Probabilidade Bivariada

Prof. Lorí Viali, Dr.
viali@mat.ufrgs.br
http://www.mat.ufrgsbr/~viali/

1/3

Motivação

Em muitas situações precisamos lidar com duas ou mais variáveis aleatórias ao mesmo tempo. Por exemplo o comprimento e a largura de uma determinada peça.

Distribuições Multivariadas

Uma distribuição de probabilidade pode ser unidimensional ou n-dimensional. Distribuições n-dimensionais ($n \geq 2$) são denominadas de distribuições multivariadas.

VA n-dimensional

Se $X_1, X_2, ..., X_n$ forem "n" funções, cada uma associando um número real a cada resultado $s \in S$, denominaremos $(X_1, X_2, ..., X_n)$ de variável aleatória n-dimensional.

Um caso especial de distribuição multivariada envolve uma variável aleatória bidimensional que é denominada de distribuição bivariada.

VADB

A variável (X, Y) será uma **variável aleatória discreta bidimensional** se os valores possíveis de (X, Y) forem finitos ou infinitos numeráveis, isto é, os valores possíveis são (x_i, y_j) com $i = 1, 2, 3, ...$ e $j = 1, 2, 3, ...$

A função de probabilidade

A cada variável aleatória discreta bidimensional está associada uma função de probabilidade que satisfaz as seguintes condições:

(i) $p(x_i, y_j) \geq 0$ para i, j = 1, 2, 3, ...

(ii) $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p(x_i, y_j) = 1$

A função "p" definida para todo (x_i, y_j) no contradomínio de (X, Y) é denominada de **função de probabilidade** de (X, Y).

A distribuição conjunta

A coleção dos pares:

$[(x_i, y_j), p(x_i, y_j)]$, i, j = 1, 2, 3, ...

é, denominada de distribuição de probabilidade conjunta da variável aleatória discreta bidimensional (X, Y).

Exemplo:

Uma pequena fábrica opera com dois turnos diários. Em um estudo sobre o padrão de ausências ao trabalho as duas variáveis aleatórias de interesse são: X = número de faltas no turno da manhã e Y = número de ausências no turno da tarde do mesmo dia.

Baseado numa longa série de registros, de funcionários, o diretor de pessoal, determinou a distribuição conjunta de X e Y mostrada na tabela (próxima lâmina).

Distribuição Conjunta

X \ Y	0	1	2	3	Σ
0	0,05	0,05	0,10	0,00	**0,20**
1	0,05	0,10	0,25	0,10	**0,50**
2	0,00	0,15	0,10	0,05	**0,30**
Σ	**0,10**	**0,30**	**0,45**	**0,15**	**1,00**

Interpretação

Na tabela o valor 0,25 da célula (X = 1, Y = 2) significa que em 25% dos dias um trabalhador faltou no turno da manhã e dois faltaram no turno da tarde. O valor de 20% da soma da primeira linha, indica que em 20% dos dias ninguém faltou no turno da manhã.

Da mesma forma, o valor de 45% na quarta coluna, indica que em 45% dos dias, dois trabalhadores não compareceram no turno da tarde.

Exercício:

Suponha que uma palavra da frase: "**O Grêmio é e sempre será o melhor time gaúcho**" é selecionada ao acaso.

Sejam:

X = tamanho da palavra e

Y = número de vogais da palavra, duas variáveis aleatórias.

Determinar a distribuição conjunta de (X, Y).

Distribuição Conjunta

X \ Y	1	2	3
1	0,4	0,0	0,0
4	0,0	0,2	0,0
6	0,0	0,2	0,2

Distribuições Marginais

A cada variável bidimensional (X, Y) estão associados duas variáveis aleatórias X e Y.

Os valores de X considerados em conjunto com as probabilidades que aparecem na última coluna à direita formam a distribuição marginal de X e os valores de Y considerados com as probabilidades da última linha da tabela formam a distribuição marginal de Y.

