Noções de Probabilidade

Francisco Marcelo M. da Rocha e Airlane P. Alencar

22 de Agosto de 2019

Índice

Experimento Aleatório: procedimento que, ao ser repetido sob as mesmas condições, pode fornecer resultados diferentes.

Exemplos:

1. Resultado do lançamento de um dado equilibrado;
2. Sorteado um estudante da escola e perguntar se ele é fumante ou não.
3. Condições climáticas no próximo domingo;
4. Tipo sanguíneo de um habitante escolhido ao acaso;
5. Um lote de ações é comprado por R$ 100,00. Você deseja observar o preço que esse lote de ações pode ser vendido daqui a um ano;
6. Dois motoristas em uma rodovia do estado de São Paulo são selecionados aleatoriamente e verifica-se se estão usando o cinto de segurança.

Espaço Amostral (Ω): : conjunto de todos os resultados possíveis de um experimento aleatório.

1. Lançamento de um dado equilibrado.
   $\Omega = \{1, 2, 3, 4, 5, 6\}$
2. Sorteado um estudante da escola e perguntar se ele é fumante ou não.
   $\Omega = \{\text{Fumante, Não Fumante}\}$
3. Tipo sanguíneo de um habitante de Osasco escolhido ao acaso.
   $\Omega = \{\text{A, B, AB, O}\}$
4. Um lote de ações é comprado por R$ 100,00. A que preço esse lote de ações pode ser vendido em um ano.
   $\Omega = \{x \in \mathbb{R} \mid x \geq 0\}$
5. Se um motorista estiver usando cinto de segurança usaremos a letra C, caso contrário S. $\Omega = \{\text{CC, CS, SC, SS}\}$

Eventos: são todos os subconjuntos do espaço amostral Ω e serão representados pelas letras latinas maiúsculas A, B, C, ...

$\emptyset$ (conjunto vazio): **evento impossível**

Ω: **evento certo**

Exemplo:

Resultado do lançamento de um dado equilibrado.

Espaço Amostral: $\Omega = \{1, 2, 3, 4, 5, 6\}$ **Alguns eventos**

$$
\begin{array}{lcl}
A: \text{sair face par} & \Longrightarrow & A = \{2, 4, 6\} \subset \Omega \\
B: \text{sair face maior do que 2} & \Longrightarrow & B = \{3, 4, 5, 6\} \subset \Omega \\
C: \text{sair face 2} & \Longrightarrow & C = \{2\} \subset \Omega
\end{array}
$$

Operações com eventos

Sejam A e B dois eventos de Ω.

- A união dos eventos A e B, $A \cup B$: representa a ocorrência de pelo menos um dos eventos, A ou B.
- A intersecção dos eventos A e B, $A \cap B$: representa a ocorrência simultânea dos eventos A e B.
- Os eventos A e B são **disjuntos** ou **mutuamente exclusivos** quando não têm elementos em comum, ou seja,

$$A \cap B = \emptyset$$

- A e B são complementares se sua intersecção é vazia e sua união é o espaço amostral, isto é,

$$A \cap B = \emptyset \text{ e } A \cup B = \Omega$$

- O complementar de A é representado por A^c.

Exemplo:

Lançamento de um dado equilibrado.

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Eventos: $A = \{2, 4, 6\}$, $B = \{3, 4, 5, 6\}$ e $C = \{1\}$

- Sair uma face par e maior que 2.

$$A \cap B = \{2, 4, 6\} \cap \{3, 4, 5, 6\} = \{4, 6\}$$

- Sair uma face par ou face 1.

$$A \cup C = \{2, 4, 6\} \cup \{1\} = \{1, 2, 4, 6\}$$

- Não sair face 1.

$$C^c = \{2, 3, 4, 5, 6\}$$

Probabilidade

Probabilidade: é um valor entre 0 e 1 que mede a incerteza associada à ocorrência do evento.

Como atribuir probabilidade aos eventos do espaço amostral?

Temos duas abordagens possíveis:

- **Proporção de ocorrências** de um evento: é o número de vezes que esse evento ocorre dividido pelo total de vezes que o experimento é realizado.
- **Suposições teóricas**: nessa abordagem a atribuição de probabilidade a um evento e feita baseando-se em características teóricas do experimento.

Probabilidade

Atribuição da probabilidade:

Exemplo: Uma empresa de tecnologia tem dois departamentos, o comercial com 12 colaboradores e de pesquisa e desenvolvimento com 15 colaboradores. Um colaborador é selecionado ao acaso para participar de um congresso no exterior. Qual é a probabilidade de se escolher um colaborador do setor comercial?

