ACH3553 - Estatística I

Aula 9: Variáveis Aleatórias

Alexandre Ribeiro Leichsenring
alexandre.leichsenring@usp.br

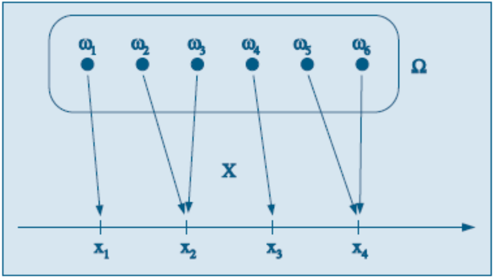
Organização

1. Variáveis aleatórias
   - Introdução

2. Valor médio e variância de variáveis aleatórias
   - Valor médio
   - Variância
   - Função de distribuição acumulada

Variável aleatória

Definição. Uma função X, definida no espaço amostral Ω e com valores num conjunto enumerável de pontos da reta é dita uma variável aleatória discreta.



A cada valor x_i da v.a. X associa-se sua probabilidade de ocorrência $\mathbf{P}(X = x_i)$, dada pela probabilidade do evento A de Ω cujos elementos correspondem ao valor x_i:

$$\mathbf{P}(X = x_i) = \mathbf{P}(A),$$

onde

$$A = \{\omega_1, \omega_2, \ldots\} \subset \Omega$$

Variável aleatória

Uma variável aleatória X, do tipo discreto, está bem caracterizada se indicarmos os possíveis valores que ela pode assumir:

$$x_1, x_2, \ldots, x_n, \ldots$$

e as respectivas probabilidades:

$$p(x_1), p(x_2), \ldots, p(x_n), \ldots$$

Ou seja, se conhecermos sua função de probabilidade:

$$(x, p(x))$$

ou

$$(x, \mathbf{P}(X = x))$$

.

Exemplo

Lançamos 3 vezes uma moeda. Se Y é o número de caras obtidas, então determine sua função de probabilidades.

Solução

Y é uma variável aleatória que assume os seguintes valores:

$$\begin{aligned} \{\bar{C},\bar{C},\bar{C}\} &\quad\Rightarrow\quad Y=0 \\ \{C,\bar{C},\bar{C}\},\{\bar{C},C,\bar{C}\},\{\bar{C},\bar{C},C\} &\quad\Rightarrow\quad Y=1 \\ \{C,C,\bar{C}\},\{C,\bar{C},C\},\{C,C,\bar{C}\} &\quad\Rightarrow\quad Y=2 \\ \{C,C,C\} &\quad\Rightarrow\quad Y=3 \end{aligned}$$

Função de probabilidades:

y	$p(y)$
0	1/8
1	3/8
2	3/8
3	1/8
Total	1,00

Exemplo

Lançamos 3 vezes uma moeda. Se Y é o número de caras obtidas, então determine sua função de probabilidades.

Solução

Y é uma variável aleatória que assume os seguintes valores:

$$\begin{aligned}
\{\bar{C}, \bar{C}, \bar{C}\} &\quad\Rightarrow\quad Y = 0 \\
\{C, \bar{C}, \bar{C}\}, \{\bar{C}, C, \bar{C}\}, \{\bar{C}, \bar{C}, C\} &\quad\Rightarrow\quad Y = 1 \\
\{C, C, \bar{C}\}, \{C, \bar{C}, C\}, \{C, C, \bar{C}\} &\quad\Rightarrow\quad Y = 2 \\
\{C, C, C\} &\quad\Rightarrow\quad Y = 3
\end{aligned}$$

Função de probabilidades:

y	$p(y)$
0	1/8
1	3/8
2	3/8
3	1/8
Total	1,00

Exercício

Uma urna tem 3 bolas azuis e 2 brancas. Retiramos 2 bolas sem reposição. Seja X o número de bolas brancas retiradas. Determine a distribuição de X.

Solução

$$\begin{aligned} \mathbf{P}(X=0) &= \frac{3}{10} \\ \mathbf{P}(X=1) &= \frac{6}{10} \\ \mathbf{P}(X=2) &= \frac{1}{10} \end{aligned}$$

x	0	1	2
$\mathbf{P}(X=x)$	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$

Exercício

Uma urna tem 3 bolas azuis e 2 brancas. Retiramos 2 bolas sem reposição. Seja X o número de bolas brancas retiradas. Determine a distribuição de X.

Solução

$$\begin{aligned}\mathbf{P}(X=0) &= \frac{3}{10}\\ \mathbf{P}(X=1) &= \frac{6}{10}\\ \mathbf{P}(X=2) &= \frac{1}{10}\end{aligned}$$

x	0	1	2
$\mathbf{P}(X=x)$	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$

Valor médio de uma variável aleatória

Exemplo

Uma indústria combina peças diferentes na montagem de conjuntos. Ela compra de dois fornecedores diferentes e tem lucros de venda diferentes de acordo com a compatibilidade das peças concatenadas. A tabela abaixo indica os valores que ele obtém de lucro e em qual proporção das peças montadas cada valor ocorre.

