Revisão de Probabilidade e Estatística

Aula 3 - Parte 1

Aishameriane Schmidt

PPGECO/UFSC

Março de 2018.

Programa

Aula 3

1. Algumas pendências das aulas anteriores

Programa

Aula 3

1. Algumas pendências das aulas anteriores
2. Continuação de variáveis aleatórias conhecidas: Poisson, Exponencial, Normal, Gamma, Beta

Programa

Aula 3

1. Algumas pendências das aulas anteriores
2. Continuação de variáveis aleatórias conhecidas: Poisson, Exponencial, Normal, Gamma, Beta
3. Distribuições conjuntas e condicionais
    - Esperança condicional
    - ~~Teorema de Bayes para variáveis aleatórias~~
4. Função geradora de momentos

Programa

Aula 3

1. Algumas pendências das aulas anteriores
2. Continuação de variáveis aleatórias conhecidas: Poisson, Exponencial, Normal, Gamma, Beta
3. Distribuições conjuntas e condicionais
    - Esperança condicional
    - ~~Teorema de Bayes para variáveis aleatórias~~
4. Função geradora de momentos
5. ~~Distribuições multivariadas~~[a]
    - ~~Multinomial~~
    - ~~A normal multivariada~~

[a] *Isto não está com distribuições conjuntas porque eu acho que vai ser melhor saber função geradora de momentos antes.*

Pendências

Quando não há elemento de A que satisfaça P, o conjunto B não tem nenhum elemento e é denominado *conjunto vazio* ($\emptyset$).

Definição

Definimos $\emptyset$ da seguinte forma:

$$\forall x, x \neq \emptyset \qquad \textit{ERRADO}$$

Pendências

Quando não há elemento de A que satisfaça P, o conjunto B não tem nenhum elemento e é denominado *conjunto vazio* ($\emptyset$).

Definição

Definimos $\emptyset$ da seguinte forma:

$$\forall x, x \neq \emptyset \qquad \textit{ERRADO}$$

$$\forall x, x \notin \emptyset$$

(lê-se: qualquer que seja x, x não percence ao vazio)

Pendências

Teorema

Lei do Estatístico Inconsciente

$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x) \cdot p_X(x)$

Para quais funções $g(\cdot)$ isso é válido?

Pendências

Teorema

Lei do Estatístico Inconsciente

$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x) \cdot p_X(x)$

Para quais funções $g(\cdot)$ isso é válido?

Definição

Composição de Funções [Lima, 1982]

Sejam $f : A \to B$ e $g : C \to D$ funções tais que a imagem de f está contido na imagem de g, i.e., $B \subset C$. Neste caso, podemos definir a função composta $g \circ f : A \to D$, que consiste em aplicar primeiro f e depois g.

Pendências

Teorema

Lei do Estatístico Inconsciente

$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{X}} g(x) \cdot p_X(x)$

Para quais funções $g(\cdot)$ isso é válido?

Definição

Composição de Funções [Lima, 1982]

Sejam $f : A \to B$ e $g : C \to D$ funções tais que a imagem de f está contido na imagem de g, i.e., $B \subset C$. Neste caso, podemos definir a função composta $g \circ f : A \to D$, que consiste em aplicar primeiro f e depois g.

Conclusão: Quando usamos a composta no teorema já estamos implicitamente assumindo que $R(X) \subset Dom(g)$!

Pendências

Lema

A probabilidade de nada ocorrer é zero

$$\mathbb{P}(\emptyset) = 0$$

Pendências

Lema

A probabilidade de nada ocorrer é zero

$$\mathbb{P}(\emptyset) = 0$$

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{n>1} \mathbb{P}(\emptyset)$$

$$\mathbb{P}(\Omega) - \mathbb{P}(\Omega) = \sum_{n>1} \mathbb{P}(\emptyset)$$

Pendências

Lema

A probabilidade de nada ocorrer é zero

$$\mathbb{P}(\emptyset) = 0$$

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{n>1} \mathbb{P}(\emptyset)$$

$$\mathbb{P}(\Omega) - \mathbb{P}(\Omega) = \sum_{n>1} \mathbb{P}(\emptyset)$$

$$0 = \sum_{n>1} \mathbb{P}(\emptyset) = \lim_{k\to\infty} \sum_{n=1}^{k} \mathbb{P}(\emptyset)$$

$$0 = \lim_{k\to\infty} \mathbb{P}(\emptyset) \cdot k$$

Se $\mathbb{P}(\emptyset)$ for diferente de 0, então $\lim_{k\to\infty} \mathbb{P}(\emptyset) \cdot k$ é diferente de zero, uma contradição!