Probabilidade

O experimento aleatório terá seu **modelo probabilístico** especificado quando estabelecemos: um **espaço amostral**, $\boldsymbol{\Omega}$ e uma medida de probabilidade, $\mathbb{P}(\omega)$, para cada ponto amostral ω.

Propriedades da Probabilidade:

- $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}(\emptyset) = 0$;
- Para qualquer evento $A \subseteq \Omega$, $0 \leq \mathbb{P}(A) \leq 1$;
- Se A e B são eventos mutuamente exclusivos, então

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B);$$

- Se $A_1, A_2, \ldots, A_n, \ldots$ são eventos de Ω 2 a 2 disjuntos ($A_i \cap A_j = \emptyset$ se $i \neq j$), então a probabilidade da união desses eventos é igual a soma das probabilidades desses eventos.

Probabilidade

Exemplo: Lançamento de dado

- Espaço amostral, $\Omega = \{1, 2, \ldots, 6\}$.
- $0 \leq \mathbb{P}(\{i\}) \leq 1$;
- $\mathbb{P}(\Omega) = \mathbb{P}(\{1, 2, \ldots, 6\}) = \sum_{i=1}^{n} \mathbb{P}(\{i\}) = 1$;
- Se os resultados do experimento são igualmente prováveis, então

$$\mathbb{P}(\{\omega_i\}) = \frac{1}{6};$$

- Se os pontos de Ω são igualmente prováveis e $A = \{1, 3, 5\}$, então

$$\mathbb{P}(A) = \mathbb{P}(1) + \mathbb{P}(3) + \mathbb{P}(5) = \frac{3}{6} = \frac{N(A)}{N(\Omega)}; \tag{1}$$

Probabilidade

Exemplo 1.

A tabela a seguir apresenta a distribuição de alunos diplomados em 2002, segundo nível de ensino e tipo de instituição, no município de São Paulo.

Nível	Instituição		Total
	Pública	Privada	
Fundamental	144.548	32.299	176.847
Médio	117.945	29.422	147.367
Superior	5.159	56.124	61.283
Total	267.652	117.845	385.497

Um aluno diplomado em 2002 é selecionado ao acaso. Vamos calcular a probabilidade do aluno selecionado ter se formado em cada nível e, em seguida, a probabilidade do aluno selecionado ter estudado em instituição pública.

Probabilidade

Exemplo 1.

Espaço amostral Ω: conjunto formado pelos 385.497 alunos diplomados em 2002 no município de São Paulo.

Eventos de interesse:

- Fundamental: o aluno se formou no ensino fundamental;
- Médio: o aluno se formou no ensino médio;
- Superior: o aluno se formou no ensino superior;
- Pública: aluno se formou em instituição pública.

Probabilidade

Exemplo 1.

$$\mathbb{P}(\text{Fundamental}) = \frac{N(\text{Fundamental})}{N(\Omega)} = \frac{176.847}{385.497} = 0,4588$$

$$\mathbb{P}(\text{Medio}) = \frac{N(\text{Medio})}{N(\Omega)} = \frac{147.367}{385.497} = 0,3823$$

$$\mathbb{P}(\text{Superior}) = \frac{N(\text{Superior})}{N(\Omega)} = \frac{61.497}{385.497} = 0,1590$$

$$\mathbb{P}(\text{Publico}) = \frac{N(\text{Publico})}{N(\Omega)} = \frac{267.652}{385.497} = 0,6943$$

Qual é a probabilidade do aluno escolhido ter se formado no ensino médio e numa instituição pública?

$$\mathbb{P}(\text{Medio} \cap \text{Publico}) = \frac{N(\text{Medio} \cap \text{Publico})}{N(\Omega)} = \frac{117.945}{385.497} = 0,3060$$

Probabilidade

Regra da Adição

Sejam A e B dois eventos de Ω. Então,

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

Caso particular:

$$\mathbb{P}(\Omega) = 1 = \mathbb{P}(A \cup A^c) \Longrightarrow \mathbb{P}(A) = 1 - \mathbb{P}(A^c)$$

Exemplo 1.

Qual é a probabilidade do aluno ter se formado no ensino médio ou numa instituição pública?

- $\mathrm{Medio} \cup \mathrm{Publico}$: aluno formado no ensino médio ou em inst. pública.
- $\mathbb{P}(\mathrm{Medio} \cup \mathrm{Pub}) = \frac{147.367}{385.497} + \frac{267.652}{385.497} - \frac{117.945}{385.497} = \frac{297.074}{385.497} = 0,7706$

Probabilidade

Probabilidade Condicional

Dados dois eventos A e B, a probabilidade condicional de A dado que ocorreu B é denotada por $\mathbb{P}(A|B)$ e definida por

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \ \mathbb{P}(B) > 0.$$

Da definição de probabilidade condicional obtemos a **regra do produto de probabilidades**

$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}(A \mid B)$$

ou

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B \mid A)$$

Probabilidade

Probabilidade Condicional: Exemplo 1.

