Sumário

Variáveis aleatórias

Em probabilidade, uma função X que associa a cada evento do espaço amostral um número real $X(\omega) \in \mathbb{R}$, é denominada uma **variável aleatória** (VA).

Uma variável aleatória pode ser classificada como **discreta** ou **contínua**, dependendo do domínio dos valores de X.

Exemplo: o número de alunos em uma sala é uma variável aleatória (discreta), denotada por X (maiúsculo). Uma observação dessa variável é denotada pela respectiva letra minúscula, *e.g.*, $x = 50$ alunos.

Em geral, denotamos a probabilidade de uma V.A. X assumir determinado valor x como

$$P[X] \quad \text{ou} \quad P[X = x]$$

Variáveis aleatórias

Dada a realização de um experimento aleatório qualquer, com um certo espaço de probabilidade, desejamos estudar a **estrutura probabilística** de quantidades associadas à esse experimento.

Note que antes da realização de um experimento, **não sabemos seu resultado**, entretanto seu espaço de probabilidade pode ser previamente estabelecido.

Dessa forma, podemos atribuir probabilidades aos *eventos* desse espaço amostral, dando origem ao conceito de **variável aleatória**.

Distribuições de probabilidade

Existem diversos *modelos probabilísticos* que procuram descrever vários tipos de variáveis aleatórias: são as **distribuições de probabilidade de variáveis aleatórias** (discretas ou contínuas).

A distribuição de probabilidades de uma VA X é, portanto, uma descrição das probabilidades associadas com os possíveis valores de X. Os valores que X assume determinam o **suporte** (S) da VA.

- **Variáveis discretas** $\rightarrow$ suporte em um conjunto de valores enumeráveis (finitos ou infinitos)
- **Variáveis contínuas** $\rightarrow$ suporte em um conjunto não enumerável de valores

Distribuições de probabilidade

Denomina-se de **distribuição de probabilidade** de alguma variável aleatória, a **regra** geral que define a

- **função de probabilidade** (fp) (V.A.s discretas), ou a
- **função densidade de probabilidade** (fdp) (V.A.s contínuas)

para a variável de interesse.

Existem muitas distribuições de probabilidade, mas algumas merecem destaque por sua importância prática.

Estas distribuições também são chamadas de **modelos probabilísticos**.

Sumário

Definição

A **função de probabilidade** (fp) da VA discreta X, que assume os valores $x_1, x_2, \ldots, x_n, \ldots$, é a função que atribui probabilidades a cada um dos possíveis valores: $\{[x_i, p(x_i)], i = 1, 2, ...\}$, ou seja,

$$P[X = x_i] = p(x_i) = p_i, \quad i = 1, 2, \ldots$$

com as seguintes propriedades:

1. A probabilidade de cada valor deve estar entre 0 e 1

$$0 \leq p(x_i) \leq 1, \quad \forall\, i = 1, 2, ...$$

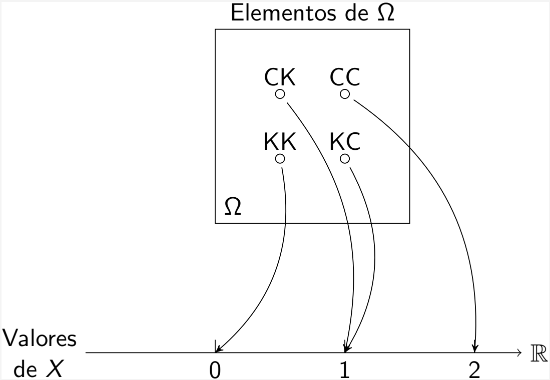
2. A soma de todas as probabilidades é igual a 1

$$\sum_i p(x_i) = 1$$

Exemplo 3.3

Experimento

Lançamento de duas moedas. X = número de resultados cara (C)



Exemplo 3.3

Podemos montar uma tabela de casos favoráveis para a cada possível valor da variável aleatória X = número de resultados cara (C)