Logo, $P(\emptyset) = 0$.

Pendências

Lema

A probabilidade de nada ocorrer é zero

$$\mathbb{P}(\emptyset) = 0$$

$$\mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \sum_{n>1} \mathbb{P}(\emptyset)$$

$$\mathbb{P}(\Omega) - \mathbb{P}(\Omega) = \sum_{n>1} \mathbb{P}(\emptyset)$$

$$0 = \sum_{n>1} \mathbb{P}(\emptyset) = \lim_{k\to\infty} \sum_{n=1}^{k} \mathbb{P}(\emptyset)$$

$$0 = \lim_{k\to\infty} \mathbb{P}(\emptyset) \cdot k$$

Se $\mathbb{P}(\emptyset)$ for diferente de 0, então $\lim_{k\to\infty} \mathbb{P}(\emptyset) \cdot k$ é diferente de zero, uma contradição!

Logo, $P(\emptyset) = 0$.

$$0 = \sum_{n>1} \mathbb{P}(\emptyset)$$

Pelo critério do termo geral, sabemos que como a série converge, o limite quando $n \to \infty$ do termo geral será igual a 0.

Uma vez que o TG é constante, $P(\emptyset) = 0$.

Pendências

Teorema

Probabilidade condicional é uma probabilidade

$\mathbb{P}[\cdot|B]$ *é uma função de probabilidade! (ela satisfaz os axiomas de Kolmogorov)*

A demonstração está nas notas de aula (assim como outros teoremas que decorrem desse fato).

Pendências

Seja $\lambda > 0$ e X uma v.a. com f.d.p. dada por:

$$f(x) = \begin{cases} \lambda^{-1} e^{-\frac{x}{\lambda}} & \text{se } x > 0. \\ 0 & \text{caso contrário.} \end{cases} \tag{1}$$

Teorema

- *f_X é de fato uma densidade (integra 1 e é sempre positiva);*
- *$\mathbb{E}[X] = \lambda$.*

Observação 1: λ é o parâmetro da distribuição e é chamado de *parâmetro de escala*. Se X é a duração que um determinado sistema funciona, então λ unidades de tempo é o tempo esperado de duração.

Pendências

Seja $\lambda > 0$ e X uma v.a. com f.d.p. dada por:

$$f(x) = \begin{cases} \lambda^{-1} e^{-\frac{x}{\lambda}} & \text{se } x > 0. \\ 0 & \text{caso contrário.} \end{cases} \tag{1}$$

Teorema

- *f_X é de fato uma densidade (integra 1 e é sempre positiva);*
- $\mathbb{E}[X] = \lambda$.

Observação 1: λ é o parâmetro da distribuição e é chamado de *parâmetro de escala*. Se X é a duração que um determinado sistema funciona, então λ unidades de tempo é o tempo esperado de duração.
Observação 2: Se definirmos $\beta = \lambda^{-1}$, então $\mathbb{E}[X] = \beta$. Neste caso, β é chamado de *taxa*. Se X for um determinado evento que ocorre a uma taxa β, então o tempo entre dois eventos tem média β.

