3.3. Variáveis aleatórias discretas

3.3.1. Conceitos básicos: funções, valor esperado, variância

3.3.1.1. Função de probabilidade

Um exemplo simples de variável aleatória discreta (VAD) é o número de *caras* obtidas no lançamento de três moedas. Este número estará entre 0 e 3, e podemos calcular as probabilidades de ocorrerem cada um destes valores (os cálculos foram feitos na seção **3.1.2**). Estas probabilidades são mostradas na Tabela **1**.

Tabela 1. Número de caras no lançamento de 3 moedas

X	p(x)
0	0,125
1	0,375
2	0,375
3	0,125
Σ	1,000

Representaremos a variável "número de caras" por *X*. As probabilidades associadas a cada valor desta variável formam uma *distribuição de probabilidades*, que mostra como as probabilidades estão distribuídas entre os diversos valores (compare esta distribuição com as *distribuições de frequências*, que vimos na Seção 2.1.4). O que nos interessa é construir um *modelo* para a variável. Este modelo será definido pela *função de probabilidades*, uma função que permite calcular a probabilidade associada a cada valor *x* da variável *X*. Representaremos esta função por *p(x)* (com "x" minúsculo):

$$P(X = x) = p(x) = P(\text{variável } X \text{ assumir um valor } x)$$

No exemplo do lançamento de três moedas,

$$p(x) = P(\text{“número de caras” ser igual a } x)$$

Há dois detalhes que é importante notar aqui. Primeiro, geralmente é usada esta convenção: o nome da variável é representado por uma letra maiúscula, enquanto o valor que ela assume num determinado experimento é representado por uma letra minúscula. Por exemplo, se lançamos a moeda três vezes e contamos as caras (*X*) e obtemos duas caras, teremos então x=2 neste experimento. Portanto, *X* é uma variável; *x* é um número, um dos valores possíveis de *X*. Esta distinção será útil mais tarde.

Segundo, alguns autores usam *f(x)* para representar a função de probabilidades de qualquer variável, outros preferem usar *p(x)* para VADs e *f(x)* para VACs. Uma vez que estes dois tipos de variáveis têm que ter tratamentos diferentes, iremos manter esta distinção ao longo deste textos; as funções *f(x)* serão apresentadas na seção **3.3.1**.

A função de probabilidades tem que satisfazer às seguintes propriedades, que são fáceis de compreender:

$$\Sigma p(x) = 1$$
$$0 \leq p(x) \leq 1$$

Isto é, o somatório de todas as probabilidades tem que ser igual a 1 (como o somatório das *frequências relativas* em uma tabela de distribuição de frequências tem que ser igual a 1). Além disso, as probabilidades têm todas que estar no intervalo entre 0 e 1, o que é exigido pela própria definição axiomática de probabilidade (Seção **3.1.1**).

3.3.1.2. Função de distribuição

Outra função importante, derivada da função de probabilidades, é a *função de distribuição*, ou *função acumulada de probabilidade*, definida como:

$$F(x_i) = P(X \leq x_i) = \sum_{k=1}^{i} x_k p(x_k)$$

Esta função dá a probabilidade acumulada de todos os valores da variável iguais ou menores que x_i ; isto é, a probabilidade de a variável assumir qualquer valor igual ou menor que x_i. A Tabela **2** mostra a $F(x_i)$ para o exemplo das três moedas mencionado acima. Note que $F(x_i)$ é o equivalente da frequência acumulada F, na distribuição de frequências. Esta função não será muito importante no estudo de VADs, mas será fundamental no estudo das VACs, visto mais adiante.

Tabela 2. Número de caras no lançamento de três moedas

X	p(x)	F(x)
0	0,125	0,125
1	0,375	0,500
2	0,375	0,875
3	0,125	1,000
∑	1,000	-

3.3.1.3. Valor esperado de uma VAD

A partir da função de probabilidades podem ser deduzidas as expressões para dois parâmetro importantes de um modelo, o *valor esperado* e a *variância*.