Qual é a probabilidade do aluno escolhido ser formado no ensino médio sabendo-se que é de instituição pública?

$$\mathbb{P}(Medio \mid Pub) = \frac{\mathbb{P}(Medio \cap Publico)}{\mathbb{P}(Publico)} = \frac{\frac{117.945}{385.497}}{\frac{267.652}{385.497}} = \frac{117.945}{267.652} = 0,4407$$

Exemplo 2.

Numa creche uma sala tem 5 crianças: 3 meninas e 2 meninos. Duas crianças são sorteadas sucessivamente. Calcule as seguintes probabilidades:

- A segunda criança selecionada é uma menina dado que a primeira é um menino.
- A segunda criança selecionada é uma menina.

Probabilidade

Exemplo 2. Creche: Eventos

- o1: a primeira criança selecionada é um menino;
- o2: a segunda criança selecionada é um menino;
- a1: a primeira criança selecionada é uma menina;
- a2: a segunda criança selecionada é uma menina;

Resolução.

$$\mathbb{P}(a2 \mid o1) = \frac{3}{4}$$

$$\begin{aligned}
\mathbb{P}(a2) &= \mathbb{P}((o1 \cap a2) \cup (a1 \cap a2)) = \mathbb{P}(o1 \cap a2) + \mathbb{P}(a1 \cap a2) \\
&= \mathbb{P}(o1) \times \mathbb{P}(a2 \mid o1) + \mathbb{P}(a1) \times \mathbb{P}(a2 \mid a1) \\
&= \frac{2}{5} \times \frac{3}{4} + \frac{3}{5} \times \frac{2}{4} = \frac{12}{20} = \frac{3}{5}
\end{aligned}$$

Probabilidade

Independência de eventos

Dois eventos A e B são independentes, se a informação da ocorrência ou não de B não altera a probabilidade da ocorrência de A

$$\mathbb{P}(A \mid B) = \mathbb{P}(A), \;\; \mathbb{P}(B) > 0.$$

Se A e B não são independentes, são chamados eventos dependentes.

Forma equivalente:

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(B) \times \mathbb{P}(A \mid B) \\ &= \mathbb{P}(A) \times \mathbb{P}(B).\end{aligned}$$

Probabilidade

Exemplo 3.

Qual a probabilidade de que o aluno sorteado se forme no ensino superior?

Agora, qual a probabilidade de seja do superior, dado que se formou em instituição pública?

Os eventos são independentes?

$$\mathbb{P}(\text{Superior}) = \frac{61283}{385497} = 0,1590 = 15,9\%$$

$$\mathbb{P}(\text{Superior}|\text{Pub}) = \frac{5159}{267652} = 0,0193 = 1,93\%$$

São dependentes.

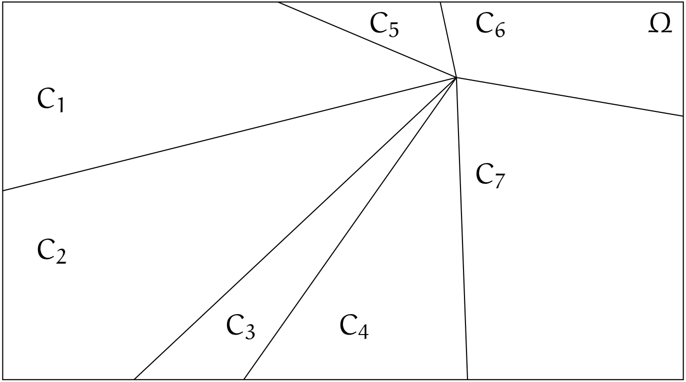
Probabilidade

Partição do espaço amostral.

Os eventos $C_1, C_2, \ldots, C_k$ formam uma partição do espaço amostral, se eles não têm intersecção entre si e se sua união é igual ao espaço amostral. Isto é,

$$C_i \cap C_j = \emptyset \ \text{ para } \ i \neq j \ \text{ e } \ \bigcup_{i=1}^{k} C_i = \Omega$$

Probabilidade



Probabilidade

Teorema de Bayes.

Suponha que os eventos $C_1, C_2, \ldots, C_k$ formem uma partição do espaço amostral e que suas probabilidades sejam conhecidas. Suponha, ainda, que para um evento A, se conheçam as probabilidades $\mathbb{P}(A \mid C_i)$ para todo $i = 1, 2, \ldots, k$. Então, para qualquer j,

$$\mathbb{P}(C_i \mid A) = \frac{\mathbb{P}(A \mid C_j) \times \mathbb{P}(C_j)}{\sum_{i=1}^{k} \mathbb{P}(A \mid C_j) \times \mathbb{P}(C_j)}$$

Probabilidade

Exemplo 2.

