Variáveis aleatórias: Introdução e variáveis aleatórias discretas (parte 1)

Notas de Aula da Professora Verónica González-López, digitadas por Beatriz Cuyabano, Pós-Graduação IMECC/UNICAMP, com modificações do Prof. Caio Azevedo

Introdução e motivação

- Vamos utilizar uma definição menos "rigorosa" porém mais simples de entender.
- Seja $(\Omega, \mathcal{A}, \mathcal{P})$ um espaço de probabilidade.
- Definição: uma variável aleatória (va ou v.a) consiste em uma variável, tal que, para cada conjunto de valores (ou valores individuais) que ela pode assumir, é possível atribuir probabilidades.
- É útil para representar (os) resultados de interesse associados a um experimento aleatório.

Introdução e motivação

- 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,

$$X : \Omega \to \mathcal{R}, (X(\omega) = a_{\omega}, \forall\, \omega \in \Omega, a_{\omega} \in \mathcal{R}).$$

- De uma outra forma, podemos entender uma variável aleatória como sendo uma função mensurável, ou seja

$$\forall A \in \mathcal{B}, B = \{\omega : X(\omega) \in A\} \in \mathcal{A}.$$

Introdução e motivação

- À rigor, existem cinco tipos de variáveis aleatórias: **discreta**, **contínua**, **absolutamente contínua**, mista e degenerada.
  - Essencialmente, estudaremos:
  - Discreta.
  - (Absolutamente contínua) contínua. Em nosso curso, não faremos distinção entre as duas
  - Degenerada (caso particular do tipo discreto e (absolutamente) contínuo).

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. Então

$$
\begin{aligned}
\omega_1 \to X &= 0, (X(\omega_1) = 0) \\
\omega_2 \,\text{ou}\, \omega_3 \to X &= 1, (X(\omega_2) = X(\omega_3) = 1) \\
\omega_4 \to X &= 2, (X(\omega_4) = 2)
\end{aligned}
$$

Variáveis Aleatórias Discretas (vad ou v.a.d.)

- Uma variável aleatória discreta é uma variável aleatória que assume valores (imagem) em um conjunto finito ($\{x_1, x_2, ..., x_n\}$) ou infinito enumerável ($\{x_1, x_2, x_3, ...\}$). Ou seja:

$$X : \Omega \rightarrow A, A \subset \mathcal{R},$$

em que A é finito ou infinito enumerável.

Variáveis Aleatórias Discretas (vad ou v.a.d.)

- 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: estudar a variável N : número de filhos.
- Suponha que uma família será selecionada aleatoriamente nessa região e o número de filhos (N) averiguado.
- Assim, a variável aleatória N pode assumir valores em $\Omega = \{0, 1, 2, 3, 4, 5\}$, de sorte que:
  - $P(N = 0) = 0,2$.
  - $P(N = 1) = 0,3$.
  - $P(N = 2) = 0,35$.
  - $P(N = 3) = P(N = 4) = P(N = 5) = \dfrac{0,15}{3} = 0,05$.

Variáveis Aleatórias Discretas

- Finalmente, determinamos uma "função de probabilidade" (ou 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

- Note que $\sum_{i=1}^{6} p_i = 1$.
- Uma outra forma de representar a tabela acima é através de

$$
\begin{aligned}
f_N(n) &= P(N=n) = 0,20\mathbb{1}_{\{0\}}(n) + 0,30\mathbb{1}_{\{1\}}(n) + 0,35\mathbb{1}_{\{2\}}(n) \\
&+ 0,05\mathbb{1}_{\{3,4,5\}}(n)
\end{aligned}
$$

Função de Probabilidade

- A função de probabilidade (fdp) ou função densidade de probabilidade (fdp), no caso discreto, é uma uma função que atribui, a cada valor que uma variável aleatória discreta X pode assumir, sua probabilidade de ocorrência.
- Seja A o suporte (conjunto de valores que a variável aleatória pode assumir, ou seja, sua imagem).
- Note que $A = \{x_1, x_2, ..., x_n\}$ (finito) ou $A = \{x_1, x_2, ...\}$ (infinito enumerável).