O *valor esperado* de uma VAD (também chamado de *valor médio, expectância* ou *esperança matemática*), representado por *E(X)*, é a média dos *n* valores que uma variável pode assumir, ponderados por suas probabilidades. O valor esperado é definido pela expressão:

$$E(X) = \sum_{i=1}^{n} x_i p(x_i) \quad (1)$$

Esta expressão surge do conceito de média ponderada (Seção 2.2.1). A *média ponderada* é uma forma de média em que cada valor de uma variável tem um *peso* diferente, definido segundo algum critério especificado. Suponhamos por exemplo que o número de carros que passaram por um cruzamento a cada minuto foi anotado durante um período de oito minutos, e o resultado encontrado foi o mostrado na Tabela **3**.

Tabela 3. Número de carros que passam por um cruzamento, por minuto

X	f	fr
0	4	0,500
1	2	0,250
2	1	0,125
3	1	0,125
Σ	8	1,000

O número médio de carros pelo cruzamento, por minuto, ponderado pelas frequências absolutas, será igual a:

$$média = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{0 \times 4 + 1 \times 2 + 2 \times 1 + 3 \times 1}{4 + 2 + 1 + 1} = \frac{7}{8} = 0{,}875$$

Note que n é o número de valores diferentes que a variável pode assumir neste exemplo, cada um representado por uma linha na tabela (n=4), não o número de observações feitas (Σf=8). Note também que uma média (ponderada ou simples) não tem necessariamente que ser um valor que a variável pode assumir; no exemplo, a variável só assume valores inteiros, mas o valor esperado não é inteiro.

O cálculo também pode ser feito usando as frequências relativas:

$$média = \frac{\sum_{i=1}^{n} x_i fr_i}{\sum_{i=1}^{n} fr_i} = \sum_{i=1}^{4} x_i fr_i = 0 \times 0{,}500 + 1 \times 0{,}250 + 2 \times 0{,}125 + 3 \times 0{,}125 = 0{,}875$$

Para uma VAD, o valor esperado é a média de sua distribuição de probabilidades, com cada valor x_i ponderado por sua probabilidades $p(x_i)$. Como o somatório de $p(x_i)$ é igual a 1, chegamos à expressão do valor esperado:

$$E(X) = \frac{\sum_{i=1}^{n} x_i p(x_i)}{\sum_{i=1}^{n} p(x_i)} = \sum_{i=1}^{n} x_i p(x_i)$$

Na seção **3.1.3** vimos um exemplo no qual X era o *número de caras no lançamento de três moedas,* cuja distribuição de probabilidades foi dada na Tabela **1**. O valor esperado de X é calculado como:

$$E(X) = \sum_{i=1}^{n} x_i p(x_i) = 0 \times 0{,}125 + 1 \times 0{,}375 + 2 \times 0{,}375 + 3 \times 0{,}125 = 1{,}5$$

Para que serve este conceito de valor esperado? Podemos dizer, simplificando, que o valor esperado é o valor que se espera que a VA assuma em média, se o experimento for repetido muitas vezes. Se lançarmos três moedas várias vezes, em média esperaremos encontrar em média 1,5 caras em cada lançamento. Veremos agora um exemplo de como o conceito de valor esperado se aplica a um problema de VAD.

Exemplo 1

Lançamos três dados e contamos o número deles que mostram a face 6. As probabilidades associadas a este experimento são mostradas na Tabela **4** (foram calculadas na seção **3.1.6**)

Tabela 4. Lançamento de três dados (X: número deles que mostram a face 6)

X	p(x)
0	0,5787
1	0,3472
2	0,0694
3	0,0046
Σ	1,0000

O valor esperado da variável X é calculado como:

$$E(X) = \sum_{i=1}^{4} x_i p(x_i) = 0 \times 0{,}5787 + 1 \times 0{,}3472 + 2 \times 0{,}0694 + 3 \times 0{,}0046 = 0{,}5$$

Exemplo 2

Se um jogador aposta no número que sai numa roleta (Fig. **1**) e acerta recebe um prêmio de 35 para 1. Numa roleta européia, que tem 37 casas (numeradas de 0 a 36), qual será o lucro esperado do jogador em cada rodada? Se o jogador aposta na "cor" (preta ou vermelha), recebe um prêmio de apenas 1 para 1 (as roletas têm 18 números em casas pretas, 18 em casas vermelhas, e um zero numa casa verde; o jogador não pode apostar neste zero). Qual destes dois tipos de aposta é mais vantajoso para o apostador?