Sabendo que a segunda criança selecionada é uma menina. Qual é a probabilidade de que a primeira criança selecionada aleatoriamente seja um menino?

$$\mathbb{P}(o_1 \mid a_2) = \frac{\mathbb{P}(o_1 \cap a_2)}{\mathbb{P}(a_2)} = \frac{\mathbb{P}(o_1) \times \mathbb{P}(a_2 \mid o_1)}{\mathbb{P}(a_2)}$$

$$= \frac{\frac{2}{5} \times \frac{3}{4}}{\frac{12}{20}} = \frac{1}{2}$$

Probabilidade

Sensibilidade e Especificidade

Para um determinado exame, definimos a sensibilidade e a especificidade respectivamente como:

$$
\begin{aligned}
s &= P(\text{Teste positivo}|\text{Doente}) = P(T+|D+) \\
e &= P(\text{Teste negativo}|\text{Não Doente}) = P(T-|D-)
\end{aligned}
$$

Um teste ergométrico foi realizado com o objetivo de detectar doença coronariana (Wiener, 1979 em Soares e Siqueira). O diagnóstico preciso de doença coronariana foi determinado por angioplastia (padrão ouro).

Doença	Teste ergométrico		
coronariana	Positivo (T+)	Negativo (T-)	Total
Presente (D+)	815	208	1023
Ausente (D-)	115	327	442
Total	930	535	1465

Probabilidade Condicional

Sensibilidade e Especificidade

Doença coronariana	Teste ergométrico Positivo (T+)	Negativo (T-)	Total
Presente (D+)	815	208	1023
Ausente (D-)	115	327	442
Total	930	535	1465

$$
\begin{aligned}
s &= \frac{815}{1023} = 0,797 \\
e &= \frac{327}{442} = 0,740
\end{aligned}
$$

Probabilidade Condicional

Metástase de carcinoma hepático

Lind e Singer(1986) (Siqueira e Soares) estudam a qualidade da tomogragia computadorizada para detectar metástase de carcinoma de fígado. O padrão ouro é a laparotomia.

Metástase	Tomografia Computadorizada		
	Positivo (T+)	Negativo (T-)	Total
Presente (D+)	52	15	67
Ausente (D-)	9	74	83
Total	61	89	150

Encontre a sensibilidade e a especificidade.

O que importa para o paciente e para o médico é a probabildade do paciente estar doente dado que o exame deu positivo! Esse valor é denominado o Valor Preditivo Positivo (VPP). Calcule o VPP supondo que a prevalência da metástase de carcinoma de fígado seja igual a 2%.

Probabilidade Condicional - VPP

Metástase de carcinoma hepático

Metástase	Tomografia Computadorizada		
	Positivo (T+)	Negativo (T-)	Total
D+	Verdadeiro Positivo 52	Falso negativo 15	67
D-	Falso positivo 9	Verdadeiro Negativo 74	83
Total	61	89	150

$$
\begin{aligned}
s &= \frac{52}{67} = 0,776 \qquad e = \frac{74}{83} = 0,892 \qquad p = P(D+) = 0,02 \\
VPP &= P(D + |T+) = \frac{P(D + \cap T+)}{P(T+)} \\
&= \frac{P(T + |D+)P(D+)}{P(T + |D+)P(D+) + P(T + |D-)P(D-)} \\
&= \frac{sp}{sp + (1 - e)(1 - p)} = \frac{0,78.0,02}{0,78.0,02 + (1 - 0,89)0,98} = 0,13
\end{aligned}
$$

Probabilidade Condicional - VPN

Metástase de carcinoma hepático

Metástase	Tomografia Computadorizada		
	Positivo (T+)	Negativo (T-)	Total
Presente (D+)	52	15	67
Ausente (D-)	9	74	83
Total	61	89	150

$$
\begin{aligned}
s &= \frac{52}{67} = 0,776 \qquad e = \frac{74}{83} = 0,892 \qquad p = P(D+) = 0,02 \\
VPN &= P(D-|T-) = \frac{P(D-\cap T-)}{P(T-)} \\
&= \frac{P(T-|D-)P(D-)}{P(T-|D+)P(D+) + P(T-|D-)P(D-)} \\
&= \frac{e(1-p)}{(1-s)p + e(1-p)} = \frac{0,89.0,98}{0,89.0,98 + (1-0,78)0,02} = 0,99
\end{aligned}
$$

Probabilidade Condicional - VPN

Metástase de carcinoma hepático

Metástase	Tomografia Computadorizada		
	Positivo (T+)	Negativo (T-)	Total
Presente (D+)	52	15	67
Ausente (D-)	9	74	83
Total	61	89	150

$$s = \frac{52}{67} = 0,776 \qquad e = \frac{74}{83} = 0,892 \qquad p = P(D+) = 0,02$$
$$VPP = P(D+|T+) = 0,13$$
$$VPN = P(D-|T-) = 0,99$$

Se o exame deu positivo, a probabilidade de estar mesmo com metástase é só de 13%!
Probabilidade de Falso positivo em 87% dos casos!
O teste acerta bem se der negativo. Falso negativo só para 1% dos casos.