Distribuição Exponencial

Densidade Acumulada

Teorema

Se $X \sim Exp(\lambda)$, então sua f.d.a. é dada por

$$F_X(x) = \mathbb{P}(X \leq x) = 1 - e^{-\frac{x}{\lambda}}$$

- **Observação:** Note que a f.d.a. nos diz a probabilidade de X ser menor ou igual que um valor qualquer, de maneira que seu complementar é a probabilidade de X ultrapassar um determinado valor;

Distribuição Exponencial

Densidade Acumulada

Teorema

Se $X \sim Exp(\lambda)$, então sua f.d.a. é dada por

$$F_X(x) = \mathbb{P}(X \leq x) = 1 - e^{-\frac{x}{\lambda}}$$

- **Observação:** Note que a f.d.a. nos diz a probabilidade de X ser menor ou igual que um valor qualquer, de maneira que seu complementar é a probabilidade de X ultrapassar um determinado valor;
- Se X é o tempo até um equipamento falhar:
  - Usamos a f.d.a. para calcular a probabilidade de que o equipamento queime em menos de 10 minutos;

Distribuição Exponencial

Densidade Acumulada

Teorema

Se $X \sim Exp(\lambda)$, então sua f.d.a. é dada por

$$F_X(x) = \mathbb{P}(X \leq x) = 1 - e^{-\frac{x}{\lambda}}$$

- **Observação:** Note que a f.d.a. nos diz a probabilidade de X ser menor ou igual que um valor qualquer, de maneira que seu complementar é a probabilidade de X ultrapassar um determinado valor;
- Se X é o tempo até um equipamento falhar:
  - Usamos a f.d.a. para calcular a probabilidade de que o equipamento queime em menos de 10 minutos;
  - E usamos o complementar para calcular a probabilidade de durar pelo menos 10 minutos.

Distribuição Exponencial

Falta de memória

Teorema

Se $X \sim Exp(\lambda)$, então

$$\mathbb{P}(X > t + s|X > t = \mathbb{P}(X > s)$$

- Se estamos avaliando tempo entre clientes que chegam em uma agência bancária e queremos saber a probabilidade de um cliente chegar 30 minutos depois da agência estar aberta há 3 horas, isso é a mesma coisa que calcular a probabilidade de um cliente chegar meia hora depois do banco abrir.

Distribuição Exponencial

Exercício

Sejam $X_1 \sim \text{Exp}(\lambda_1)$ e $X_2 \sim \text{Exp}(\lambda_2)$ duas v.a. aleatórias independentes e seja $Z = \min\{X_1, X_2\}$.

Encontre a f.d.a. de Z. Que distribuição é essa?

Distribuição Exponencial

Exercício

Sejam $X_1 \sim \text{Exp}(\lambda_1)$ e $X_2 \sim \text{Exp}(\lambda_2)$ duas v.a. aleatórias independentes e seja $Z = \min\{X_1, X_2\}$.

Encontre a f.d.a. de Z. Que distribuição é essa?
Dica: Utilize a definição: Se X e Y são independentes, então $\mathbb{P}(X \cap Y) = \mathbb{P}(X)\mathbb{P}(Y)$.

Definição

Seja X uma v.a. discreta, tomando valores em $\mathbb{N}$. Dizemos que X tem distribuição de Poisson se sua função massa de probabilidade (f.m.p.) é dada por:

$$p_x = \mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad k \in \{1, 2, 3, \cdots\} \tag{2}$$

Onde k representa o número de ocorrências de interesse e λ a taxa média de ocorrência. Utilizamos a notação $X \sim Poisson(\lambda)$.

Distribuição de Poisson

Observações

Para que a distribuição de Poisson seja apropriada para modelar um fenômeno X, as seguintes condições devem ser atendidas:

a. X é o número de ocorrências em um determinado intervalo de tempo e assume valores em $\{1, 2, 3, \cdots\}$;

Distribuição de Poisson

Observações

Para que a distribuição de Poisson seja apropriada para modelar um fenômeno X, as seguintes condições devem ser atendidas:

a. X é o número de ocorrências em um determinado intervalo de tempo e assume valores em $\{1, 2, 3, \cdots\}$;

b. as ocorrências acontecem de forma independente, isto é, o fato de um carro já ter passado em um pedágio não altera a probabilidade do próximo carro passar;

Distribuição de Poisson
Observações

Para que a distribuição de Poisson seja apropriada para modelar um fenômeno X, as seguintes condições devem ser atendidas:

a. X é o número de ocorrências em um determinado intervalo de tempo e assume valores em $\{1, 2, 3, \cdots\}$;

b. as ocorrências acontecem de forma independente, isto é, o fato de um carro já ter passado em um pedágio não altera a probabilidade do próximo carro passar;

c. a taxa λ com que os eventos ocorrem deve ser constante;