Distribuições Marginais

No caso discreto a obtenção da distribuição marginal de X é dado por:

$p(x_i) = P(X = x_i)$

$= P(X = x_i, Y = y_1 \text{ ou } Y = y_2 \text{ ou } ....)$

$= \sum_{j=1}^{\infty} p(x_i, y_j)$

Definição:

Se (X, Y) é uma VA discreta bidimensional, então as coleções de pares:

[x, p(x) = P(X = x)] e

[y, p(y) = P(Y = y)]

são denominadas de **Distribuições Marginais**.

Exemplo:

Considerando a distribuição conjunta (X, Y) onde X = faltas no turno da manhã e Y = faltas no turno da tarde, tem-se, as seguintes distribuições marginais.

Distribuições Marginais

x	p(x)
0	0,20
1	0,50
2	0,30
Σ	**1,00**

y	p(y)
0	0,10
1	0,30
2	0,45
3	0,15
Σ	**1,00**

Exercício:

Considerando a distribuição conjunta (X, Y) onde X = tamanho da palavra e Y = número de vogais, determine as distribuições marginais.

Solução

X \ Y	1	2	3	Σ
1	0,4	0,0	0,0	**0,4**
4	0,0	0,2	0,0	**0,2**
6	0,0	0,2	0,2	**0,4**
Σ	**0,4**	**0,4**	**0,2**	**1,0**

Distribuições Condicionais

No estudo descritivo com distribuições conjuntas foram calculados os percentuais em relação as linhas e as colunas. Na probabilidade este cálculo é denominado de Distribuições Condicionais.

Definição:

Seja x_i um valor da VAD X tal que $p(x_i) > 0$. A probabilidade:

$P(Y = y_j \mid X = x_i) =$

$P(X = x_i, Y = y_j) / P(X = x_i) =$

$P(y_j \mid X = x_i)$ para $j = 1, 2, \ldots, n$

é denominada probabilidade condicional de $Y = y_j$, dado que $X = x_i$.

Exemplo:

Com base na tabela do número de ausências ao trabalho nos turnos da manhã e da tarde, determine a distribuição condicional de $P(Y \mid x = 0)$.

$P(Y = y \mid x = 0) =$

$= P(Y = y; x = 0)/P(x = 0) =$

$= P(Y \mid x = 0).$

Assim para y = 0, 1, 2 e 3, tem-se:

$p(0 \mid x = 0) = P(y = 0; x = 0)/P(x = 0) =$

$= 0,05/0,20 = 0,25.$

$p(1 \mid x = 0) = P(y = 1; x = 0)/P(x = 0) =$

$= 0,05/0,20 = 0,25.$

p(2 | x = 0) = P(y = 2; x = 0)/P(x = 0) = = 0,10/0,20 = 0,50.

p(3 | x = 0) = P(y = 3; x = 0)/P(x = 0) = = 0/0,20 = 0.

y	p(y | x = 0)
0	0,25
1	0,25
2	0,50
∑	1,00

Exercício:

Qual é a distribuição do número de vogais se o tamanho da palavra é seis.

Solução

X \ Y	1	2	3	Σ
1	0,4	0,0	0,0	**0,4**
4	0,0	0,2	0,0	**0,2**
6	0,0	0,2	0,2	**0,4**
Σ	**0,4**	**0,4**	**0,2**	**1,0**

p(2 | x = 6) = P(y = 2; x = 6)/P(x = 6) = = 0,2/0,4 = 0,5.

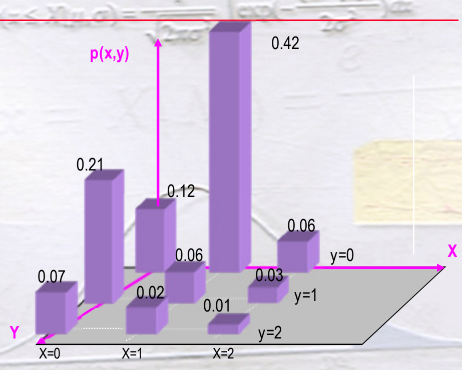
p(3 | x = 6) = P(y = 3; x = 6)/P(x = 6) = = 0,2/0,40 = 0,5.

