\left(\tilde{\sigma^2}\right) = Var\left(\tilde{\sigma^2}\right) + \left\{\beta_{\sigma^2}(\tilde{\sigma^2})\right\}^2$$

Obtendo a $Var\tilde{sigma}^2$:

$$\bar{\ } \atop n$$

$$\sigma^4 \frac{}{n^2 Var\left(\frac{\sum\limits_{i=1}^{n}}{\sigma^2}\right) = \frac{\sigma^4}{n^2} 2(n-1)}$$

$$\frac{Var(\tilde{\sigma^2}) = Var\left(\sum_{i=1}\left(X_i - \bar{X}\right)^2\right)}{n\frac{\sigma^4}{n^2}Var\left(\frac{\sum\limits_{i=1}^{n}}{\sigma^2}\right) = \frac{\sigma^4}{n^2}2(n-1)}$$

$$\begin{aligned}
EQM_{\sigma^2}(\tilde{\sigma^2}) &= \frac{\sigma^4}{n^2}2(n-1) + \left[-\frac{\sigma^2}{n}\right]^2 \\
&= \frac{\sigma^n}{n^2}2(n-1) + \frac{\sigma^4}{n^2} \\
&= \frac{\sigma^2}{n^2}\left\{2n-2+1\right\} \\
&= \frac{\sigma^4}{n^2}\left\{2n-1\right\}
\end{aligned}$$