Função de Probabilidade

- Notação (A finito ou infinito enumerável):

$$P(X = x) = f_X(x) = h(x)\mathbb{1}_A(x),$$

em que $\mathbb{1}_A(x)$ é a usual função indicadora e $h(.)$ é uma função que atribui a probabilidade para cada $x \in A$.

- Note que

$$p_i = P(X = x_i) = f_X(x_i) = \begin{cases} h(x_i), & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases}$$

ou seja, a fdp está definida $\forall x \in \mathcal{R}$.

Função de Probabilidade

- Eventualmente (A finito):

$$
\begin{aligned}
f_X(x) &= \sum_{i=1}^{n} P(X = x_i)\mathbb{1}_{\{x_i\}}(x) \\
&= P(X = x_1)\mathbb{1}_{\{x_1\}}(x) + P(X = x_2)\mathbb{1}_{\{x_2\}}(x) + \ldots \\
&+ P(X = x_n)\mathbb{1}_{\{x_n\}}(x) \\
&= p_1\mathbb{1}_{\{x_1\}}(x) + p_2\mathbb{1}_{\{x_2\}}(x) + \ldots \\
&+ p_m\mathbb{1}_{\{x_n\}}(x).
\end{aligned}
$$

Função de Probabilidade

- Ainda (A finito):

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

em que $p_i = P(X = x_i) = f_X(x_i)$.

- Para que $f_X(.)$ seja, de fato, uma fdp, devemos ter
  - $0 \leq f_X(x) \leq 1, \forall x \in A$.
  - $\sum_{x \in A} f_X(x) = 1$.

Exemplo

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

Exemplo

- Como se comporta a variável **custo de obra de fundações**?
  - Vamos assumir que as alterações ocorrem independentemente entre cada um dos três intervalos de 5 metros
  - Seja o evento A : ocorre alteração no subsolo em cada intervalo.
  - O processo todo demanda três etapas $\Rightarrow 2^3 = 8$ possibilidades.

Exemplo

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

Exemplo

- Note que associamos a cada evento do espaço amostral um valor da variável aleatória C (custo), sendo que eventos diferentes podem corresponder a um mesmo valor de C.
- Suporte de C : $c_1 = 100$, $c_2 = 150$, $c_3 = 200$, $c_4 = 250$.
  - $P(C = c_1) = P\left(A^C A^C A^C\right) = 0,729.$
  - $P(C = c_2) = P\left(AA^C A^C \cup A^C AA^C \cup A^C A^C A\right) = 3 \times 0,081 = 0,243.$
  - $P(C = c_3) = P\left(AAA^C \cup AA^C A \cup A^C AA\right) = 3 \times 0,009 = 0,027.$
  - $P(C = c_4) = P(AAA) = 0,001.$

Exemplo

- Assim, temos que a fdp de C é dada por:

C	100	150	200	250
p_i	0,729	0,243	0,027	0,001

- Analogamente,

$$\begin{aligned} f_X(x) &= 0,729\mathbb{1}_{\{100\}}(c) + 0,243\mathbb{1}_{\{150\}}(c) \\ &+ 0,027\mathbb{1}_{\{200\}}(c) + 0,001\mathbb{1}_{\{250\}}(c). \end{aligned}$$

- O estudo da fdp de C pode auxiliar na previsão de gastos e na elaboração de orçamentos, por exemplo.

Exemplo

- Uma moeda é lançada duas vezes.
  - $\Omega = \{CC, CX, XC, XX\}$, em que C = cara e X = coroa.
  - Seja N : número de caras em dois lançamentos
  - Temos, assim, que a respectiva fdp é dada por

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

ou, analogamente,

$$f_x(x) = \frac{1}{4}\mathbb{1}_{\{0,2\}}(x) + \frac{1}{2}\mathbb{1}_{\{1\}}(x).$$

Função de Distribuição Acumulada (fda ou f.d.a.)