Resposta:

Se o jogador aposta \$ 1 em um número qualquer, a variável X (lucro do jogador) pode assumir dois valores: se ele acerta, recebe \$ 35 de prêmio; se ele erra, perde o \$ 1 que apostou (portanto, o lucro é negativo: $X=-1$). A probabilidade de ele acertar é de 1/37; a de ele errar, 36/37. A função de probabilidades de X é mostrada na Tabela **5**.

Tabela 5

X	p(x)
35	1/37
-1	36/37
Σ	1

Figura 1. Mesa de roleta européia

O valor esperado do lucro será portanto de

$$E(X) = \sum_{i=1}^{n} x_i p(x_i) = 35 \times \frac{1}{37} + (-1) \times \frac{36}{37} = -\frac{1}{37} = -0{,}027$$

Se o jogador aposta na cor, a variável X terá dois valores possíveis: X=1 (se ele acerta) e X=−1 (se ele erra). A função de probabilidades é mostrada na Tabela 6.

Tabela 6

X	p(x)
1	18/37
-1	19/37
Σ	1

O valor esperado do lucro :

$$E(X) = \sum_{i=1}^{n} x_i p(x_i) = 1 \times \frac{18}{37} + (-1) \times \frac{19}{37} = -\frac{1}{37} = -0{,}027$$

Tabela 7

X	p(x)
2	12/37
-1	25/37
Σ	1

Um exemplo adicional: o jogador pode também apostar na *coluna*. Há três colunas na mesa; a primeira contém os números 1, 4, 7, etc., a segunda contém os números 2, 5, 8, etc., e a terceira contém os números restantes. Se o jogador aposta numa coluna, e o número sorteado pela roleta é um dos que está naquela coluna, o jogador ganha 2 para 1 (como está escrito ao pé de cada coluna: *2 to 1*). A variável X terá dois valores possíveis: X=2 (se ele acerta) e X=−1 (se ele erra). A função de probabilidades é mostrada na Tabela 7.

O valor esperado do lucro é então:

$$E(X)=\sum_{i=1}^{n} x_i p(x_i) = 2\times\frac{12}{37}+(-1)\times\frac{25}{37}=-\frac{1}{37}=-0{,}027$$

O lucro do jogador será portanto igual nos três casos, e sempre negativo; ou seja, a cada vez que o jogador aposta $1, ele está em média perdendo $ 0,027. Existem várias outras maneira de apostar numa roleta, e todas levam ao mesmo resultado. As roletas americanas são diferentes das européias, pois além de um zero (0), têm também um duplo zero (00). Que diferença você acha que isto faz, do ponto de vista do jogador?

Exemplo 3

Uma pessoa paga $ 100 para fazer um seguro, e receberá $ 1000 caso se acidentar dentro de um ano. Se as estatísticas dizem que a probabilidade de uma pessoa se acidentar neste prazo é de 0,05, qual será o lucro esperado anual da companhia de seguros, por pessoa segurada ?

Resposta

A variável X: lucro da companhia (por pessoa, por ano) pode assumir dois valores: X = 100 (se a pessoa não se acidenta) e X = −900 (se a pessoa se acidenta, a companhia lhe devolve os 100, e mais 900). A probabilidade de a pessoa se acidentar (e portanto a companhia ter prejuízo) é de 0,05; de não se acidentar, igual a 0,95. A função de probabilidades de X é mostrada na Tabela 8.

Tabela 8

X	p(x)
100	0,95
-900	0,05
Σ	1,00

O valor esperado do lucro é:

$$E(X)=\sum_{i=1}^{n} x_i p(x_i) = 100\times 0{,}95+(-900)\times 0{,}05 = 50$$

Portanto, a companhia seguradora tem um lucro médio de $ 50 por ano, por pessoa segurada. (Existe aliás uma área de estudo, a *Atuária* - bastante relacionada com a Estatística -, dedicada a avaliar o risco de operações financeiras, como os seguros ou qualquer forma de investimento, com base em conceitos da teoria de probabilidades).

3.3.1.4. Variância de uma VAD

Qualquer variável, aleatória ou não, tem dispersão (se não houvesse dispersão, não seria uma variável, e sim uma constante!). A variância de uma VAD, representada por V(X), é definida como:

$$V(X) = \sum_{i=1}^{n} \left(x_i - E(X)\right)^2 p(x_i) \tag{2}$$

onde n é o número de valores diferentes que a variável pode assumir.

