

Tópico: Variáveis Aleatórias

Variáveis Aleatórias Discretas

- Variável que assume valores em um conjunto enumerável
 $\{x_1, x_2, x_3, \dots\}$
- Exemplo: Número de filhos - Com dados do último censo, a assistente social de um censo de saúde constatou que para as famílias da região:
 - 20% não têm filhos
 - 30% têm 1 filho
 - 35% têm 2 filhos
 - 5% têm 3 filhos
 - 5% têm 4 filhos
 - 5% têm 5 filhos

Variáveis Aleatórias Discretas

- Interesse: Estudo da variável $N =$ número de filhos
- Suponha que uma família será selecionada aleatoriamente nessa região e o número de filhos averiguado, o interesse é o estudo de N .
- N pode assumir os valores em $\Omega = \{0, 1, 2, 3, 4, 5\}$
 - $P(N = 0) = 0.2$
 - $P(N = 1) = 0.3$
 - $P(N = 2) = 0.35$
 - $P(N = 3) = P(N = 4) = P(N = 5) = \frac{0.15}{3} = 0.05$

Variáveis Aleatórias Discretas

- Finalmente, determinamos uma "função de probabilidade" para N :

N	0	1	2	3	4	5
x_i	x_1	x_2	x_3	x_4	x_5	x_6
p_i	0.20	0.30	0.35	0.05	0.05	0.05

- Tal que $\sum_{i=1}^6 p_i = 1$

Função Discreta de Probabilidade

- Função que atribui a cada valor da variável discreta X sua probabilidade
- $P(X = x_i) = p_i$, em que x assume os valores em $\Omega = \{x_1, x_2, \dots\}$
- $0 \leq p_i \leq 1$
- $\sum_i p_i = 1$

X	x_1	x_2	x_3	\dots
p_i	p_1	p_2	p_3	\dots

Exemplos

- 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 (mudanças para mais ou para menos), medidas corretivas serão necessárias, encarecendo o custo da obra.
 - admite-se que a probabilidade de ocorrência de alterações é de 0.1 para cada 5 metros
 - o custo básico inicial é de 100 UPCs e será acrescido de 50k, com k representando o número de alterações observadas

Exemplos

- Como se comporta a variável Custo de Obra de fundações?
- Assumimos que as alterações ocorrem independentemente entre cada um dos três intervalos de 5 metros
- A = ocorrência de alterações em cada intervalo
- 3 etapas $\Rightarrow 2^3 = 8$ possibilidades

Exemplos

Evento	Probabilidade	Custo
AAA	$(0.1)^3 = 0.001$	250
AA ^C A	$(0.1)^2(0.9) = 0.009$	200
AA ^C A	$(0.1)^2(0.9) = 0.009$	200
AA ^C A ^C	$(0.1)(0.9)^2 = 0.081$	150
A ^C AA	$(0.1)^2(0.9) = 0.009$	200
A ^C AA ^C	$(0.1)(0.9)^2 = 0.081$	150
A ^C A ^C A	$(0.1)(0.9)^2 = 0.081$	150
A ^C A ^C A ^C	$(0.9)^3 = 0.729$	100

Exemplos

- Note que associamos a cada evento do espaço amostral um valor da variável C (custo), e eventos diferentes podem corresponder ao mesmo valor de C .
- $c_1 = 100$, $c_2 = 150$, $c_3 = 200$, $c_4 = 250$
 - $P(C = c_1) = P(A^C A^C A^C) = 0.729$
 - $P(C = c_2) = P(AA^C A^C \cup A^C AA^C \cup A^C A^C A) = 3 \times 0.081 = 0.243$
 - $P(C = c_3) = P(AAA^C \cup AA^C A \cup A^C AA) = 3 \times 0.009 = 0.027$
 - $P(C = c_4) = P(AAA) = 0.001$

Exemplos

C	100	150	200	250
p_i	0.729	0.243	0.027	0.001

- O comportamento de C estudado através da probabilidade de ocorrência pode auxiliar na previsão de gastos e na elaboração de orçamentos

Exemplos

- Uma moeda é lançada duas vezes
- N = número de caras em dois lançamentos
- $\Omega = \{CC, CX, XC, XX\}$, em que C = cara e X = coroa

N	0	1	2
p_i	$P(XX) = \frac{1}{4}$	$P(CX \cup XC) = \frac{2}{4} = \frac{1}{2}$	$P(CC) = \frac{1}{4}$

Função de Distribuição Acumulada

- Um grupo de 1000 crianças foi analisado para determinar a efetividade de uma vacina contra um tipo de alergia. As crianças recebiam uma dose de vacina e após um mês passavam por um novo teste. Se tivessem tido alguma reação alérgica, recebiam outra dose (caso contrario o tratamento era concluído). Ao fim de 5 doses, foram consideradas imunizadas.
- Variável de interesse: $X =$ número de doses

Doses (X)	1	2	3	4	5
Freq.	245	288	256	145	66

Função de Distribuição Acumulada

- Uma criança é sorteada ao acaso, qual será a probabilidade dela ter recebido 2 doses?

$$P(X = 2) = \frac{288}{1000} = 0.288$$

Doses (X)	1	2	3	4	5
p_i	0.245	0.288	0.256	0.145	0.066

- Qual a probabilidade da criança ter recebido até duas doses?