Lucro (R$)	Proporção
15	0,56
10	0,23
5	0,02
-5	0,19
Total	1

Qual o lucro médio por conjunto montado que a indústria deve conseguir?

Solução

O lucro médio será dado por:

$$\text{Lucro médio} = 0,56 \times 15 + 0,23 \times 10 + 0,02 \times 5 + 0,19 \times (-5) = 9,85$$

Valor médio de uma variável aleatória

Exemplo

Uma indústria combina peças diferentes na montagem de conjuntos. Ela compra de dois fornecedores diferentes e tem lucros de venda diferentes de acordo com a compatibilidade das peças concatenadas. A tabela abaixo indica os valores que ele obtém de lucro e em qual proporção das peças montadas cada valor ocorre.

Lucro (R$)	Proporção
15	0,56
10	0,23
5	0,02
-5	0,19
Total	1

Qual o lucro médio por conjunto montado que a indústria deve conseguir?

Solução

O lucro médio será dado por:

$$\text{Lucro médio} = 0,56 \times 15 + 0,23 \times 10 + 0,02 \times 5 + 0,19 \times (-5) = 9,85$$

Definição

O **valor médio** ou (valor esperado) de uma variável aleatória discreta X é definido por

$$\mathbf{E}(X) = \sum_{i=1}^{n} x_i \, \mathbf{P}(X = x_i) = \sum_{i=1}^{n} x_i \, p_i$$

onde x_i são os possíveis valores que a variável X pode assumir.

Obs. A expressão acima é semelhante àquela utilizada para a média de valores observados, onde no lugar das probabilidades (teóricas) p_i tínhamos as frequências relativas f_i.

Exemplo

Calcule o número esperado de caras no lançamento de 3 moedas.

Solução Seja Y = número de caras nos 3 lançamentos, então vimos que a função de probabilidades de Y é dada por:

y	$p(y)$
0	1/8
1	3/8
2	3/8
3	1/8
Total	1

O **valor médio** de Y é dado por:

$$\begin{aligned}
\mathbf{E}(Y) &= \sum_{i=1}^{n} y_i \cdot \mathbf{P}(Y = y_i) \\
&= 0 \cdot \mathbf{P}(Y=0) + 1 \cdot \mathbf{P}(Y=1) + 2 \cdot \mathbf{P}(Y=2) + 3 \cdot \mathbf{P}(Y=3) \\
&= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} \\
&= \frac{3}{2}
\end{aligned}$$

Exemplo

Calcule o número esperado de caras no lançamento de 3 moedas.

Solução Seja Y = número de caras nos 3 lançamentos, então vimos que a função de probabilidades de Y é dada por:

y	$p(y)$
0	1/8
1	3/8
2	3/8
3	1/8
Total	1

O **valor médio** de Y é dado por:

$$\begin{aligned}
\mathbf{E}(Y) &= \sum_{i=1}^{n} y_i \cdot \mathbf{P}(Y = y_i) \\
&= 0 \cdot \mathbf{P}(Y = 0) + 1 \cdot \mathbf{P}(Y = 1) + 2 \cdot \mathbf{P}(Y = 2) + 3 \cdot \mathbf{P}(Y = 3) \\
&= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} \\
&= \frac{3}{2}
\end{aligned}$$

Variância de uma variável aleatória

Além do valor médio (ou simplesmente *média*) definido acima, podemos considerar outras medidas de posição e variabilidade, como a mediana e o desvio padrão.

Definição

Chamamos de *variância* da variável aleatória X o valor

$$\mathrm{Var}(X) = \sum_{i=1}^{n} (x_i - \mathbf{E}(X))^2 \cdot \mathbf{P}(X = x_i)$$

Obs. O *desvio padrão* de X, $\mathrm{DP}(X)$, é definido como a raiz quadrada da variância.

Exemplo

Vamos calcular a variância do lucro por conjunto montado no exemplo do empresário.

Solução

Já sabemos que $\mathbf{E}(X) = 9,85$ e conhecemos a função de probabilidade de X:

x	$p(x)$
15	0,56
10	0,23
5	0,02
-5	0,19

A variância do lucro será dada por:

$$\begin{aligned}
\text{Var}(X) &= \sum_{i=1}^{n}(x_i - \mathbf{E}(X))^2 \cdot \mathbf{P}(X = x_i) \\
&= (15 - 9,85)^2 \times 0,56 + (10 - 9,85)^2 \times 0,23 + (5 - 9,85)^2 \times 0,02 + \\
&\quad +(-5 - 9,85)^2 \times 0,19 \\
&= 57,23
\end{aligned}$$

Função de distribuição acumulada

Definição

Dada a variável aleatória X, chamaremos de função de distribuição acumulada (f.d.a), ou simplesmente função de distribuição $F(x)$ à função

$$F(x) = \mathbf{P}(X \leq x)$$

A Função de Distribuição de Probabilidade nos fornece a probabilidade de que a variável aleatória assuma valores iguais ou menores do que um valor estabelecido.

