MPEI 2015-2016

Aula 11

Distribuições Condicionais

Distribuições condicionais

- As distribuições e densidades de probabilidade condicionais são definidas através da definição de prob. Condicional de eventos
- Seja, por exemplo, $\{X \leq x\}$ um evento e $\{Y \leq y\}$ outro
- $F_{X|Y\leq y} =$
- $= \frac{P(X\leq x, Y\leq y)}{P(Y\leq y)}$
- $= \frac{F_{X,Y}(x,y)}{F_Y(y)}$
- $= \frac{F_{X,Y}(x,y)}{F_{X,Y}(\infty,y)}$
  - ou seja função distribuição acumulada conjunta a dividir pela f. distribuição marginal
- É uma função de x, para qualquer y fixo

Função densidade de probabilidade condicionada

- Para qualquer y fixo, pode-se diferenciar relativamente a x,
  - obtendo-se a função densidade de probabilidade da variável aleatória X condicionada ao evento $\{Y \leq y\}$

$$f_{X|Y\leq y} = \frac{\partial F_{X|Y\leq y}(x,y)}{\partial y}$$

Funções de probabilidade condicional

- A probabilidade da variável X num acontecimento A **conhecendo um valor exacto y que a variável Y assume** pode ser determinado usando a definição
- $P(X\ em\ A\ |\ Y = y) = \frac{P(X\ em\ A\ , Y = y)}{P(Y = y)}$
  - Cujo cálculo depende muito do tipo da variável Y
    - Se Y for contínua $P(Y = y) = 0$ ☹
  - Daremos atenção aos casos com Y discreto

Condicionada numa v. a. discreta

- Y discreta, $Y = y_k$:
  - $F_X(x|y_k) = \frac{P(X \leq x, Y=y_k)}{P(Y=y_k)}$, $P(Y = y_k) \neq 0$

- Se $F_X(x|y_k)$ for derivável, a função de densidade condicional será
- $f(x|y_k) = \frac{d}{dx} F_X(x|y_k)$

Ambas as v. a. discretas

- Se ambas variáveis forem discretas teremos
- $p_X(x_i|y_k) = P(X = x_i|Y = y_k)$
- $= \frac{P(X=x_i,Y=y_k)}{P(Y=y_k)} = \frac{p_{XY}(x_i,y_k)}{p_Y(y_k)}$,
  - para y_k para os quais $p_Y(y_k) > 0$
  - Define-se $p_X(x_i|y_k) = 0$ quando $p_Y(y_k) = 0$

- De forma similar teremos $p_Y(y_k|x_i) = \frac{p_{XY}(x_i,y_k)}{p_X(x_i)}$
- A probabilidade de qualquer evento A dado $Y = y_k$ é dada por:
- $P[\, X \; em \; A \mid Y = y_k] = \sum_{x_i em\ A} \quad p_Y(x_i|y_k)$

...

- De $p_X(x_i|y_k) = \frac{p_{XY}(x_i,y_k)}{p_Y(y_k)}$ obtém-se:
- $p_{XY}(x_i, y_k) = p_{X|Y}(x_i|y_k) \quad p_Y(y_k)$

- Somando ambos os lados para todos os valores de Y obtemos:
- $p_X(x) = \sum_y p_{XY}(x, y\ ) = \sum_y p_{X|Y}(x|y) p_Y(y)$
- Que é uma **generalização da Lei da Probabilidade total**

- Nota: se X e Y independentes $p_{X|Y}(x|y) = p_X(x)$.

Exemplo simples

- Consideremos X e Y como resultados do lançamento de 2 dados honestos
- Qual o valor de $p_X(6|5)$?
- Como sabemos:
  - $p_{XY}(x_i, y_k) = \frac{1}{36}$, para todos os pares i,k
  - $p_Y(y_k) = \frac{1}{6}$

- Resolução:
- = $\frac{p_{XY}(x_i,y_k)}{p_Y(y_k)} = \frac{p_{XY}(6,5)}{p_Y(5)} = \ \ldots = 1/6$

Exemplo 2

- Considere I e P como variáveis aleatórias relativas ao tempo em Inglaterra e Portugal, respectivamente, e a distribuição conjunta da tabela.
- Qual o valor de P(I="Chuva" | P="Sol")?
- = $p_I("Chuva"|"Sol")$
- $= \frac{p_{IP}("Chuva","Sol")}{p_P("Sol")}$
- $= \frac{0{,}05}{0{,}7}$
- Qual o valor de P(P="Sol" | I="Chuva")?