Distribuição de Poisson

Observações

Para que a distribuição de Poisson seja apropriada para modelar um fenômeno X, as seguintes condições devem ser atendidas:

a. X é o número de ocorrências em um determinado intervalo de tempo e assume valores em $\{1, 2, 3, \cdots\}$;

b. as ocorrências acontecem de forma independente, isto é, o fato de um carro já ter passado em um pedágio não altera a probabilidade do próximo carro passar;

c. a taxa λ com que os eventos ocorrem deve ser constante;

d. dois eventos não podem ocorrer no mesmo instante;

Distribuição de Poisson

Observações

Para que a distribuição de Poisson seja apropriada para modelar um fenômeno X, as seguintes condições devem ser atendidas:

a. X é o número de ocorrências em um determinado intervalo de tempo e assume valores em $\{1, 2, 3, \cdots\}$;

b. as ocorrências acontecem de forma independente, isto é, o fato de um carro já ter passado em um pedágio não altera a probabilidade do próximo carro passar;

c. a taxa λ com que os eventos ocorrem deve ser constante;

d. dois eventos não podem ocorrer no mesmo instante;

e. a probabilidade de que um evento ocorra em um intervalo de tempo é proporcional ao comprimento deste intervalo.

Distribuição de Poisson

- Vamos verificar se de fato p_X define uma função massa;
- Vamos calcular a média e a variância de $X \sim \mathcal{P}(\lambda)$.

Distribuição Gama

Definição

Função Gama

$\Gamma : \mathbb{R}^+ \to \mathbb{R}^+$ é chamada de função Gama e é tal que:

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$$

Distribuição Gama

Definição

Função Gama

$\Gamma : \mathbb{R}^{+} \rightarrow \mathbb{R}^{+}$ é chamada de função Gama e é tal que:

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$$

Propriedades: (demonstração em [Stern and Izbicki, 2016])

1. Para $a \geq 1$, temos que $\Gamma(a) = (a-1)\Gamma(a-1)$
2. Se $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$

Distribuição Gama

Definição

Função Gama

Dizemos que uma variável aleatória X tem distribuição Gama com parâmetros(k, λ) e denotamos por $X \sim \text{Gama}(k, \lambda)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(k)\lambda^k} x^{k-1} e^{-\frac{x}{\lambda}} & \text{, if } x > 0 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em est. bayesiana, a distribuição Gama é utilizada como priori para modelar variâncias (ou precisões);

Distribuição Gama

Definição

Função Gama

Dizemos que uma variável aleatória X tem distribuição Gama com parâmetros(k, λ) e denotamos por $X \sim \text{Gama}(k, \lambda)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(k)\lambda^k} x^{k-1} e^{-\frac{x}{\lambda}} & \text{, if } x > 0 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em est. bayesiana, a distribuição Gama é utilizada como priori para modelar variâncias (ou precisões);
  - No modelo de regressão linear múltiplo, a priori Gama para a precisão dos erros com uma priori Normal para os coeficientes junto com verossimilhança normal resulta em uma posteriori conjunta Normal-Gama.

Distribuição Gama

Definição

Função Gama

Dizemos que uma variável aleatória X tem distribuição Gama com parâmetros(k, λ) e denotamos por $X \sim \text{Gama}(k, \lambda)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(k)\lambda^k} x^{k-1} e^{-\frac{x}{\lambda}} & \text{, if } x > 0 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em est. bayesiana, a distribuição Gama é utilizada como priori para modelar variâncias (ou precisões);
  - No modelo de regressão linear múltiplo, a priori Gama para a precisão dos erros com uma priori Normal para os coeficientes junto com verossimilhança normal resulta em uma posteriori conjunta Normal-Gama.
- Se $X \sim \text{Exp}(\lambda)$, então $X \sim Gama(1, \lambda)$
  - Dizemos que a distribuição exponencial é um *caso particular* da distribuição Gama.