Probabilidade

Exercícios. - Altman. Drum, Christacapoulos, 1972

2. Um exame de imagem é utilizado para detectar doença do fígado, que é bem menos invasivo que uma biópsia realizada pela Patologia.

Patologia	Exame US fígado		
	Positivo (T+)	Negativo (T-)	Total
Doente (D+)	231	27	258
Normal (D-)	32	54	86
Total	263	81	344

Comente sobre as medidas de concordância entre diagnótico real e do exame (s e e).
Calcule a probabilidade da pessoa estar doente dado que o exame deu positivo se a prevalência for de 25%. E se a prevalência for de 75%.

Probabilidade

Exercícios. - Altman. Weiss et al., 1985

3. Um exame para medir anticorpos de HIV utilizou dados de pacientes com AIDS e doadores sem AIDS. ELISA=Enzyme-linked immunosorbent array.
   Ratio of the mean absorbance of a pair of test samples divided by the mean absorbance of 8 negative control wells.
   Para considerar paciente com resultado positivo, é necessário definir um valor de corte. Considero positivo se a razão for maior que um determinado valor.

Probabilidade

Exercícios. - Altman, D.. Weiss et al., 1985

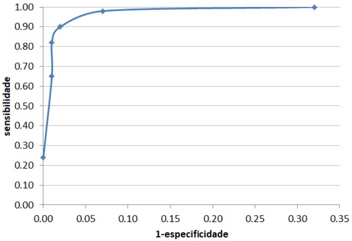
Ratio	Healthy	AIDS
<2.0	202	0
2-2.99	73	2
3-3.99	15	7
4-4.99	3	7
5-5.99	2	15
6-11.99	2	36
>=12	0	21
Total	297	88

- Se considero positivo se a razão for maior que 3, qual é a sensibilidade e a especificidade.
- Calcule a probabilidade da pessoa estar doente dado que o exame deu positivo se a prevalência for de 1%.
- Não faz sentido usar os totais de pacientes sadios e doentes para calcular a prevalência.

Probabilidade

Exercícios. - Altman D. Weiss et al., 1985

Valor de corte	Sensibilidade	Especificidade	PPV	NPV
2	1.00	0.68	0.03	1.000
3	0.98	0.93	0.12	0.9997
4	0.90	0.98	0.28	0.999
5	0.82	0.99	0.39	0.998
6	0.65	0.99	0.49	0.996
12	0.24	1.00	1.00	0.992



Probabilidade

Exercícios.

3. Durante viagens frequentes para certa cidade, um vendedor viajante se hospeda no hotel A 50% das vezes, no hotel B, 30% vezes e no hotel C, 20% das vezes. Quando o vendedor chega ao hotel, há algum problema com a reserva 3% das vezes no hotel A, 6% das vezes no hotel B e 10% das vezes no hotel C. Suponha que o vendedor viaje para essa cidade.
   - a. Ache a probabilidade de que o vendedor se hospede no hotel A e tenha um problema com a reserva.
   - b. Ache a probabilidade de que o vendedor tenha problema com a reserva.
   - c. Suponha que o vendedor tenha problema com a reserva. Qual é a probabilidade de que o vendedor tenha se hospedado no hotel A?

Leitura

Leia o Capítulo .

Capítulo 5 do livro Estatística Básica
Altman!

Variável Aleatória

Variável Aleatória

Como as variáveis de interesse estão sujeitas à aleatoriedade, nós as denominamos variáveis aleatórias.
Para caracterizá-las consideramos sua distribuição de probabilidades.
Para variáveis discretas, associamos uma probabilidade à cada possível valor.

- face do dado: cada face com prob. 1/6;
- X: número de filhos - $P(X=0)=1/4$,$P(X=1)=2/4$, $P(X=2)=1/4$.

Distribuição Bernoulli

Altman, 1991 p. 63

- Simples: só duas possíveis categorias:
  sangue tipo B (prob=0,08 $\rightarrow X = 1$) e não B (prob=0,92 $\rightarrow X = 0$).
- Qual a média ou esperança ou valor esperado dessa variável X?
  Média Ponderada
  $E(X) = 1x0,08 + 0x0,92 = 0,08 = p$
- Qual a variância de X?
  $Var(X) = E(X-p)^2 = E(X^2) - p^2 = p - p^2 = p(1-p)$

Distribuição Binomial

Altman, 1991 p. 63

Amostra aleatória de 2 pessoas → 2 variáveis independentes X_1, X_2.