$$P(X \leq 2) = P(X = 1) + P(X = 2) = 0.245 + 0.288 = 0.533$$

Função de Distribuição Acumulada

- Em geral, a função de distribuição acumulada (f.d.a.) de uma variável discreta X é definida por $F(x) = P(X \leq x)$
- Assim, se X assume os valores em $\Omega = \{x_1, x_2, \dots, x_n\}$:

$$F(x_1) = P(X = x_1)$$

$$F(x_2) = P(X = x_1) + P(X = x_2)$$

$$\vdots$$

$$F(x_n) = P(X = x_1) + \dots + P(X = x_n)$$

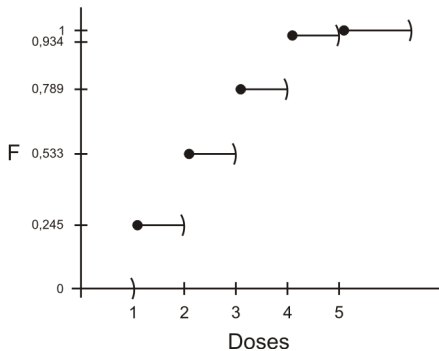
Função de Distribuição Acumulada

- Retomando o exemplo das vacinas, notemos que a f.d.a. de $X =$ número de doses é definido para qualquer valor real, logo:

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.245, & 1 \leq x < 2 \\ 0.533, & 2 \leq x < 3 \\ 0.789, & 3 \leq x < 4 \\ 0.934, & 4 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$

Função de Distribuição Acumulada

- f.d.a. de $X =$ número de doses:



Medidas de Posição para Variáveis Aleatórias Discretas

- A média, valor esperado ou esperança de uma variável aleatória discreta X , cuja f.d.p. é dada por $P(X = x_i) = p_i$ é dada pela expressão:

$$\mu_X = E(X) = \sum_{i \geq 1} x_i p_i$$

- A mediana (Md) é o valor médio que satisfaz $P(X \geq Md) \geq \frac{1}{2}$ e $P(X \leq Md) \geq \frac{1}{2}$
- A moda (Mo) é o valor da variável X que tem maior probabilidade de ocorrência:

$$P(X = Mo) = \max\{p_1, p_2, \dots\}$$

Medidas de Posição para Variáveis Aleatórias Discretas

- Exemplo:

X	-5	10	15	20
p_i	0.3	0.2	0.4	0.1

- $\mu_X = (-5) \times 0.3 + 10 \times 0.2 + 15 \times 0.4 + 20 \times 0.1 = 8.5$
- $Mo(X) = 15$
- $P(X \leq 10) = P(X \geq 15) = 0.5$, então a mediana
 $Md(X) = \frac{10+15}{2} = 12.5$
- Obs: note que nem a média (8.5) nem a mediana (12.5) são valores assumidos pela variáveis X .

Medidas de Posição para Variáveis Aleatórias Discretas

- Exemplo:

X	2	5	8	15	20
p_i	0.1	0.3	0.2	0.2	0.2

- $\mu_X = 10.3$
- $Mo(X) = 5$
- $Md(X) = 8$

Medidas de Posição para Variáveis Aleatórias Discretas

- Exemplo: Seja $Y = 5X - 10$

Y	0	15	30	65	90
p_i	0.1	0.3	0.2	0.2	0.2

- $\mu_Y = 41.5$
- $Mo(Y) = 15$
- $Md(Y) = 30$
- Note que, como $Y = 5X - 10$:

$$\mu_Y = 5\mu_X - 10$$

$$Mo(Y) = 5Mo(X) - 10$$

$$Md(Y) = 5Md(X) - 10$$

Medidas de Dispersão para Variáveis Aleatórias Discretas

- Seja uma variável aleatória discreta X , cuja f.d.p. é dada por $P(X = x_i) = p_i$ e média μ_X , a variância é a ponderação pelas respectivas probabilidades, dos desvios relativos à média elevados ao quadrado:

$$\sigma_X^2 = \text{Var}(X) = \sum_{i \geq 1} (x_i - \mu_X)^2 p_i$$

- Extraindo a raiz quadrada da variância, obtemos o desvio padrão, representado por σ_X

Medidas de Dispersão para Variáveis Aleatórias Discretas

- Exemplo:

X	2	5	8	15	20
p_i	0.1	0.3	0.2	0.2	0.2

- $\mu_X = 10.3$
- $\sigma_X^2 = 39.61$
- $\sigma_X = 6.294$

Medidas de Dispersão para Variáveis Aleatórias Discretas

- Exemplo: em um estudo sobre incidência de câncer, foi registrado para cada paciente com esse diagnóstico o número de casos de câncer em parentes próximos. Os dados de 26 pacientes são os seguintes:

Paciente	Incidência	Paciente	Incidência	Paciente	Incidência	Paciente	Incidência
1	2	8	3	15	5	22	4
2	5	9	3	16	2	23	0
3	0	10	2	17	2	24	0
4	2	11	0	18	3	25	3
5	1	12	1	19	2	26	3
6	5	13	1	20	1		
7	3	14	4	21	5		

Medidas de Dispersão para Variáveis Aleatórias Discretas

- Pela tabela, concluímos que X assume valores em $\Omega = \{0, 1, 2, 3, 4, 5\}$
- Estudos anteriores assumem que a incidência de câncer em parentes próximos pode ser teoricamente modelada pela seguinte função discreta:

incidência (x_i)	0	1	2	3	4	5
$p_i = P(X = x_i)$	0.1	0.1	0.3	0.3	0.1	0.1
$F(x_i) = P(X \leq x_i)$	0.1	0.2	0.5	0.8	0.9	1