- Exemplo: um grupo de 1000 crianças foi analisado para determinar a efetividade de uma vacina contra um tipo específico 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, eram consideradas imunizadas.
- Variável de interesse: X : número de doses recebidas. Distribuição de frequências:

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

Função de Distribuição Acumulada

- Distribuição de probabilidades ("estimada"):

Doses (X)	1	2	3	4	5
p_i	0,245	0,288	0,256	0,145	0,066

ou, analogamente

$$\begin{aligned} P(X = x) &= 0,245\mathbb{1}_{\{1\}}(x) + 0,288\mathbb{1}_{\{2\}}(x) + 0,256\mathbb{1}_{\{3\}}(x) \\ &= 0,145\mathbb{1}_{\{4\}}(x) + 0,066\mathbb{1}_{\{5\}}(x) \end{aligned}$$

Função de Distribuição Acumulada

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

$$P\left(X=2\right)=\frac{288}{1000}=0,288.$$

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

$$P\left(X\leq 2\right)=P\left(X=1\right)+P\left(X=2\right)=0,245+0,288=0,533.$$

Função de Distribuição Acumulada

- Em geral, a função de distribuição acumulada (fda ou f.d.a.) de uma variável aleatória (qualquer), digamos X, é definida por

$$F_X(x) = P(X \leq x), \forall x \in \mathcal{R}. \tag{1}$$

- Obs: alguns livros (artigos) usam a seguinte definição :

$$F_X(x) = P(X < x), \forall x \in \mathcal{R}. \tag{2}$$

- Obs: No caso discreto, pode haver diferença (quase sempre há) entre os resultados obtidos pelas definições (1) e (2).

Função de Distribuição Acumulada

- Assim, se X é uma vad assumindo os valores em $\Omega = \{x_1, x_2, ..., x_n\}$ ou $\Omega = \{x_1, x_2, ..., x_n, ...\}$, temos que

$$\begin{aligned} F(x_1) &= P(X = x_1) \\ F(x_2) &= P(X = x_1) + P(X = x_2) \\ &\vdots \\ F(x_n) &= P(X = x_1) + ... + P(X = x_n). \end{aligned}$$

- Ou seja (no caso discreto, $\forall x \in \mathcal{R}$)):

$$F_X(x) = P(X \leq x) = \sum_{y \leq x} P(X = y) = \sum_{y \leq x} f_X(y).$$

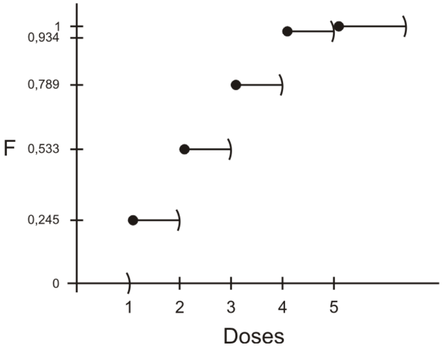
Função de Distribuição Acumulada

- Retomando o exemplo das vacinas, notemos que a f.d.a. de X : número de doses recebidas, é definida para qualquer valor real, ou seja:

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

Função de Distribuição Acumulada

- Graficamente, a fda anterior, é esboçada por:



Propriedades de uma fda (qualquer)

- Obs: quando dissermos que uma dada fda é discreta, significa que está associada à uma vad. O análogo vale para uma fda contínua (veremos adiante no curso).
- Seja $F_X(.)$ uma fda qualquer (discreta ou contínua), então:
  1. $F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$, e $F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1$.
  2. $F_X(.)$ é monótona não decrescente, ou seja, $\forall a < b, (a, b) \in \mathcal{R}^2$, $F_X(a) \leq F_X(b)$.
  3. $F_X(.)$ é contínua à direita, ou seja, $\lim_{0 < h \to 0} F_X(x + h) = F_X(x)$.
- Prova: exercício.

Propriedades de uma fda (qualquer)

- Exercício: Provar que uma legítima *fda* respeita os axiomas de Kolmogorov.
- À rigor, qualquer função, digamos F, tal que $F : \mathcal{R} \rightarrow [0, 1]$, e que apresente as propriedades 1,2,3 (slide anterior) é uma legítima fda.
- Uma vez que tenhamos $F_X(.)$, é possível obter a respectiva fdp através de:

$$f_X(x) = F_X(x) - \lim_{0<h\rightarrow 0} F_X(x+h), \forall x \in \mathcal{R}.$$

Função de sobrevivência (fds ou f.d.s)

- A fds de uma va (qualquer) X, é dada por:

$$S_X(x) = 1 - F_X(x) = P(X > x), \forall x \in \mathcal{R}.$$

- No caso discreto, a definição acima converte-se em:

$$S_X(x) = \sum_{y>x} P(X = y) = \sum_{y>x} f_X(y), \forall x \in \mathcal{R}$$

- O nome vem do fato desta quantidade ser muito utilizada na área de análise de sobrevivência.