I \ P →	SOL	CHUVA
SOL	0,5	0,2
CHUVA	0,05	0,05
NEBULOSIDADE	0,15	0,05

- Se Y for contínua surge o problema do denominador ser 0
- A definição da função de distribuição neste caso será:

$$F_Y(x|y) = \lim_{h \to 0} F_Y\,(x|\, \boldsymbol{y} < \boldsymbol{Y} \leq \boldsymbol{y} + \boldsymbol{h})$$

- Sendo $F_Y(x|y)$ derivável, demonstra-que:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

função densidade de probabilidade condicionada a Y=y

Funções de variáveis aleatórias

E respectivas expectâncias

Funções de v. a. múltiplas

- Muitas vezes encontramos **problemas em que temos uma transformação das v. a.** $X_1, X_2, \ldots, X_n$ que produz variáveis aleatórias $Y_1, Y_2, \ldots, Y_m$
- O caso mais simples é termos uma função escalar de várias variáveis aleatórias

$$Z = g(X_1, X_2, \ldots, X_n)$$

A função de distribuição acumulada de Z é determinada como sabemos calculando a probabilidade do conjunto $\{Z \leq z\}$

i.e. A região R_z do espaço n-dimensional tal que

$R_z = \{\mathbf{x}: g(\mathbf{x}) \leq \mathbf{z}\}$ $\qquad\qquad$ $\mathbf{x} = (\boldsymbol{x_1}\ \boldsymbol{x_2}\ \ldots\ \boldsymbol{x_n})$

- Logo
  - $F_Z(z) = \int_{\mathrm{x} \in R_Z} \ldots \int f_X(\mathrm{x})\, d\mathrm{x}$

Expectância de funções de v. aleatórias

- Expectância de uma função de 2 var. aleatórias

  $Z = g(X, Y)$

  $$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, dx\, dy$$

- Para o caso de variáveis discretas :
  - $E[Z] = \sum_i \sum_n g(x_i, y_n)\, p_{X,Y}(x_i, y_n)$

- Resultado generalizável para uma função com um número arbitrário de variáveis aleatórias

  $Z = g(\mathbf{X})$ em que $\mathbf{X}$ (bold) é um vector

  $$E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

Exemplo

- $Z = g(X, Y) = X + Y$ *i.e.* Soma de 2 v.a.
- $E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f_{X,Y}(x, y)\, dx\, dy$
- $= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f_{X,Y}(x, y)\, dx\, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{X,Y}(x, y)\, dx\, dy$
- $= \int_{-\infty}^{\infty} x\, f_X(x)\, dx + \int_{-\infty}^{\infty} y f_Y(y)\, dy$
- $= \boldsymbol{E(X) + E(Y)}$
- Resultado válido quer as variáveis **sejam independentes ou não**
- Mostra (conjuntamente com a propriedade de escala, $E(aX) = aE(X)$) que a expectância é um **operador linear**

Momentos de funções de variáveis aleatórias

- Momento de ordem n de uma função escalar de um vector aleatório:

  $Z = g(\mathbf{X})$ em que $\mathbf{X}$ (bold) é um vector

  $$E[\boldsymbol{Z^n}] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \boldsymbol{g^n(\mathbf{x})} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

- Aplicando para obter a variância temos:
- $\mathrm{Var}[Z] = E[Z^2] - E^2[Z]$
- $= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \boldsymbol{g^2(\mathbf{x})} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} - \left(\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \boldsymbol{g(\mathbf{x})} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}\right)^{\mathbf{2}}$

Expectância condicional

- A expectância condicional de X dado $\mathrm{Y} = y$ é definida por

$$E\left[X \mid y\right] = \int_{-\infty}^{\infty} x f_X(X \mid y)\,dx\,.$$

- Se X e Y forem ambas discretas temos

$$E\left[X \mid y\right] = \sum_{x_j} x p_X(x_j \mid y).$$

Soma e combinação linear de variáveis aleatórias

É um caso particular das funções de v.a. com muito interesse

Motivação

- Se somarmos duas variáveis aleatórias X_1 e X_2 quais as características da variável aleatória $S = X_1 + X_2$?
  - Em termos de momentos ?
  - Em termos de função de distribuição ?

- E no caso geral $S_n = X_1 + X_2 + \cdots + X_n$?