X	Casos favoráveis	Casos favoráveis (relativo)
0	1	1/4
1	2	2/4
2	1	1/4
Total	4	1

Assim podemos associar a cada valor de X sua **probabilidade** correspondente, como resultado do número relativo de casos favoráveis

$$P[X = 0] = 1/4$$
$$P[X = 1] = 2/4 = 1/2$$
$$P[X = 2] = 1/4$$

Exemplo 3.3

Dessa forma, a **distribuição de probabilidade** da variável aleatória $X =$ número de resultados cara (C) é a tabela

X	$P[X = x_i] = p(x_i)$
0	1/4
1	1/2
2	1/4
Total	1

Repare que as propriedades da função de probabilidade estão satisfeitas:

1. As probabilidades $p(x_i)$ estão entre 0 e 1
2. A soma de todas as probabilidades $p(x_i)$ é 1

Exemplo 3.1

Com dados do último censo, a assistente social de um Centro de Saúde constatou que para as famílias da região, 20% não tem filhos, 30% tem um filho, 35% tem dois, e as restantes se dividem igualmente entre três, quatro ou cinco filhos.

Descreva a função de probabilidade da VA N definida como número de filhos.

Exemplo 3.2

Na construção de um certo prédio, as fundações devem atingir 15 metros de profundidade e, para cada 5 metros de estacas colocadas, o operador anota se houve alteração no ritmo de perfuração previamente estabelecido. Essa alteração é resultado de mudanças para mais ou para menos, na resistência do subsolo.

Nos dois casos, medidas corretivas serão necessárias, encarecendo o custo da obra. Com base em avaliações geológicas, admite-se que a probabilidade de ocorrência de alterações é de 0.1 para cada 5 metros.

O custo básico inicial é de 100 UPC (unidade padrão de construção) e será acrescida de $50k$, com k representando o número de alterações observadas. Como se comporta a va custo das obras de fundações?

Função de distribuição de probabilidade

Em muitas situações, é útil calcularmos a probabilidade **acumulada** até um certo valor.

Definimos a **função de distribuição** ou **função acumulada de probablidade** de uma VA X pela expressão:

$$F(x) = P(X \leq x)$$

para qualquer número real x.

Exemplo 3.5

Uma população de 1000 crianças foi analisada num estudo para determinar a efetividade de uma vacina contra um tipo de alergia.

No estudo, as crianças recebiam uma dose da vacina e, após um mês, passavam por um novo teste. Caso ainda tivessem tido alguma reação alérgica, recebiam outra dose da vacina. Ao fim de 5 doses todas as crianças foram consideradas imunizadas.

Os resultados completados estão na tabela a seguir.

	1	2	3	4	5
Freq.	245	288	256	145	66

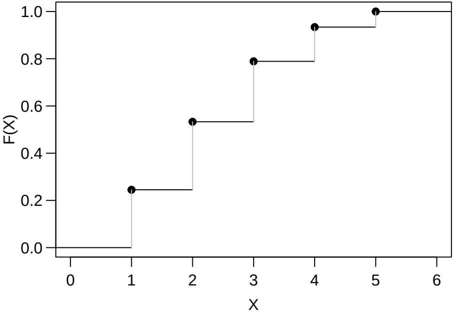
Para uma criança sorteado ao acaso qual a probabilidade dela ter recebido 2 doses? E até 2 doses?

Exemplo 3.5

Tabela de frequência:

	n_i	f_i	f_{ac}
1	245	0.245	0.245
2	288	0.288	0.533
3	256	0.256	0.789
4	145	0.145	0.934
5	66	0.066	1.000
Sum	1000	1.000	

Grafico de $F(X)$:



Exemplo 3.5

Assim,

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = 0,533$$

Note que podemos escrever

$$F(x) = P(X \leq x) = 0,533 \quad \text{para} \quad 2 \leq x < 3$$

E os valores completos da função de distribuição são:

$$F(x) = \begin{cases} 0 & \text{se } x < 1 \\ 0,245 & \text{se } 1 \leq x < 2 \\ 0,533 & \text{se } 2 \leq x < 3 \\ 0,789 & \text{se } 3 \leq x < 4 \\ 0,934 & \text{se } 4 \leq x < 5 \\ 1 & \text{se } x \geq 5 \end{cases}$$