Exercício

Calcule a função distribuição do número de caras (Y) no lançamento de 3 moedas.

Solução

Como vimos, a distribuição de Y é dada por

y	0	1	2	3
$\mathbf{P}(Y = y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Logo a **função de distribuição** de Y é dada por

y	0	1	2	3
$\mathbf{P}(Y \leq y)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	$\frac{8}{8}$

Exercício

Calcule a função distribuição do número de caras (Y) no lançamento de 3 moedas.

Solução

Como vimos, a distribuição de Y é dada por

y	0	1	2	3
$\mathbf{P}(Y = y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Logo a **função de distribuição** de Y é dada por

y	0	1	2	3
$\mathbf{P}(Y \leq y)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	$\frac{8}{8}$

Leitura

O conceito de **variável aleatória** está relacionada à descrição de modelos probabilísticos para variáveis de interesse (fenômenos reais).

Exemplo

Um empresário pretende estabelecer uma firma para montagem de um produto composto de uma esfera e um cilindro. As partes são adquiridas em fábricas diferentes (A e B), e a montagem consiste em juntar as duas peças e pintá-las. O produto acabado deve ter o comprimento (definido pelo cilindro) e a espessura (definida pela esfera) dentro de certos limites.

O empresário deseja estudar a viabilidade do seu empreendimento tendo uma ideia do lucro por peça montada.

Sabe-se que cada componente pode ser classificado como bom, longo ou curto, conforme sua medida esteja dentro da especificação, maior ou menor que a especificada, respectivamente. Além disso, foram obtidos dos fabricantes o preço de cada componente (R$5,00) e as probabilidades de produção de cada componente com as características bom, longo e curto:

Exemplo (continuação)

Produto		Fábrica A (Cilindro)	Fábrica B (Esfera)
Dentro das especificações	bom (B)	0,80	0,70
Maior que as especificações ...	longo (L)	0,10	0,20
Menor que as especificações ..	curto (C)	0,10	0,10

Se o produto final apresentar algum componente curto (C), ele será irrecuperável, e o conjunto será vendido como sucata a R$ 5,00. Cada componente longo (L) poderá ser recuperado a um custo adicional de R$ 5,00.

Se o preço de venda for R$ 25,00, como seria a distribuição de frequências da variável X: lucro por conjunto montado?

Solução

A construção dessa distribuição de frequências vai depender de certas *suposições* que faremos sobre o comportamento do sistema considerado. Com base nessas suposições, estaremos trabalhando com um *modelo* da realidade.

A distribuição que obtivermos será uma distribuição teórica, tanto mais próxima da distribuição de frequências real quanto mais fiéis à realidade forem as suposições.

Suposição: Classificação dos cilindros e das esferas são independentes.

Espaço amostral para a montagem dos conjuntos segundo as características de cada componente e suas probabilidades:

Produto	Probabilidade	Lucro por montagem (X)
BB	0,56	15
BL	0,16	10
BC	0,08	-5
LB	0,07	10
LL	0,02	5
LC	0,01	-5
CB	0,07	-5
CL	0,02	-5
CC	0,01	-5

Resumindo, vemos que X pode assumir os seguintes valores:

$$\begin{aligned} 15,\ \text{se ocorrer o evento } A_1 &= \{BB\} \\ 10,\ \text{se ocorrer o evento } A_2 &= \{BL, LB\} \\ 5,\ \text{se ocorrer o evento } A_3 &= \{LL\} \\ -5,\ \text{se ocorrer o evento } A_4 &= \{BC, LC, CB, CL, CC\} \end{aligned}$$

Cada um desses eventos tem uma probabilidade associada:

$$\mathbf{P}(A_1) = 0,56 \quad \mathbf{P}(A_2) = 0,23$$
$$\mathbf{P}(A_3) = 0,02 \quad \mathbf{P}(A_4) = 0,19$$

Modelo teórico para a distribuição do lucro por conjunto montado:

x	$p(x)$
15	0,56
10	0,23
5	0,02
-5	0,19
Total	1

A função $(x, p(x))$ é chamada *função de probabilidade da v.a. X*

Exercício

Considere a variável aleatória:

$$Y = \text{custo de recuperação de cada conjunto produzido}$$

Como será a função de probabilidade de Y?

Obs: Note que o custo de recuperação pode ser zero no caso de conjunto bom.

Solução

y	$p(y)$
0	0,75
5	0,23
10	0,02
Total	1

Exercício

Considere a variável aleatória:

$$Y = \text{custo de recuperação de cada conjunto produzido}$$

Como será a função de probabilidade de Y?

Obs: Note que o custo de recuperação pode ser zero no caso de conjunto bom.

Solução

y	$p(y)$
0	0,75
5	0,23
10	0,02
Total	1