Média da soma de n variáveis

- Sejam $X_1, X_2, \dots, X_n$, n variáveis aleatórias e $S_n = X_1 + X_2 + \cdots + X_n$ a sua soma
- Teorema: **A média da soma de n variáveis é igual à soma das médias**
- Demonstração

$$\begin{aligned} E[S_n] &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (\sum_{j=1}^{n} x_j) f_{X_1 \dots X_n}(x_1, \dots, x_n) dx_1 \dots dx_n = \\ &= \sum_{j=1}^{n} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_j f_{X_1 \dots X_n}(x_1, \dots, x_n) dx_1 \dots dx_n = \\ &= \sum_{j=1}^{n} \int_{-\infty}^{+\infty} x_j f_{X_j}(x_j) dx_j = \sum_{j=1}^{n} E[X_j] \end{aligned}$$

Variância da soma de n variáveis

- Considerando da mesma forma $S_n = X_1 + X_2 + \cdots + X_n$:
- Teorema: **A variância da soma de n variáveis aleatórias é dada pela soma de todas as variâncias e covariâncias**

$$Var(S_n) = \sum_{i=1}^{n} Var(X_i) + \sum_{\substack{j=1 \\ j \neq k}}^{n} \sum_{k=1}^{n} Cov(X_j, X_k)$$

- Demonstração:

$$\begin{aligned} Var\left(S_n\right) &= E\left[\sum_{j=1}^{n}\left(X_j - E\left[X_j\right]\right)\sum_{k=1}^{n}\left(X_k - E\left[X_k\right]\right)\right] = \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} E\left[\left(X_j - E\left[X_j\right]\right)\right] E\left[\left(X_k - E\left[X_k\right]\right)\right] \end{aligned}$$

...

- Se as variáveis **são independentes**, $Cov(X_j, X_k) = 0$, para todo o $j \neq k$, pelo que:

- $\boldsymbol{Var(S_n) = \sum_{i=1}^{n} Var(X_i)}$
  - **Variância da soma igual a soma das variâncias**

- Se para além de independentes forem **identicamente distribuídas (IID)**
  e tivermos $E[X_i] = \mu$ e $Var(X_i) = \sigma^2, i = 1,2, \dots, n$
  a média e variância da soma são dadas por:

- $\mathrm{E}[S_n] = n\,\mu$ e $Var(S_n) = n\,\sigma^2$

Função de distribuição da soma de 2 variáveis aleatórias independentes

- Caso discreto (2 v.a. Discretas X e Y)
- Fazendo $\mathrm{Z} = X + Y$
- $p_Z(z) = P(X + Y = z)$