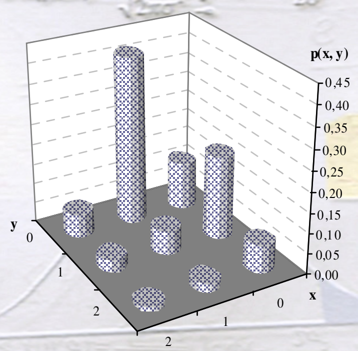
y	P(y | x = 6)
2	0,5
3	0,5
Σ	**1,0**

Representação gráfica

Representar graficamente a distribuição bivariada apresentada na tabela abaixo:

Y \ X	0	1	2
0	0,12	0,42	0,06
1	0,21	0,06	0,03
2	0,07	0,02	0,01





Descrevendo a distribuição

- A distribuição conjunta pode ser descrita pela média, variância e desvio padrão de cada uma das variáveis;
- Isto é feito através das distribuições marginais.

Descrevendo a distribuição

Por exemplo, considerando a distribuição representada graficamente, tem-se:

x	p(x)	p(y)
0	0,4	0,6
1	0,5	0,3
2	0,1	0,1

E(X) = 0,70

V(X) = 0,41

E(Y) = 0,50

V(Y) = 0,45

Ela, pode ainda ser descrita em termos do relacionamento entre as duas variáveis. Para descrever o relacionamento entre as duas variáveis existe a covariância e o coeficiente de correlação.

A covariância

A covariância entre as variáveis aleatórias X e Y é dada por: COV(X, Y) = E[(X – E(X)(Y – E(Y)]. Ou seja, é o valor médio do produto dos desvios de X e Y em relação as suas médias.

Se X assumir os valores $x_1, x_2, \ldots, x_n$ e Y os valores $y_1, y_2, \ldots, y_m$ e $P(X = x_i, Y = y_j) = p(x_i, y_j)$, então a expressão:

COV(X, Y) = E[(X – E(X)(Y – E(Y)] =

= E(XY) – E(X)E(Y).

Pode ser escrita como:

$$COV(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}[x_i - E(X)][y_j - E(Y)].p(x_i, y_j) =$$

$$= E(XY) - E(X)E(Y) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i\, y_j . p(x_i, y_j) - \sum_{i=1}^{n} x_i . p(x_i) . \sum_{j=1}^{m} y_j . p(y_j)$$

Exemplo:

Calcular a covariância da distribuição anterior.

$COV(X, Y) = \Sigma\Sigma(X- \mu_X)(Y- \mu_Y)p(x, y) =$

$= (0 - 0.7)(0 - 0.5)p(0, 0) + \ldots$

$+ \ldots +$

$= (2 - 0.7)(2 - 0.5)p(2, 2) =$ **-0.15**

O Coeficiente de Correlação

A covariância varia no intervalo $(-\infty, +\infty)$ e é, portanto difícil de interpretar. Por isto trabalha-se, em geral, com o coeficiente de correlação, que é calculado da seguinte forma: $\rho = COV(X, Y)/\sigma_X\sigma_Y$.

A grande vantagem da utilização do coeficiente de correlação é que ele varia no intervalo [-1; +1], sendo assim de fácil interpretação.

Quando mais próximo de -1 ou +1 estiver o coeficiente maior é a associação linear entre as variáveis. Um coeficiente próximo de zero, indica ausência de relacionamento linear.