X1	X2	Prob	S=X1+X2	Prob
0	0	0.8464	0	0.8464
1	0	0.0736	1	0.1472
0	1	0.0736		
1	1	0.0064	2	0.0064
		1		1

Qual a distribuição do número de pessoas com sangue tipo B em amostra aleatória de n pessoas?

$P(Y = X_1 + X_2 + \ldots + X_n = y) = ?, \; y = 0, 1, \ldots, n$

Distribuição Binomial

Altman, 1991 p. 68

- Variáveis independentes $X_1, X_2, \ldots, X_n \rightarrow Y = X_1 + X_2 + \ldots + X_n$

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}, y = 0, 1, \ldots, n$$

- ex: n=10, p=0,08.

$$\begin{aligned}
P(X=0) &= \binom{10}{0} 0,08^0 0,92^1 0 = 0,92^{10} = 0,4344 \\
P(X=1) &= \binom{10}{1} 0,08^1 0,92^9 = 100,080,92^9 = 0,3777 \\
P(X=2) &= \binom{10}{2} 0,08^2 0,92^8 = \frac{10x9}{2} 0,08^2 0,92^8 = 0,1478
\end{aligned}$$

Distribuição Poisson

Altman, 1991 p. 68

- Parece distribuir-se com distribuição Poisson com $\lambda = 0,51$

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

x	Crimes	%	Óbitos	%	P(X=x) Poisson
0	114	61,3%	220	60,3%	0,600
1	56	30,1%	113	31,0%	0,306
2	11	5,9%	23	6,3%	0,078
3	4	2,2%	8	2,2%	0,013
4+	1	0,5%	1	0,3%	0,002
	186		365		

Variáveis Aleatórias Discretas

Número de chegadas durante a madrugada

Parece distribuir-se segundo uma distribuição chamada de Poisson em que

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

sendo a média de X igual a $\lambda = 3$ chegadas por madrugada.

- A probabilidade de não chegar ninguém é $P(X = 0) = e^{-\lambda} = e^{-3} = 0,04979$.
- A probabilidade de chegar alguém é $P(X \geq 1) = 1 - P(X = 0) = 0,95$.
- Probabilidade de chegar mais que 2 pacientes é...

Distribuição Poisson

Número de crimes na lua nova em Nova Delhi e Mortes por dia em hospital em Montreal (Altman, 1991 p. 68)

Parece distribuir-se com distribuição Poisson com $\lambda = 0.51$

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

x	Crimes	%	Óbitos	%	P(X=x) Poisson
0	114	61.3%	220	60.3%	0.600
1	56	30.1%	113	31.0%	0.306
2	11	5.9%	23	6.3%	0.078
3	4	2.2%	8	2.2%	0.013
4+	1	0.5%	1	0.3%	0.002
	186		365		

Variável Aleatória Contínua

Definição (Bussab e Morettin, 2013):

Uma função X, definida sobre o espaço amostral Ω e assumindo valores num intervalo de números reais, é denominada ***Variável Aleatória Contínua***.

Exemplos:

- Altura dos alunos.
- Custo de um sinistro.
- Tempo utilizado na realização de um trabalho.
- Preço de uma ação no fechamento.

Função Densidade de Probabilidade

Definição:

$f(x)$ é uma função densidade de probabilidade (f.d.p.) para uma variável aleatória contínua X, se satisfaz duas condições:

1. $f(x) \geq 0$, para todo $x \in (-\infty, \infty)$;
2. A área definida abaixo de $f(x)$ é igual a 1.

$$\int_{-\infty}^{\infty} f(x) = 1.$$

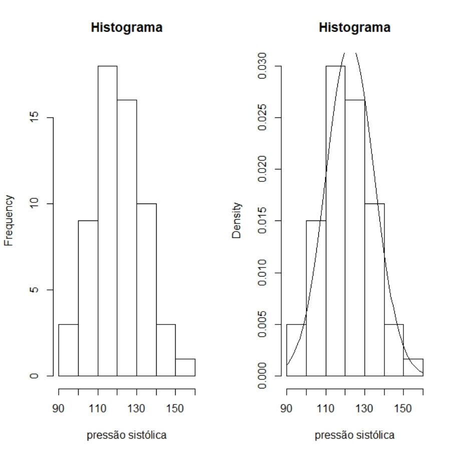
A função $f(x)$ é maior nas regiões em que X tem maior probabilidade de ocorrer.