Função de sobrevivência (fds ou f.d.s)

- Note que

$$F_X(x) + S_X(x) = 1, \forall x \in \mathcal{R}.$$

- Seja $S_X(.)$ uma fds qualquer (discreta ou contínua), então:
  1. $S_X(-\infty) = \lim_{x \to -\infty} S_X(x) = 1$, e $S_X(+\infty) = \lim_{x \to +\infty} S_X(x) = 0$.
  2. $S_X(.)$ é monótona não crescente, ou seja, $\forall a < b, (a, b) \in \mathcal{R}^2$, $S_X(a) \geq S_X(b)$.
  3. $S_X(.)$ é contínua à esquerda, ou seja, $\lim_{0 < h \to 0} S_X(x - h) = S_X(x)$.
- Demonstração: exercício.

Outras propriedades

- Sejam $a < b \in \mathcal{R}$ e X uma vad, então:

$$
\begin{aligned}
P(a \leq X \leq b) &= P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a), \\
P(a < X < b) &= P(X < b) - P(X < a), \\
P(a < X \leq b) &= P(X \leq b) - P(X < a) = F_X(b) - P(X < a), \\
P(a \leq X < b) &= P(X < b) - P(X \leq a) = P(X < b) - F_X(a).
\end{aligned}
$$

- Dem: Exercício

Outras propriedades

- Sejam $a < b \in \mathcal{R}$ e X uma vad, então:

$$
\begin{aligned}
P(X \leq a \cup X \geq b) &= P(X \leq a) + P(X \geq b) = F_X(a) + P(X \geq b), \\
P(X < a \cup X > b) &= P(X < a) + P(X > b) = P(X < a) + S_X(b), \\
P(X \leq a \cup X > b) &= P(X \leq a) + P(X > b) = F_X(a) + S_X(b), \\
P(X < a \cup X \geq b) &= P(X < a) + P(X \geq b).
\end{aligned}
$$

- Dem: Exercício

Transformações de variáveis aleatórias

- Seja X uma vad e defina $Y = g(X)$, de sorte que Y seja uma va. Então:
  - Se $g(.)$ for uma função um a um, então:

$$f_Y(y) = P(Y = y) = P(g(X) = y) = P(X = g^{-1}(y)) = f_X(g^{-1}(y)).$$

  - Se $g(.)$ não for uma função um a um, então:

$$\begin{aligned} f_Y(y) &= P(Y = y) = P(g(X) = y) = \sum_{x:g(x)=y} P(X = x) \\ &= \sum_{x:g(x)=y} f_X(x) \end{aligned}$$

Transformações de variáveis aleatórias

- A demonstração (à rigor feita no slide anterior) da propriedade anterior vem dos seguintes fatos:
  - A fdp de X (consequentemente de Y) já fornece(m) a própria distribuição de probabilidade.
  - Dividirmos os casos em que $g(.)$ é e não é uma transformação um a um.

Transformações de variáveis aleatórias

- Exemplo. Seja X, de sorte que:

X	0	1	2	4
$P(X = x)$	0,3	0,2	0,4	0,1

- Defina $Y = X^2$, então:

Y	0	1	4	16
$P(Y = y)$	0,3	0,2	0,4	0,1

- Ou seja, $P(Y = y) = P(X = \sqrt{y}), \forall y \in \mathcal{R}$.

Transformações de variáveis aleatórias

- Exemplo. Seja X, de sorte que:

X	-1	1	2	4
$P(X = x)$	0,3	0,2	0,4	0,1

- Defina $Y = X^2$, então:

Y	1	4	16
$P(Y = y)$	0,5	0,4	0,1

- Ou seja, $P(Y = y) = \sum_{x:x^2=y} P(X = x), \forall y \in \mathcal{R}$.