Exemplo:

Para a tabela, do exemplo anterior, o coeficiente de correlação será:

$$\rho = \frac{COV(X,Y)}{\sigma_X \sigma_Y} = \frac{-0{,}15}{\sqrt{0{,}41.045}} = -0{,}35$$

Estas expressões para o cálculo da covariância e do coeficiente de correlação não são práticas. Existe uma maneira mais simples do obter estes resultados. Antes, porém, é necessário definir mais alguns conceitos.

Funções de Variáveis Aleatórias

- Função de uma variável aleatória;
- Função de duas variáveis aleatórias:
  - Resultado unidimensional;
  - Resultado bidimensional.

01. Variável Aleatória Discreta

Função de uma variável aleatória

Seja X uma variável aleatória discreta com fp $p(x_i)$. Seja $Y = f(x)$. Se X for monótona, então $y_i = f(x_i)$, onde x_i são os valores de X, com probabilidades: $P(Y = y_i) = P(X = x_i)$

Se X não for monótona, então aos valores possíveis $y_i = f(x_i)$ de Y se associará a probabilidade igual a soma das probabilidades dos valores de X pertencente à imagem inversa de y_i por f.

Exemplo um:

Determinar a distribuição da variável Y = 3X, dada a distribuição de X da tabela:

x	1	3	5
p	0,4	0,1	0,5

Solução:

Como Y = 3X é monótona, a distintos valores de X correspondem distintos valores de Y. Assim:

y	3	9	15
q	0,4	0,1	0,5

Exemplo dois:

Determinar a distribuição da variável $Y = X^2$, se a distribuição de X é a da tabela:

x	-1	-2	1	2
p	0,3	0,1	0,2	0,4

Solução:

Como $Y = X^2$ não é monótona, a correspondência entre os valores de X e de Y não é biunívoca. Então, por definição, a probabilidade de cada y_i será igual a soma das probabilidades dos valores de X correspondendo a y_i, isto é:

$$P(Y = 1) = P(X = -1) + P(X = 1) =$$
$$= 0{,}3 + 0{,}2 = 0{,}5$$
$$P(Y = 4) = P(X = -2) + P(X = 2) =$$
$$= 0{,}1 + 0{,}4 = 0{,}5$$

y	1	4
p	0,5	0,5

02. Funções de duas variáveis aleatórias

O problema consiste em dada a distribuição de probabilidade conjunta de (X, Y), determinar qual é a distribuição de probabilidade de Z = H(X, Y).

Considere a variável Z = H(X, Y), uma função de duas variáveis aleatórias X e Y. Z = Z(s) é também uma variável aleatória, pois:

- Executa-se o experimento **E** obtendo "s";
- Calculam-se os números X(s) e Y(s);
- Calcula-se o número Z = H[X(s), Y(s)].

O valor de Z depende de "s", o resultado original do experimento. Ou seja, Z = Z(s) é uma função que associa um número real Z(s) a todo resultado $s \in S$, em consequência Z é uma variável aleatória.

Se (X, Y) for discreta, o problema é simples. Suponha que (X, Y) seja uma VADB. Então, as seguintes variáveis aleatórias unidimensionais são de interesse:

- X + Y;
- X - Y
- XY;
- X/Y;
- min(X, Y);
- max(X, Y);
- etc.

Exemplo:

Considere a distribuição conjunta da seguinte variável Bidimensional:

X \ Y	0	1	2	3	Σ
0	0,05	0,05	0,10	0,00	**0,20**
1	0,05	0,10	0,25	0,10	**0,50**
2	0,00	0,15	0,10	0,05	**0,30**
Σ	**0,10**	**0,30**	**0,45**	**0,15**	**1,00**

Determinar a distribuição da variável aleatória discreta U = min(X, Y).

Para obter a distribuição de probabilidade de U adota-se o seguinte procedimento:

- Determina-se os valores possíveis de U.

  Neste caso, são: 0, 1, 2.

