Probabilidade II
Aula 2

Março de 2009

Mônica Barros, D.Sc.

Conteúdo

Independência

- Sejam X_1 e X_2 variáveis aleatórias contínuas ou discretas. Dizemos que X_1 e X_2 são **independentes** se sua densidade **conjunta** pode ser fatorada como o **produto das** respectivas densidades **marginais**. Isto é:

$$f(x_1, x_2) = f_1(x_1) f_2(x_2)$$

- Em particular, no caso de duas variáveis discretas, esta igualdade deve valer para todos os valores possíveis de ambas as variáveis, e portanto, para demonstrar a dependência entre duas variáveis, basta mostrar que a igualdade não se verifica para algum par de valores de x_1 e x_2.

Independência

- **Conseqüências**
- Se X_1 e X_2 são independentes então:

  1) As densidades condicionais são iguais às densidades marginais correspondentes.

  2) Os momentos (média, variância, etc ...) condicionais são iguais aos momentos correspondentes calculados a partir das densidades marginais.

  3) Se X_1 e X_2 são independentes com densidades marginais $f_1(x_1)$ e $f_2(x_2)$ então:

Independência

- **Conseqüências**

$Pr(a < X_1 < b, c < X_2 < d) = Pr(a < X_1 < b). Pr(c < X_2 < d)$ para quaisquer $a < b$, $c < d$.

4) Sejam X_1 e X_2 variáveis aleatórias independentes com densidades marginais $f_1(x_1)$ e $f_2(x_2)$. Então, se todos os valores esperados abaixo existirem:

$E[u(X_1)v(X_2)] = E[u(X_1)].E[v(X_2)]$

Independência

- **Conseqüências**

O valor esperado de um produto de funções de variáveis aleatórias onde cada função depende apenas de uma das variáveis aleatórias é igual ao produto dos valores esperados das funções individuais sempre que as variáveis aleatórias X_1 e X_2 forem independentes.

Independência

- **Demonstração – caso contínuo**

$$E\{u(X)v(Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x)v(y)f(x,y)dxdy$$

- **Como por hipótese X e Y são independentes, a densidade conjunta é o produto das densidades marginais de X e Y e então:**

$$E\{u(X)v(Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x)v(y)f_X(x)f_Y(y)dxdy =$$

$$= \left\{\int_{-\infty}^{\infty} u(x)f_X(x)dx\right\}\left\{\int_{-\infty}^{\infty} v(y)f_Y(y)dy\right\} = E(u(X)).E(v(Y))$$

Independência

- **Um caso particular importante ocorre quando u(.) e v(.) são as funções identidade, o que leva a:**

- **Se X e Y são independentes então:**
- **$E(X.Y) = E(X).E(Y) = \mu_X . \mu_Y$**

ENCE Escola Nacional de Ciências Estatísticas IBGE

Independência

- A noção de independência se estende de maneira imediata para um conjunto de n variáveis.

- Dizemos que X_1, X_2, ..., X_n são independentes se sua densidade (ou função de probabilidade) conjunta é o produto das densidades (funções de prob.) marginais.

ENCE Escola Nacional de Ciências Estatísticas IBGE

Independência

- E quais são alguns dos efeitos da independência?
  - Condicionais são iguais as marginais;
  - Valores esperados de produtos de funções de cada um dos X_i ´s individualmente são iguais aos produtos dos respectivos valores esperados.

ENCE Escola Nacional de Ciências Estatísticas IBGE

Exemplo 1

- Sejam X e Y variáveis discretas com função de probabilidade conjunta do Exemplo 1 da aula passada:

Pr(X = x) →	X = 0	X = 1	X = 2	X = 3
Pr(Y = y) ↓				
Y = 0	0.2	0.1	0.1	0
Y = 1	0.1	0.1	0.1	0.1
Y =2	0	0.1	0.1	0

- X e Y são independentes? Por que?

  Por exemplo,

  Pr(X=0) = 0.2 + 0.1 + 0 = 0.3 e

  Pr(Y = 0) = 0.2 + 0.1 + 0.1 + 0 = 0.4

  Mas, Pr(X = 0, Y = 0) = 0.2 ≠ (0.3)(0.4).

  Logo, X e Y não são independentes - (são dependentes)

ENCE Escola Nacional de Ciências Estatísticas IBGE

Covariância

- A covariância entre as variáveis X e Y é definida como:

$$COV(X,Y) = E[(X-\mu_X).(Y-\mu_Y)] = E[X.Y - \mu_X.Y - \mu_Y.X + \mu_X.\mu_Y] =$$
$$= [E(X.Y) - \mu_X.E(Y) - \mu_Y.E(X) + \mu_X.\mu_Y] =$$
$$= E(X.Y) - \mu_X.\mu_Y$$

- Onde μ_X = E(X) , μ_Y = E(Y) são as médias de X e Y, computadas a partir das densidades marginais.
- Note que a covariância é um valor esperado que envolve simultaneamente as duas variáveis, por isso ela deve ser calculada a partir da densidade conjunta de X e Y.