O conceito de *variância* de uma VAD é baseado no mesmo conceito de variância de uma amostra (seção **2.2.2.4**), definida por:

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{X})^2$$

Para variáveis discretas representados em tabelas de frequências, cada valor da variável deve ser ponderado por sua frequência:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^2 f_i}{\sum_{i=1}^{n} f_i}$$

Por exemplo, se usarmos os dados da Tab. **3** que registra o número de carros que passam por um cruzamento, por minuto, num período total de oito minutos. A variância desta amostra pode ser calculada como:

$$\overline{X} = 0{,}875$$

$$s^2 = \frac{\sum_{i=1}^{4} (x_i - 0{,}875)^2 f_i}{\sum_{i=1}^{4} f_i} =$$

$$s^2 = \frac{1}{4+2+1+1}\left[(0-0{,}875)^2 \times 4 + (1-0{,}875)^2 \times 2 + (2-0{,}875)^2 \times 1 + (3-0{,}875)^2 \times 1\right]$$

$$s^2 = 1{,}11$$

Se usarmos frequências relativas, ao invés das frequências absolutas, a expressão é mais simples (já que o somatório das freqüências relativas é sempre igual a 1):

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^2 f_i}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} (x_i - \overline{X})^2 fr_i \tag{3}$$

A idéia que dá origem a definição de variância de um VAD é a mesma: a variância é o somatório ponderado dos quadrados dos desvios de cada valor da variável X em relação à sua média (no caso das VADs, em relação a seu valor esperado). A diferença é que nas VADs a ponderação é feita pelas *probabilidades* de cada valor, ao invés de pelas *freqüências relativas*. Comparando a eq. (**3**) com a eq. (**2**). vemos que as diferenças são que a média $\overline{X}$ de uma amostra é substituída pelo valor esperado de uma VA, e que as frequências relativas *fr* são substituídas pelas probabilidades *p(x)*.

Note que a média ponderada do quadrado dos desvios

$$(x_i - E(X))^2$$

nada mais é do que o valor esperado do quadrado dos desvios de X em relação a seu valor esperado; podemos portanto re-escrever a variância como na eq. (**4**):

$$V(X) = E\left[(X - E(X))^2\right] \tag{4}$$

Demonstra-se que está fórmula por sua vez pode ser re-escrita como na eq. (**5**), que às vezes é útil os cálculos.

$$V(X) = E(X^2) - [E(X)]^2 \tag{5}$$

Exemplo 1 (cont.)

Retornemos ao Exemplo 1 da seção anterior (lançamento de três dados). As probabilidades associadas a este experimento são mostradas na Tabela **4**. O valor esperado de X foi calculado como E(X) = 0,5. A variância de X será dada por:

$$V(X) = \sum_{i=1}^{n} (x_i - E(X))^2 \, p(x_i)$$

$$V(X) = (0-0{,}5)^2 \times 0{,}5787 + (1-0{,}5)^2 \times 0{,}3472 +$$
$$+ (2-0{,}5)^2 \times 0{,}0694 + (3-0{,}5)^2 \times 0{,}0046 = 0{,}4167$$

Exemplo 4

Suponha que um dado comum seja lançado, e a variável de interesse seja o número mostrado na sua face superior. Os valores possíveis desta variável são 1, 2, 3, 4, 5, e 6; as probabilidades são P=1/6, para qualquer valor. O valor esperado da variável portanto é igual a:

$$E(X) = \sum_{i=1}^{n} x_i p(x_i) = (1+2+3+4+5+6)\frac{1}{6} = 3{,}5$$

A variância desta variável será igual a:

$$V(X) = \sum_{i=1}^{n} (x_i - E(X))^2 \, p(x_i)$$

$$V(X) = (1-3{,}5)^2 \times \frac{1}{6} + (2-3{,}5)^2 \times \frac{1}{6} + (3-3{,}5)^2 \times \frac{1}{6} +$$
$$+ (4-3{,}5)^2 \times \frac{1}{6} + (5-3{,}5)^2 \times \frac{1}{6} + (6-3{,}5)^2 \times \frac{1}{6}$$
$$V(X) = 2{,}917$$