- Para determinar P(U = 0) deve-se notar que P(U = 0) se e só se, um dos seguintes eventos ocorre: (X = 0, Y = 0) ou (X = 0, Y = 1) ou (X = 0, Y = 2) ou (X = 0, Y = 3) ou (X = 1, Y = 0) ou (X = 2, Y = 0).
- Assim: P(U = 0) = 0,05 + 0,05 + 0 + 0,05 + 0,10 + 0 = 25%.

Ilustração da determinação de P(U = 0) = = P[min (X, Y) = 0].

X \ Y	0	1	2	3	Soma
0	0,05	0,05	0,10	0,00	**0,20**
1	0,05	0,10	0,25	0,10	**0,50**
2	0,00	0,15	0,10	0,05	**0,30**
Soma	**0,10**	**0,30**	**0,45**	**0,15**	**1,00**

As outras probabilidades de U podem ser obtidas de forma semelhante, tendo-se:

P(U = 1) = 0,10 + 0,25 + 0,10 + 0,15 =

= 0,60 = 60%.

P(U = 2) = 0,10 + 0,05 = 0,15 = 15%.

Determinação de P(U = 1) =

= P[min (X, Y) = 1]:

X \ Y	0	1	2	3	Soma
0	0,05	0,05	0,10	0,00	**0,20**
1	0,05	0,10	0,25	0,10	**0,50**
2	0,00	0,15	0,10	0,05	**0,30**
Soma	**0,10**	**0,30**	**0,45**	**0,15**	**1,00**

Distribuição da VAD U = min(X, Y).

u	p(u)
0	0, 25
1	0,60
2	0,15
Σ	**1,00**

Exercício um:

Determinar a distribuição conjunta das variáveis aleatórias X = Número de vogais e Y = número de consoantes de uma palavra selecionada ao acaso da seguinte frase:

"O Grêmio é e sempre será o melhor time Gaúcho".

Solução um:

X \ Y	0	2	3	4	Σ
1	0,4	0,0	0,0	0,0	**0,4**
2	0,0	0,2	0,0	0,2	**0,4**
3	0,0	0,0	0,2	0,0	**0,2**
Σ	**0,4**	**0,2**	**0,2**	**0,2**	**1,0**

Exercício dois:

Com base na distribuição conjunta, determinar as distribuições das seguintes variáveis: XY, X + Y, X – Y, min(X, Y) e max(X, Y).

(x, y)	p(x, y)	xy	x+y	x – y	min	max
(1, 0)	0,4	0	1	1	0	1
(1, 2)	0,0	2	3	-1	1	2
(1, 3)	0,0	3	4	-2	1	3
(1, 4)	0,0	4	5	-3	1	4
(2, 0)	0,0	0	2	2	0	2
(2, 2)	0,2	4	4	0	2	2
(2, 3)	0,0	6	5	-1	2	3
(2, 4)	0,2	8	6	-2	2	4
(3, 0)	0,0	0	3	3	0	3
(3, 2)	0,0	6	5	1	2	3
(3, 3)	0,2	9	6	0	3	3
(3, 4)	0,0	12	7	-1	3	4

Solução:

xy	p(xy)
0	0,4
4	0,2
8	0,2
9	0,2
Σ	**1,00**

x + y	p(x + y)
1	0,4
4	0,2
6	0,4
Σ	**1,00**

x - y	p(x - y)
0	0,4
1	0,4
-2	0,2
Σ	**1,00**

min	p(min)
0	0,4
2	0,4
3	0,2
Σ	**1,00**

Distribuições

max	p(max)
1	0,4
2	0,2
3	0,2
4	0,2
Σ	**1,00**

x	p(x)
1	0,4
2	0,4
3	0,2
Σ	**1,00**

y	p(y)
0	0,4
2	0,2
3	0,2
4	0,2
Σ	**1,00**

Exercício três:

Calcular:

E(X) e V(Y)

E(Y) e V(Y)

E(X + Y) e V(X + Y)

E(X – Y) e V(X – Y)

E(XY) e V(XY)

Covariância Revisitada

Considerando os resultados anteriores, podemos redefinir a covariância como sendo:

Covariância Revisitada

COV(X,Y) =

$= E[(X – E(X)(Y – E(Y)] =$

$= E(XY – XE(Y) – YE(X) + E(X)E(Y)] =$

$= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) = \mathbf{E(XY) - E(X)E(Y)}$.