Exemplo 3.6

Num estudo sobre a incidência de câncer foi registrado, para cada paciente com esse diagnóstico, o número de casos de câncer em parentes próximos (pais, irmãos, tios, filhos, primos e sobrinhos). A distribuição de frequência para 26 pacientes é a seguinte:

	0	1	2	3	4	5
n_i	4	4	6	6	2	4

Estudos anteriores assumem que a incidência de câncer em parentes próximos pode ser teoricamente modelada pela seguinte função discreta de probabilidade:

	0	1	2	3	4	5
p_i	0.1	0.1	0.3	0.3	0.1	0.1

Os dados observados concordam com o modelo teórico?

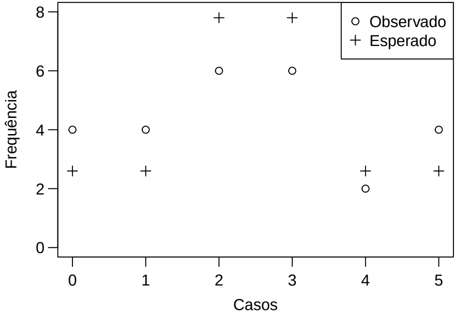
Exemplo 3.6

O número de observações de incidência que seria esperado seguindo o modelo teórico é calculado como

$$e_i = n \times p_i$$

.

	n_i	e_i
0	4	2.6
1	4	2.6
2	6	7.8
3	6	7.8
4	2	2.6
5	4	2.6
Sum	26	26.0



Exercícios recomendados

- Seção 3.1 - $3, 5$ e 6.
- Seção 3.2 - $2, 3, 5$ e 6.
- Seção 3.3 - $1, 2, 3, 4, 5$ e 6.
- Extras: Seção 3.4 - 1, 9, 10, 14, 16, 19, 21, e 26.

Sumário

Variáveis aleatórias

Em probabilidade, uma função X que associa a cada evento do espaço amostral um número real $X(\omega) \in \mathbb{R}$, é denominada uma **variável aleatória** (VA).

Uma variável aleatória pode ser classificada como **discreta** ou **contínua**, dependendo do domínio dos valores de X.

Exemplo: o número de alunos em uma sala é uma variável aleatória (discreta), denotada por X (maiúsculo). Uma observação dessa variável é denotada pela respectiva letra minúscula, *e.g.*, $x = 50$ alunos.

Em geral, denotamos a probabilidade de uma V.A. X assumir determinado valor x como

$$P[X] \quad \text{ou} \quad P[X = x]$$

Distribuições de probabilidade

Existem diversos *modelos probabilísticos* que procuram descrever vários tipos de variáveis aleatórias: são as **distribuições de probabilidade de variáveis aleatórias** (discretas ou contínuas).

A distribuição de probabilidades de uma VA X é, portanto, uma descrição das probabilidades associadas com os possíveis valores de X. Os valores que X assume determinam o **suporte** (S) da VA.

- **Variáveis discretas** $\rightarrow$ suporte em um conjunto de valores enumeráveis (finitos ou infinitos)
- **Variáveis contínuas** $\rightarrow$ suporte em um conjunto não enumerável de valores

Distribuições de probabilidade

Denomina-se de **distribuição de probabilidade** de alguma variável aleatória, a **regra** geral que define a

- **função de probabilidade** (fp) (V.A.s discretas), ou a
- **função densidade de probabilidade** (fdp) (V.A.s contínuas)

para a variável de interesse.

Existem muitas distribuições de probabilidade, mas algumas merecem destaque por sua importância prática.

Estas distribuições também são chamadas de **modelos probabilísticos**.