Covariância

- Da expressão anterior pode-se notar que a variância é apenas um caso particular da covariância. Por exemplo, fazendo Y = X na última expressão leva a:

$$COV(X,X) = E[(X-\mu_X).(X-\mu_X)] = E\left[(X-\mu_X)^2\right] = VAR(X)$$

- Obviamente poderíamos ter derivado a variância de Y pelo mesmo procedimento.

Covariância

- A covariância é expressa em unidades, no mínimo, "estranhas".

- A unidade de COV(X, Y) é o produto das unidades de X e de Y. Então, se X é dado em metros e Y em litros, a covariância é medida em metros*litros, o que não explica muita coisa, não é mesmo?

Covariância

- O ideal seria construir uma medida, a partir da covariância, que fosse independente de unidades, ou seja, adimensional.

- Esta medida é o coeficiente de correlação, a ser definido daqui a pouco.

Covariância

- Um outro ponto importante a respeito da covariância é: COV(X, Y) não nos diz nada a respeito de $E(X^r.Y^s)$ onde r, s ≠ 1.

- Em outras palavras: a covariância só mede a associação linear entre X e Y. Não diz nada, por exemplo, a respeito de $E(X^2.Y^5)$.

Covariância

- Já a independência é uma condição muito mais forte.
- Pelo resultado da página 5 (vide demonstração no slide 7), a independência implica não apenas em covariância zero mas também em $E(X^r.Y^s) = E(X^r)E(Y^s)$ para todo r, s para os quais os valores esperados existem.

Covariância

- Vamos deduzir algumas propriedades da covariância de funções lineares de X e Y.
- Para isso é necessário relembrar quem são a média e a variância de uma função linear de uma variável aleatória.

Covariância

- Se X é uma variável aleatória qualquer com média μ_X e variância σ_X^2 e <u>a</u> e <u>b</u> são constantes então:
- $E(aX+b) = a.\ E(X) + b = a.\ \mu_X + b$
- $VAR\ E(aX+b) = a^2.\ VAR(X) = a^2.\ \sigma_X^2$

Covariância

- $COV(aX+b,\ cY+d) =$

$= E\{(aX+b)(cY+d)\} - (a\mu_X + b)(c\mu_Y + d) =$

$= a.c.E(XY) + a.d.E(X) + b.c\ E(Y) + b.d$

$-\ \ a.c.\ \mu_X.\ \mu_Y - a.d.\ \mu_X - b.c.\ \mu_Y - b.d$

$= a.c\{E(XY) - \mu_X.\ \mu_Y\} + ad\{\mu_X - \mu_X\} + bc\{\mu_Y - \mu_Y\} =$

$= a.c.COV(X,Y)$

Coeficiente de Correlação

- **O coeficiente de correlação entre X e Y, denotado por ρ, é definido como:**

$$\rho = \frac{COV(X,Y)}{\sqrt{VAR(X)}.\sqrt{VAR(Y)}}$$

- **O coeficiente de correlação é a covariância dividida pelo produto dos desvios padrões das duas variáveis, ou seja, é uma medida padronizada (e adimensional) da covariância.**

Coeficiente de Correlação

- **Se X e Y são independentes, então ρ = 0 (X e Y são descorrelatados).**

- **A recíproca deste fato NÃO É verdadeira, ou seja, CORRELAÇÃO ZERO NÃO IMPLICA EM INDEPENDÊNCIA.**

- **A única instância em que as duas condições são eqüivalentes (correlação zero e independência) é no caso de variáveis Normais.**

Coeficiente de Correlação

- **O coeficiente de correlação entre X e Y é um valor entre -1 e +1. Além disso , ρ = +1 ou ρ = -1 se e somente se, X e Y é uma função linear de X.**

$$\rho = +1 \Leftrightarrow Y = aX + b \ \text{ onde } a > 0$$

$$\rho = -1 \Leftrightarrow Y = aX + b \ \text{ onde } a < 0$$

- **Mas, é importante notar que o coeficiente de correlação é uma medida de associação LINEAR entre as variáveis, ou seja, não diz nada sobre a relação entre X^j e Y^k onde j e k são diferentes de 1.**

Coeficiente de Correlação

- **Resultado**
- **Se Y = aX+b então o coeficiente de correlação entre X e Y é + 1 se a > 0 e -1 se a < 0.**
- **Demonstração**
- **Pelas propriedades da covariância de funções lineares (vide slide 20):**

$$COV(X,Y) = COV(X, aX + b) = a.COV(X,X) = a.VAR(X)$$

Coeficiente de Correlação

- A variância de Y é:

$$VAR(Y) = a^2.VAR(X)$$

- O desvio padrão de Y é:
  dp(Y) = |a|.dp(X)
- Assim, o coeficiente de correlação entre X e Y é:

$$\rho = \frac{COV(X,Y)}{dp(X)dp(Y)} = \frac{a.VAR(X)}{dp(X)|a|dp(X)} = \frac{a}{|a|}\frac{VAR(X)}{VAR(X)} = \frac{a}{|a|}$$

Coeficiente de Correlação

- Se a > 0 então |a| = a e assim ρ = +1
- Se a < 0 então |a| = -a e assim ρ = -1
- Resta provar apenas que os valores -1 e +1 são os limites mínimo e máximo do coeficiente de correlação. Isso pode ser feito através da desigualdade de Schwarz.