$P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x)$ corresponde à área sob a curva $f(x)$ entre os pontos a e b.

Distribuição Normal

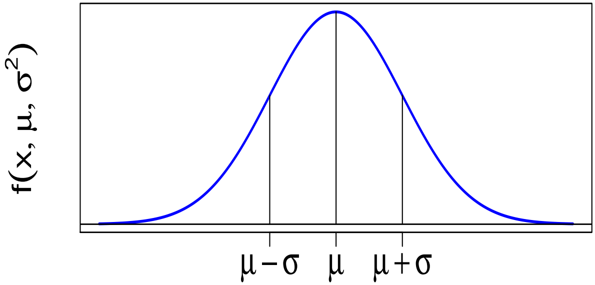
Observa-se a ocorrência de variáveis que apresentam distribuição como a distribuição Normal.

Pressão sistólica de 60 estudantes - Soares e Siqueira



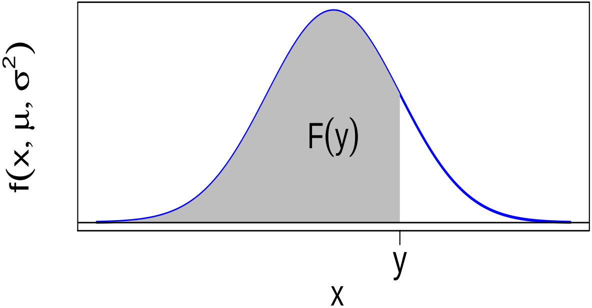
Distribuição Normal

Figura 2: Função densidade de probabilidade de uma variável aleatória normal com média μ e variância σ^2



Distribuição Normal

Figura 3: Representação gráfica da função de distribuição acumulada como área.



Distribuição Normal Padrão

- Se $X \sim \mathcal{N}(\mu; \sigma^2)$, então a variável aleatória definida por

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

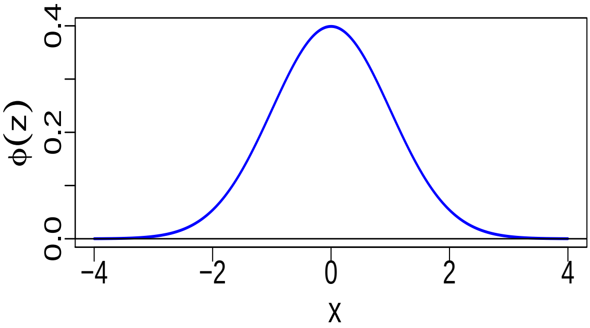
terá média $\mathbb{E}(Z) = 0$ e variância $\mathbb{Var}(Z) = 1$.

- $Z \sim \mathcal{N}(0, 1)$.
- Uma v.a. normal com média zero e variância um é chamada de **Normal Padrão**.
- A função densidade da v.a. normal padrão Z é dada por

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}, \quad \text{para} \quad -\infty < z < \infty.$$

Distribuição Normal Padrão

Figura 4: Representação gráfica da função densidade da normal padrão.



Função de Distribuição Acumulada da Normal Padrão

A função de distribuição acumulada da normal padrão é dada por

$$\Phi(t) = \mathbb{P}(Z \leq t) = \int_{-\infty}^{t} \phi(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-z^2}{2}} dz.$$

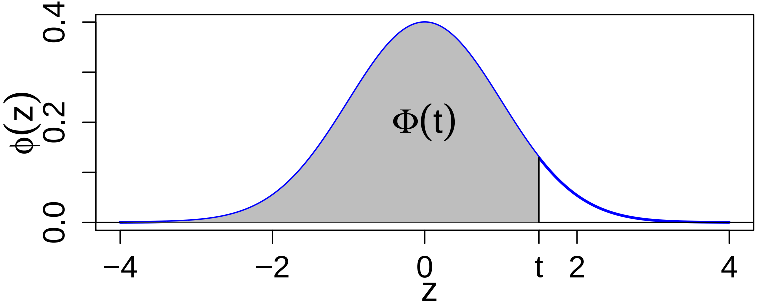


Tabela da Distribuição Normal

Se $Z \sim \mathcal{N}(0,1)$, então a probabilidade

$$\mathbb{P}(-1,3 \leq Z \leq 1,5) = \frac{1}{\sqrt{2\pi}} \int_{-1,3}^{1,5} e^{\frac{-z^2}{2}}\, dz \tag{2}$$