Variáveis aleatórias contínuas

Uma V.A. é classificada como contínua se assume valores em qualquer intervalo dos números reais, ou seja, um conjunto de valores não enumerável. Dessa forma, não é possível atribuir probabilidades para um ponto específico, apenas para intervalos da reta.

Exemplos:

- Peso de animais
- Tempo de falha de um equipamento eletrônico
- Altura da maré em uma hora específica
- Salinidade da água do mar
- Retorno financeiro de um investimento

Exemplo 6.1

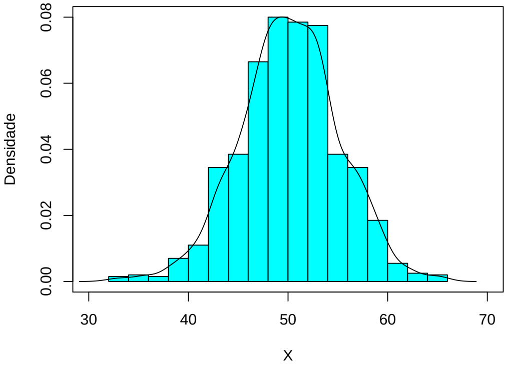
Estudos anteriores revelam a existência de uma grande lençol de água no subsolo de uma região. No entanto, sua profundidade ainda não foi determinada, sabendo-se apenas que o lençol pode estar situado em qualquer ponto entre 20 e 100 metros.

- Determine uma função para representar a variável X (profundidade do lençol de água).
- Calcule a probabilidade de encontar água em uma profundidade pelo menos igual a 25, mas inferior a 29 metros.

Função densidade de probabilidade

Não podemos atribuir probabilidades à valores específicos, pois há uma quantidade **não enumerável** (infinita) de valores em um ponto.

Atribuimos probabilidades à intervalos de valores, por meio de uma **função**. Portanto, as probabilidades são representadas por áreas.



Função densidade de probabilidade

A **função densidade de probabilidade** (fdp) atribui probabilidades à intervalos de valores do tipo $[a, b]$, e é definida por

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

com as seguintes propriedades:

i. É uma função não negativa

$$f(x) \geq 0$$

ii. A área total sob a curva deve ser igual a 1

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

Função densidade de probabilidade

Observações:

- $P[X = x] = 0$, portanto:

$$P[a \leq X \leq b] = P[a < X \leq b] = P[a \leq X < b] = P[a < X < b]$$

- Qualquer função $f(\cdot)$ que seja não negativa e cuja área total sob a curva seja igual à unidade caracterizará uma VA contínua.
- $f(x)$ **não** representa a probabilidade de ocorrência de algum evento. A área sob a curva entre dois pontos é que fornecerá a probabilidade.

Exemplo

Seja a função:

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

- Verifique se essa função é uma fdp.
- Calcule:
    - $P[X > 0]$
    - $P[X > 0,5]$
    - $P[-0,5 \leq X \leq 0,5]$
    - $P[X < -2]$
    - $P[X < 0,5]$
    - $P[X < 0 \cup X > 0,5]$

(Ver também exemplos 6.2 e 6.3).

Medidas de posição para VAs contínuas

- O valor esperado (ou média) da VA contínua X com função densidade $f(x)$, é dado pela expressão:

$$E(X) = \mu = \int_{-\infty}^{\infty} x\, f(x) dx.$$

- A mediana é o valor Md que tem a propriedade de

$$P(X \geq Md) \geq 0.5 \quad \text{e} \quad P(X \leq Md) \geq 0.5.$$

- A moda é o valor Mo tal que,

$$f(Mo) = \max_x f(x).$$

Variância para VAs contínuas

- Para uma VA X com densidade $f(x)$, a variância é dada por

$$Var(X) = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

- Expressão alternativa

$$Var(X) = \sigma^2 = E(X^2) - E(X)^2.$$

onde

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx.$$

Exemplo

Seja a função:

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

Calcule $E(X)$, $Var(X)$, $DP(X)$.

Exercícios recomendados

- Seção 6.1 - $1, 2, 3, 4$ e 5.
- Seção 6.2 - 1 a 9.