Distribuição Gama

Lema

- *A densidade Gama de fato é uma densidade de probabilidade (sempre positiva e integra 1);*
- *Se $X \sim Gama(k, \lambda)$, então:*
  - $\mathbb{E}[X] = k\lambda$
  - $Var[X] = k\lambda^2$

Distribuição Beta

Definição

Dizemos que X tem distribuição Beta com parâmetros (α, β) e denotamos por $X \sim \text{Beta}(\alpha, \beta)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1} & \text{, if } 0 < x < 1 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em bayesiana, a densidade Beta é comumente utilizada como priori para parâmetros que signifiquem elasticidades (em módulo), probabilidades e/ou o coeficiente de um processo *AR* que seja estacionário;

Distribuição Beta

Definição

Dizemos que X tem distribuição Beta com parâmetros (α, β) e denotamos por $X \sim \text{Beta}(\alpha, \beta)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1} & \text{, if } 0 < x < 1 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em bayesiana, a densidade Beta é comumente utilizada como priori para parâmetros que signifiquem elasticidades (em módulo), probabilidades e/ou o coeficiente de um processo *AR* que seja estacionário;
    -
    - A densidade Beta é conjugada da Binomial, isso significa que uma priori Beta com uma verossimilhança Binomial resulta em uma posteriori Beta.

Distribuição Beta

Definição

Dizemos que X tem distribuição Beta com parâmetros (α, β) e denotamos por $X \sim \text{Beta}(\alpha, \beta)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1} & \text{, if } 0 < x < 1 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em bayesiana, a densidade Beta é comumente utilizada como priori para parâmetros que signifiquem elasticidades (em módulo), probabilidades e/ou o coeficiente de um processo *AR* que seja estacionário;

Distribuição Beta

Definição

Dizemos que X tem distribuição Beta com parâmetros (α, β) e denotamos por $X \sim \text{Beta}(\alpha, \beta)$ se a densidade de X é

$$f_X(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1} & \text{, if } 0 < x < 1 \\ 0 & \text{, caso contrário} \end{cases}$$

- Em bayesiana, a densidade Beta é comumente utilizada como priori para parâmetros que signifiquem elasticidades (em módulo), probabilidades e/ou o coeficiente de um processo *AR* que seja estacionário;
  - A densidade Beta é conjugada da Binomial, isso significa que uma priori Beta com uma verossimilhança Binomial resulta em uma posteriori Beta.

Distribuição Beta

Lema

A densidade na Definição 14 é uma fdp válida. Em particular,

$$\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx = 1$$

Distribuição Beta

Lema

A densidade na Definição 14 é uma fdp válida. Em particular,

$$\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx = 1$$

Lema

Se $X \sim Beta(\alpha, \beta)$*, então para todo* $c, d \geq 0$ $\mathbb{E}[X^c(1-X)^d] = \frac{\Gamma(\alpha+c)\Gamma(\beta+d)}{\Gamma(\alpha+\beta+c+d)}$.
Portanto, $\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$ *e* $Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

Demonstração: Fica como exercício :)

Distribuição Normal

Definição

Dizemos que X tem distribuição Normal com parâmetros (μ, σ^2) e denotamos por $X \sim \mathrm{N}(\mu, \sigma^2)$ se a densidade de X é

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Se $\mu = 0$, e $\sigma = 1$, dizemos que X tem distribuição normal padrão.

Distribuição Normal

Definição

Dizemos que X tem distribuição Normal com parâmetros (μ, σ^2) e denotamos por $X \sim \mathrm{N}(\mu, \sigma^2)$ se a densidade de X é

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Se $\mu = 0$, e $\sigma = 1$, dizemos que X tem distribuição normal padrão.

Definição

Seja $Z \sim \mathrm{N}(0, 1)$.

$$\phi(z) = \mathbb{P}(Z \leq z) = F_Z(z)$$

ϕ não tem solução analítica.

Distribuição Normal

Lema

Se $X \sim N(\mu, \sigma^2)$, *então* $\frac{(X-\mu)}{\sigma} \sim N(0, 1)$.

Distribuição Normal

Lema

Se $X \sim N(\mu, \sigma^2)$, *então* $\frac{(X-\mu)}{\sigma} \sim N(0, 1)$.

