Variáveis Aleatórias

- Uma variável, tal que, para cada conjunto de valores (ou valores individuais) que ela assume, pode-se atribuir probabilidades.
- Resultados de interesse associados a um experimento aletório.
- Notação: letras maiúsculas (variável aleatória), X, Y, Z, etc. Letras minúsculas, (valor assumido pela variável aleatória), x, y, z etc.
- Associa cada valor do espaço amostral a um número na reta, ou seja $X \equiv X(\omega) : \Omega \to \mathcal{R}$.

Exemplo: Lançar duas moedas consecutivamente

- Notação $C = cara$, $X = coroa$.
- $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$.
- $\omega_1 = (C, C)$; $\omega_2 = (C, X)$; $\omega_3 = (X, C)$; $\omega_4 = (X, X)$
- Defina X: número de coroas observadas na amostra.

$$\begin{aligned} \omega_1 \rightarrow X &= 0 \\ \omega_2 \text{ ou } \omega_3 \rightarrow X &= 1 \\ \omega_4 \rightarrow X &= 2 \end{aligned}$$

Variáveis Aleatórias Discretas

- Variável que assume valores em um conjunto finito $\{x_1, x_2, x_3, \ldots, x_k\}$ ou infinito enumerável $\{x_1, x_2, x_3, \ldots\}$.
- 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.
  - 15% têm igualmente 3, 4 ou 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 estudar o comportamento de N.
- N pode assumir os valores em $\Omega = \{0, 1, 2, 3, 4, 5\}$
  - $P(N = 0) = 0,2$
  - $P(N = 1) = 0,3$
  - $P(N = 2) = 0,35$
  - $P(N = 3) = P(N = 4) = P(N = 5) = \dfrac{0,15}{3} = 0,05$

Variáveis Aleatórias Discretas

- Finalmente, podemos determinar uma "função de probabilidade"(distribuição 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 ou função de probabilidade

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

X	x_1	x_2	x_3	...
p_i	p_1	p_2	p_3	...

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, medidas corretivas serão necessárias, encarecendo o custo da obra.

Exemplos

- 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 UPCs (Unidades Padrão de Construção) e será acrescido de 50k, com k representando o número de alterações observadas.

Exemplos

- Como se comporta a variável Custo da Obra das 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
AAA^C	$(0,1)^2(0,9) = 0,009$	200
AA^CA	$(0,1)^2(0,9) = 0,009$	200
AA^CA^C	$(0,1)(0,9)^2 = 0,081$	150
A^CAA	$(0,1)^2(0,9) = 0,009$	200
A^CAA^C	$(0,1)(0,9)^2 = 0,081$	150
A^CA^CA	$(0,1)(0,9)^2 = 0,081$	150
$A^CA^CA^C$	$(0,9)^3 = 0,729$	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.
- $c_1 = 100,\ c_2 = 150,\ c_3 = 200,\ c_4 = 250$
  - $P\left(C = c_1\right) = P\left(A^C A^C A^C\right) = 0,729$
  - $P\left(C = c_2\right) = P\left(AA^C A^C \cup A^C AA^C \cup A^C A^C A\right) = 3 \times 0,081 = 0,243$
  - $P\left(C = c_3\right) = P\left(AAA^C \cup AA^C A \cup A^C AA\right) = 3 \times 0,009 = 0,027$
  - $P\left(C = c_4\right) = P\left(AAA\right) = 0,001$

Exemplos

- Assim, temos a seguinte distribuição de probabilidade para a variável custo:

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. Caso ainda tivessem tido alguma reação alérgica, recebiam outra dose. 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?

  $P(X = 2) = \dfrac{288}{1000} = 0,288$

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 $\Omega = \{x_1, x_2, ..., x_n\}$:

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

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

  $\vdots$

  $F(x_n) = P(X = x_1) + ... + P(X = x_n)$

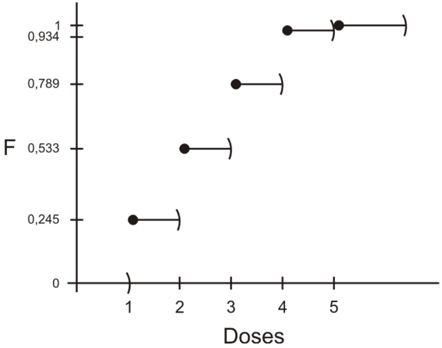
Função de Distribuição Acumulada

- 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 igual a:

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

- A mediana (Md) é o valor (médio, se existir mais de um) 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, ou seja:

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

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 $Md(X) = \dfrac{10 + 15}{2} = 12,5$
  - Obs: note que nem a média (8,5) nem a mediana (12,5) são valores assumidos pela variável 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 $Y = 5X - 10$, então a fdp de Y é dada por

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 $P\left(X = x_i\right) = p_i$ e média μ_X. Assim, a variância é a ponderação pelas respectivas probabilidades, dos desvios relativos à média elevados ao quadrado, ou seja:

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

- Extraindo a raiz quadrada (positiva) 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

  - $\mu_X = 2 \times 0,1 + 5 \times 0,3 + 8 \times 0,2 + 15 \times 0,2 + 20 \times 0,2 = 10,3$
  - $\sigma_X^2 = 0,1 \times (2 - 10,3)^2 + 0,3 \times (5 - 10,3)^2 + 0,2 \times (8 - 10,3)^2 + 0,2 \times (15 - 10,3)^2 + 0,2 \times (20 - 10,3)^2 = 39,61$
  - $\sigma_X = \sqrt{39,61} \approx 6,294$

- Uma outra forma de se calcular a variância (que também serve no contexto da análise descritiva) é $\sigma_X^2 = \sum_{i=1}^{n} x_i^2 p_i - \mu_x^2$. Qualquer uma das duas fórmulas pode ser utilizada.