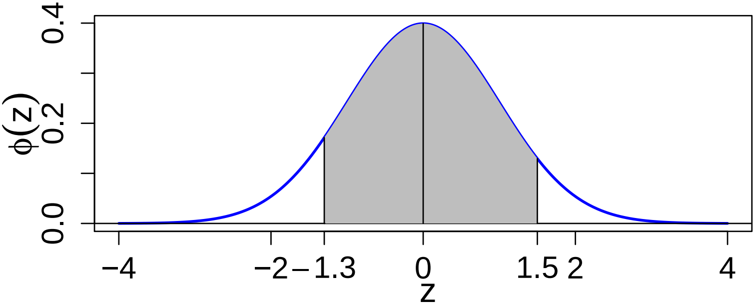


Tabela da Distribuição Normal

- A integral apresentada em (2) não tem expressão analítica.
- A integral (2) pode ser calculada por meio de métodos numéricos para integrais.
- Usaremos a função de distribuição acumulada $\mathbb{P}(Z \leq t) = \Phi(t)$ que foi tabelada. O cálculo segue da seguinte maneira:

$$\begin{aligned}\mathbb{P}(-1,3 \leq Z \leq 1,5) &= \Phi(1,5) - \Phi(-1,3) \\ &= \mathbb{P}(Z \leq 1,5) - \mathbb{P}(Z \leq -1,3) \\ &= 0,9332 - 0,0968 = 0,8364.\end{aligned}$$

Exemplo: $X \sim \mathcal{N}(\mu = 10, \sigma^2 = 4)$

- Usando a tabela

$$
\begin{aligned}
\mathbb{P}(8 \leq X \leq 12) &= \mathbb{P}\left(\frac{8-10}{2} \leq \frac{X-\mu}{\sigma} \leq \frac{12-10}{2}\right) \\
&= \mathbb{P}\left(\frac{8-10}{2} \leq Z \leq \frac{12-10}{2}\right) \\
&= \mathbb{P}(Z \leq 1) - \mathbb{P}(Z \leq -1) \\
&= \Phi(1) - \Phi(-1) \\
&= 0,8413 - 0,1587 = 0,6827
\end{aligned}
$$

Leitura

Leitura.

Capítulo 7 do livro texto

Distribuição Normal

Exercício - Soares e Siqueira - $X \sim \mathcal{N}(\mu = 120, \sigma^2 = 10^2)$

Considere que a pressão sistólica de pessoas saudáveis tem distribuição normal com média 120 mmHg e desvio padrão 10mmHg .

- Qual é a probabilidade de uma pessoa ter pressão sistólica maior que 140mmHg?

$$
\begin{aligned}
\mathbb{P}(X > 140) = 1 - \mathbb{P}(X \leq 140) = 1 - \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{140-120}{10}\right) \\
= 1 - \mathbb{P}\left(Z \leq 2\right) \\
= 1 - \Phi(2) \\
= 1 - 0,9772 = 0,0228
\end{aligned}
$$

- Quais são os limites de um intervalo simétrico em relação à média que engloba 95% dos valores de pressões sistólicos de pessoas saudáveis?

Distribuição Normal

Exercício - Soares e Siqueira - $X \sim \mathcal{N}(\mu = 120, \sigma^2 = 10^2)$

$(X \sim \mathcal{N}(\mu = 120, \sigma^2 = 10^2))$.

- Queremos encontrar a e b em torno de $\mu = 120$ tais que $P(a \leq X \leq b) = 0,95$.

$$
\begin{aligned}
\mathbb{P}(a \leq X \leq b) = &\quad \mathbb{P}\left(\frac{a-120}{10} \leq \frac{X-120}{10} \leq \frac{b-120}{10}\right) = 0,95 \\
= &\quad P(a' \leq Z \leq b') = 0,95
\end{aligned}
$$

Olhando na tabela da distribuição normal padrão, encontramos $a' = -1,96$ e $b' = +1,96$. Então,

$$
\begin{aligned}
a &= 120 - 1,96 * 10 = 100,4 \\
b &= 120 + 1,96 * 10 = 139,6
\end{aligned}
$$

Distribuição Normal

Exercício

Para variável com distribuição normal, $X \sim \mathcal{N}(\mu, \sigma^2)$, temos que a média amostral para amostra aleatória de n indivíduos é

$$\bar{X} = \frac{X_1 + \ldots + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Para amostra de 25 indivíduos, qual a probabilidade $\mathbb{P}(\bar{X} > 123)$.

Leitura

Leitura.

Ler também outras distribuições de probabilidade como a dist. Binomial que discutiremos em aula.

Capítulo 7 do livro Estatística Básica

Referências Bibliográficas

- Bussab, W.O. e Morettin, P.A. (2013). **Estatística Básica**. 8ª edição. São Paulo: Editora Saraiva.
- Magalhães, M.N. e Lima, A.C.P. (2013). **Noções de Probabilidade e Estatística**. 7ª edição. São Paulo: EDUSP.