$=\sum_{\{(x,y)|x+y=z\}} P(X = x, Y = y)$

$$= \sum_{x} P(X = x, Y = z - x)$$

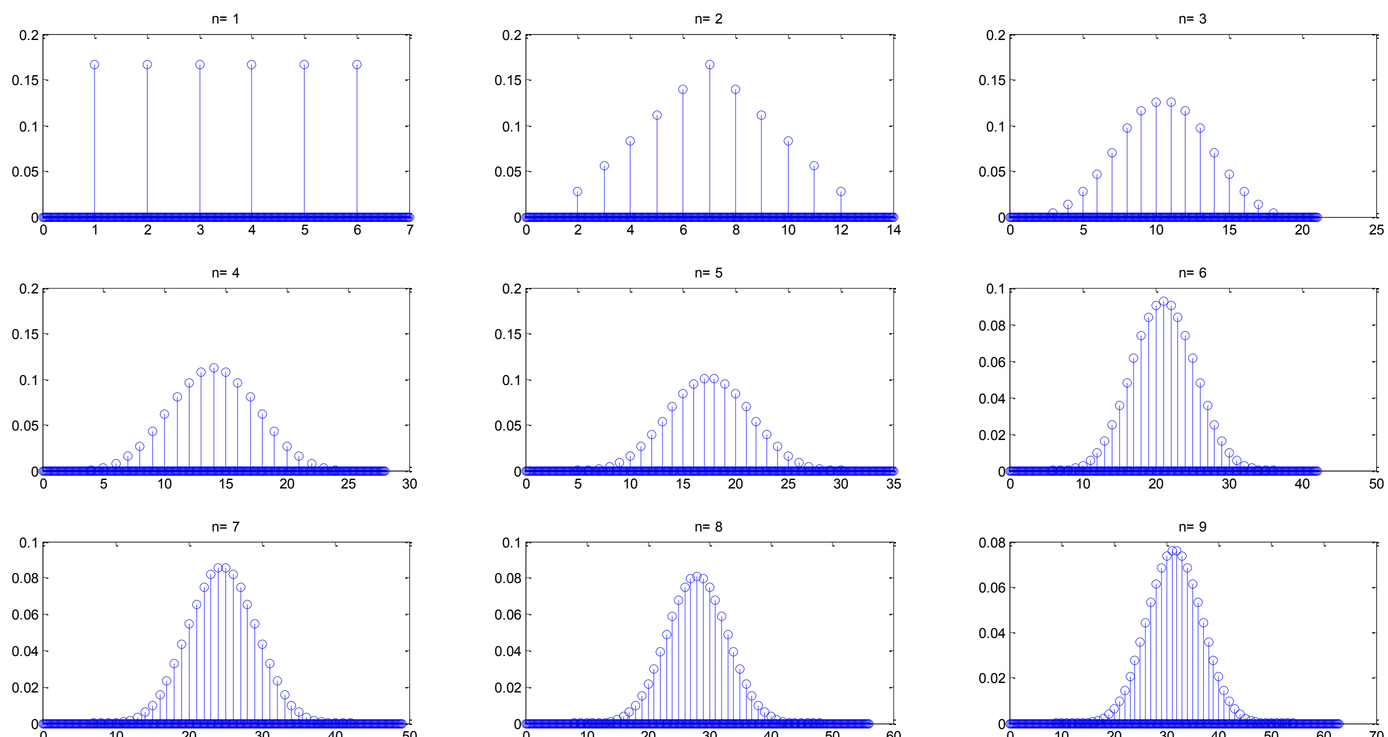
$= \sum_x p_X(x) \quad p_Y(z - x)$; devido à indep.

$= p_X\ (x)\ *\ p_Y(z)$

- Que é a convolução discreta de p_X e p_Y

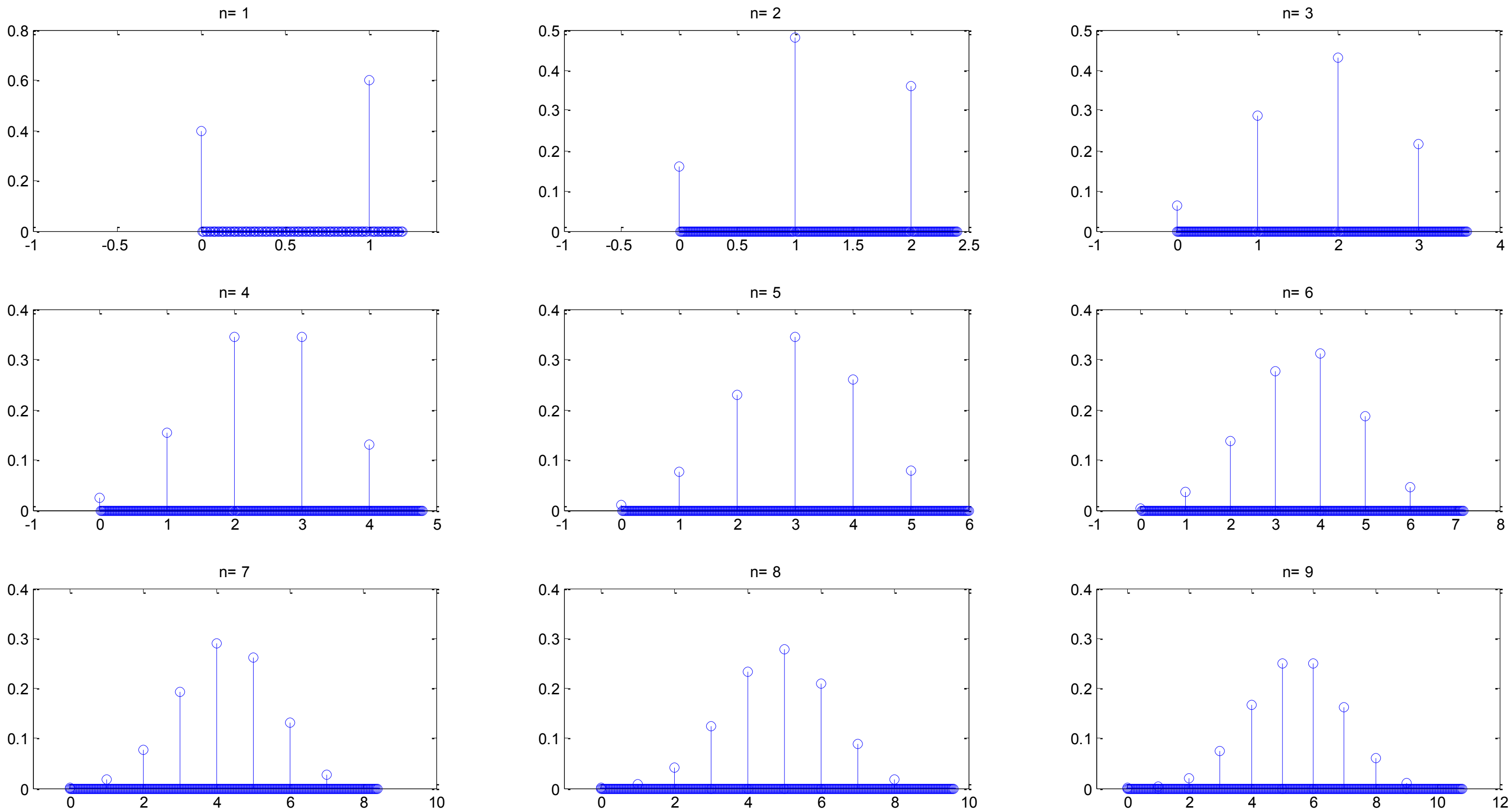
Exemplo (em Matlab)