Suponha agora que uma moeda tenha suas duas faces marcadas com os números 3 e 4. Se esta moeda é lançada, quais serão o valor esperado e a variância do número mostrado? O valor esperado será a média entre 3 e 4, obviamente igual ao do dado comum no exemplo acima. Podemos prever porém que a variância deve ser menor (basta compararmos as amplitudes totais: num dado, X varia de 1 a 6; nesta moeda, irá variar apenas de 3 a 4). De fato, calculando E(X) e V(X), obtemos:

$$E(X)=\sum_{i=1}^{n} x_i\,p(x_i)=(3+4)\frac{1}{2}=3{,}5$$

$$V(X)=\left[(3-3{,}5)^2+(4-3{,}5)^2\right]\times\frac{1}{2}=0{,}25$$

Exemplo 2 (cont.)

No exemplo 2 acima, comparamos as diferentes formas de se apostar numa roleta; a conclusão foi a de que todas as formas levam ao mesmo lucro esperado, E(X) = −1/37. Comparemos agora a variância dos lucros destas formas de aposta: aposta *en plein* (num número escolhido), na cor (preta ou vermelha), ou na coluna.

(i) aposta *en plein* (usando os dados da Tabela 5)

$$V(X)=\sum_{i=1}^{n}\left(x_i-E(X)\right)^2 p(x_i)$$

$$V(X)=\left(35-\frac{-1}{37}\right)^2\times\frac{1}{37}+\left(-1-\frac{-1}{37}\right)^2\times\frac{36}{37}\cong 34{,}080$$

(ii) aposta na cor (usando os dados da Tabela 6)

$$V(X)=\left(1-\frac{-1}{37}\right)^2\times\frac{18}{37}+\left(-1-\frac{-1}{37}\right)^2\times\frac{19}{37}\cong 0{,}999$$

(iii) aposta na coluna (usando os dados da Tabela 7)

$$V(X)=\left(2-\frac{-1}{37}\right)^2\times\frac{12}{37}+\left(-1-\frac{-1}{37}\right)^2\times\frac{25}{37}=1{,}972$$

Conclusão: as três formas de apostar dão o mesmo resultado médio (valor esperado), mas a aposta *en plein* é a que leva a um resultado com maior variância; a aposta na *cor* leva à menor variância.

3.3.1.5. Propriedades do valor esperado e da variância de uma VA

Listamos abaixo algumas das propriedades do valor esperado e da variância de uma variável aleatória. Na seção 3.3.2, iremos ilustrar algumas delas, usando um modelo de distribuição *uniforme*.

(i) Propriedades do valor esperado:

Valor esperado...

- de uma constante C:
  Se $X = C \rightarrow$ $E(X) = C$

- da soma da variável X com uma constante C:
  se $Y = X + C \rightarrow$ $E(Y) = E(C+X) = C + E(X)$

- do produto da variável X por uma constante C:
  se $Y = CX \rightarrow$ $E(Y) = E(CX) = CE(X)$

- da soma de duas ou mais variáveis
  se $Y = X_1+X_2 \rightarrow$ $E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2)$
  se $Y = X_1+X_2+...+X_k \rightarrow$ $E(Y) = E(X_1) + E(X_2) + ... +E(X_k)$

- de uma função linear da VA
  se $Y = aX + b \rightarrow$ $E(Y) = E(aX+b) = aE(X) + b$

- do produto de VAs independentes (esta regra não serve se as VAs forem dependentes probabilisticamente!)
  se $Y = X_1X_2 \rightarrow$ $E(Y) = E(X_1X_2) = E(X_1).E(X_2)$

(ii) Propriedades da variância

Variância ...

- de uma constante C:
  Se $Y = C \rightarrow$ $V(X) = 0$

- da soma da variável X com uma constante C:
  se $Y = X + C \rightarrow$ $V(Y) = V(X+C) = V(X)$

- do produto da variável X por uma constante C:
  se $Y = CX \rightarrow$ $V(Y) = V(CX) = C^2V(X)$

- da soma de VAs independentes (esta regra não serve se as VAs forem dependentes probabilisticamente!
  se $Y = X_1 + X_2 \rightarrow$ $V(Y) = V(X_1 + X_2) = V(X_1) + V(X_2)$
  se $Y = X_1 + X_2 + ... + X_k \rightarrow$ $V(Y) = V(X_1) + V(X_2) + ... + V(X_k)$