Desigualdade de Schwarz

- Será usada aqui para provar que o coeficiente de correlação é sempre um número no intervalo [-1, +1]
- Suponha que X e Y tenham 2os momentos finitos. Então E(|X.Y|) < ∞ e:

$$|E(X.Y)| \leq \sqrt{E\left(X^2\right)E\left(Y^2\right)}$$

- E a igualdade ocorre se, e somente se, uma das variáveis é um múltiplo da outra.

Desigualdade de Schwarz

- Demonstração
- Para todo número t real, X + t.Y tem um 2º. Momento finito. Considere a função:
- $g(t) = E\{ (X + tY)^2 \}$
- Esta função é finita. Podemos expandir o quadrado acima para obter:
- $g(t) = E\{ X^2 + 2.t.X.Y + t^2.Y^2 \} =$
  $= E(X^2) = 2.t.E(X).E(Y) + t^2.E(Y^2)$

ENCE Escola Nacional de Ciências Estatísticas IBGE

Desigualdade de Schwarz

- Note que g(t) é uma função quadrática em t e portanto é sempre não negativa.
- Se E(Y²) = 0 então Y = 0 com probabilidade 1 (pois do contrário a variância de Y seria negativa) e assim E(XY) = E(X.0) = 0 e a igualdade é válida na expressão do slide 27.
- Suponha então que E(Y²) > 0.

ENCE Escola Nacional de Ciências Estatísticas IBGE

Desigualdade de Schwarz

- Qual o ponto t_0 que minimiza g(t)?
- Basta resolver a condição de primeira ordem e verificar o sinal da 2ª. derivada.

$$\frac{dg(t)}{dt} = 0 \Leftrightarrow 2E(XY) + 2tE(Y^2) = 0 \Rightarrow t_0 = -\frac{E(XY)}{E(Y^2)}$$

- A segunda derivada é:

$$\frac{d^2g(t)}{dt^2} = 2.E(Y^2) > 0$$

- Indicando um ponto de mínimo.

ENCE Escola Nacional de Ciências Estatísticas IBGE

Desigualdade de Schwarz

- Avaliando a função g(.) em seu mínimo leva a:

$$g(t_0) = E(X^2) - 2\frac{E(XY)}{E(Y^2)}E(XY) + \left(-\frac{E(XY)}{E(Y^2)}\right)^2 E(Y^2) \geq 0$$

$$E(X^2) - 2\frac{(E(XY))^2}{E(Y^2)} + \frac{(E(XY))^2}{E(Y^2)} \geq 0$$

$$E(X^2) - \frac{(E(XY))^2}{E(Y^2)} \geq 0$$

$$E(X^2).E(Y^2) \geq (E(XY))^2$$

$$(E(XY))^2 \leq E(X^2).E(Y^2)$$

$$\sqrt{(E(XY))^2} \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}$$

$$|E(XY)| \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}$$

ENCE Escola Nacional de Ciências Estatísticas IBGE

Desigualdade de Schwarz

- Suponha agora que X e Y são variáveis com média zero, ou seja, X = W – μ_W e Y = Z – μ_Z.
- Aplicando a desigualdade de Schwarz nestas condições leva a:

$$\{E[(W-\mu_W)(Z-\mu_Z)]\}^2 \leq \{E(W-\mu_W)^2\}\{E(Z-\mu_Z)^2\}$$

$$\{COV(W,Z)\}^2 \leq VAR(W).VAR(Z)$$

$$\left\{\frac{COV(W,Z)}{dp(W)dp(Z)}\right\}^2 \leq 1 \Leftrightarrow \rho \leq 1$$

O que demonstra os limites para o coeficiente de correlação para quaisquer v.a.

Exemplo – duas variáveis descorrelatadas que não são independentes

- Considere as variáveis discretas X e Y cuja função de probabilidade conjunta é dada a seguir. Esta função de probabilidade corresponde a uma uniforme discreta no plano definida nos pontos (1,0), (0,1), (-1, 0) e (0,-1).

X→, Y↓	-1	0	1
-1	0	0,25	0
0	0,25	0	0,25
1	0	0,25	0

Exemplo

- Da tabela anterior segue que:
- Pr(X = -1) = Pr(X = +1) = 0,25 e Pr(X=0) = ½
- Então E(X) = 0(1/2) + 1(1/4) + (-1)(1/4) = 0

- De maneira análoga E(Y) = 0
- Também:
- Pr(XY = 0) = 1 (por que?)
- Então E(XY) = 0

Exemplo

- Se você não se convenceu dos dois últimos passos, use a "força bruta" e resolva E(XY) pela função de probabilidade conjunta.

- Logo, COV(X,Y) = E(XY) – E(X)E(Y) = 0
- Mas, X e Y são dependentes. Por exemplo, basta notar que Pr(X = 0).Pr(Y = 0) =(1/2)(1/2) = ¼ ≠ Pr(X = 0, Y = 0) = 0.