- Usando conv() e a pmf relativa à variável X correspondente ao lançamento de um dado honesto (n=1, 2, ..., 9)



Outro exemplo

- Sendo X relativa ao número de caras num lançamento de moeda não honesta
  - com probabilidade de cara = 0,6



Caso contínuo

- Sendo X e Y independentes e contínuas
- Fazendo novamente $\mathrm{Z} = X + Y$
- Para obter a função densidade prob. de Z , primeiro obtém-se a f.d.p conjunta de X e Z e depois integra-se

$$
\begin{aligned}
F_{Z|X}(z \mid x) &= \mathbf{P}(Z \le z \mid X = x) = \mathbf{P}(X + Y \le z \mid X = x) = \mathbf{P}(x + Y \le z \mid X = x) = \mathbf{P}(Y \le z - x \mid X = x) \\
&= \mathbf{P}(Y \le z - x) \qquad \text{using the independence of } X \text{ and } Y \\
&= F_Y(z - x)
\end{aligned}
$$

$$
f_{Z|X}(z \mid x) = \frac{d}{dz} F_{Z|X}(z \mid x) = \frac{d}{dz} F_Y(z - x) = f_Y(z - x)
$$

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_{X,Z}(x,z)\,dx = \int_{-\infty}^{\infty} f_X(x) f_{Z|X}(z \mid x)\,dx \\
&= \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,dx \equiv f_X * f_Y(z)
\end{aligned}
$$

Obtém-se através da convolução, agora contínua

- Para mais informação, ver, por exemplo:
- https://engineering.purdue.edu/~ipollak/ece302/SPRING12/notes/23_GeneralRVs-8_Sums.pdf

Combinações lineares de variáveis aleatórias

- Os **resultados anteriores generalizam-se** facilmente para o caso de termos uma soma pesada (combinação linear) $Y_n = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$
  - Em que $c_1, c_2, \ldots, c_n$ são constantes

- $E[Y_n] = c_1 E[X_1] + c_2 E[X_2] + \cdots + c_n E[X_n]$
- $Var(Y_n) = \sum_{i=1}^{n} c_i{}^2 Var(X_i) + \sum_i \sum_{j\ (\neq i)} c_i c_j Cov(X_i, X_j)$

- Se independentes $Var(Y_n) = \sum_{i=1}^{n} c_i{}^2 Var(X_i)$

Exemplo de aplicação

- Um semicondutor tem 3 camadas.
- Se a espessura das várias camadas tiver uma variância de 25, 40 e 30 nanómetros quadrados, qual a variância da espessura das 3 camadas ?

- Considerando X_1, X_2, X_3 como as espessuras das camadas e independência
- $Var(X_1 + X_2 + X_3) =$
  $Var(X_1) + var(X_2) + var(X_3) = 95$
- Mostrando a propagação das variâncias de todas as camadas para o resultado final
  – Os problemas somam-se ☹