Definição:

1. Se COV(X, Y) = 0, então as variáveis X e Y são ditas **Não Correlacionadas**.
2. Duas variáveis aleatórias X e Y são ditas **independentes** se: **p(x, y) = p(x).p(y),** para todos os pares de valores (x, y).

Relações e consequências

$E(X \pm Y) = E(X) \pm E(Y)$

$V(X \pm Y) = V(X) + V(Y) \pm 2Cov(X, Y)$

$E(XY) \neq E(X)E(Y)$ (em geral)

Relações e consequências

Se X e Y são independentes então:

(i) COV(X, Y) = 0;

(ii) E(XY) = E(X).E(Y)

(iii) $V(X \pm Y) = V(X) + V(Y)$

Observações:

E(XY) = E(X)E(Y) pode ser verdadeira, mas as variáveis X e Y serem dependentes.

Se Cov(X, Y) = 0, isto não quer dizer que X e Y são independentes, o contrário sim.

A independência é uma relação mais forte do que a não correlação. Dizer que duas variáveis são não correlacionadas, significa que elas não tem relacionamento linear, enquanto que independência exclui qualquer tipo de relacionamento.

A Covariância

Relembrando:

E(aX + b) = aE(X) + b.

(i) Cov(aX + b, cY + d) = acCov(X, Y)

(ii) Cov(X + Y, Z + W) = Cov(X, Z) + Cov(X, W) + Cov(Y, Z) + Cov(Y, W).

O Coeficiente de Correlação

(i) $-1 \leq \rho \leq 1$;

(ii) $\rho(X + a, Y + b) = \rho(X, Y)$;

(iii) $\rho(aX, bY) = \rho(X, Y)$;

(iv) Se $\rho(X, Y) = 1$ ou $\rho(X, Y) = -1$, então $Y = aX + b$ com $a > 0$ ou $a < 0$ respectivamente.

Exemplo:

Considere a distribuição conjunta da tabela abaixo.

Y \ X	-1	0	1
-1	1/8	1/8	1/8
0	1/8	0	1/8
1	1/8	1/8	1/8

Verifique se X e Y são independentes.

Solução:

As distribuições marginais são:

X \ Y	-1	0	1	Σ
-1	1/8	1/8	1/8	3/8
0	1/8	0	1/8	2/8
1	1/8	1/8	1/8	3/8
Σ	3/8	2/8	3/8	1

X \ Y	-1	0	1	Σ
-1	1/8	1/8	1/8	3/8
0	1/8	0	1/8	2/8
1	1/8	1/8	1/8	3/8
Σ	3/8	2/8	3/8	1

As variáveis não são independentes, pois $p(1, 1) = 1/8 \neq p(1).q(1) = (3/8)(3/8) = 9/64$.

Exercício:

Uma urna contém três bolas numeradas: 1, 2 e 3. Duas destas bolas são retiradas ao caso e sejam X = valor da primeira bola retirada e Y = valor da segunda bola retirada.

Determine:

(a) E(XY)

(b) Cov(X, Y)

(c) Var (X + Y) se as retiradas forem:

(i) Com reposição;

(ii) Sem reposição.

Referências:

GRIMMETT, G. R., SITRZAKER, D. R. *Probability and Random Processes.* Oxford (London): Oxford University Press, 1991.

MEYER, Paul L. *Probabilidade: aplicações à estatística.* Rio de Janeiro: LTC, 1998.

MOOD, Alexander M., GRAYBILL, Franklin A. *Introduccion a la Teoria de la Estadistica.* Madrid: Aguilar, 1969.