Lema

Se $X \sim N(\mu, \sigma^2)$, *então* $\mathbb{E}[X] = \mu$ *e* $Var[X] = \sigma^2$.

Densidades Conjuntas

Motivação

Considere duas v.a. discretas, X e Y, associadas ao mesmo experimento. As probabilidades que X e Y podem assumir são descritas pela função densidade de probabilidade conjunta de X e Y:

$$f_{X,Y}(x, y) = \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x, Y = y)$$

Densidades Conjuntas

Motivação

Considere duas v.a. discretas, X e Y, associadas ao mesmo experimento. As probabilidades que X e Y podem assumir são descritas pela função densidade de probabilidade conjunta de X e Y:

$$f_{X,Y}(x,y) = \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x, Y = y)$$

A f.d.p. conjunta de X e Y determina a probabilidade de qualquer evento que possa ser especificado em termos de X e Y:

$$\mathbb{P}((X,Y) \in A) = \sum_{(x,y) \in A} f_{X,Y}(x,y)$$

Densidades Conjuntas

Motivação

Considere duas v.a. discretas, X e Y, associadas ao mesmo experimento. As probabilidades que X e Y podem assumir são descritas pela função densidade de probabilidade conjunta de X e Y:

$$f_{X,Y}(x,y) = \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x, Y = y)$$

A f.d.p. conjunta de X e Y determina a probabilidade de qualquer evento que possa ser especificado em termos de X e Y:

$$\mathbb{P}((X,Y) \in A) = \sum_{(x,y)\in A} f_{X,Y}(x,y)$$

Pode-se calcular a f.d.p. somente de X ou de Y (*densidade marginal*) através da conjunta:

$$f_X(x) = \sum_y f_{X,Y}(x,y) \qquad f_Y(y) = \sum_x f_{X,Y}(x,y)$$

Definição

Sejam X e Y duas variáveis aleatórias discretas. A fmp conjunta de X e Y é $p_{X,Y}(x,y) := \mathbb{P}(X = x, Y = y)$. De forma similar, se X e Y são duas variáveis aleatórias com distribuições contínuas, a fdp conjunta de X e Y é $f_{X,Y}(x,y)$. Seja A um subconjunto de $\mathbb{R}^2$,

$$\mathbb{P}((X,Y) \in A) = \mathbb{P}(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in A\}) =$$

$$\begin{cases} \sum_{(x,y)\in A} p_{X,Y}(x,y) & \text{, se } X \text{ e } Y \text{ são discretas.} \\ \int_{(x,y)\in A} f_{X,Y}(x,y)d(x,y) & \text{, se } X \text{ e } Y \text{ são contínuas.} \end{cases}$$

Densidades Conjuntas

Exemplo

Exemplo

Considere que X e Y are são variáveis aleatórias discretas com a fmp conjunta dada na Tabela abaixo.

X/Y	0	1
0	0.2	0.4
1	0.3	0.1

Densidades Condicionais

Exemplo

Exemplo

Sejam X e Y variáveis aleatórias contínuas, com a seguinte fdp

$$f_{X,Y}(x,y) = \begin{cases} cxy & \text{, if } x > 0, y > 0, x + y < 1 \\ 0 & \text{, caso contrário} \end{cases}$$

Qual é o valor de c?

Independência de v.a.'s

Definição

Duas variáveis aleatórias contínuas X e Y são independentes se e somente se $f_{X,Y}(x, y) = f_X(x)f_Y(y)$.

Independência de v.a.'s

Definição

Duas variáveis aleatórias contínuas X e Y são independentes se e somente se $f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

Lema

Sejam X e Y duas variáveis aleatórias. X e Y são independentes se e somente se para todos $A, B \subset \mathfrak{R}$, $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.

Independência de v.a.'s

Definição

Duas variáveis aleatórias contínuas X e Y são independentes se e somente se $f_{X,Y}(x, y) = f_X(x)f_Y(y)$.

Lema

Sejam X e Y duas variáveis aleatórias. X e Y são independentes se e somente se para todos $A, B \subset \mathfrak{R}$, $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.

Lema

Duas variáveis contínuas, X e Y, são independentes se e somente se a distribuição conjunta pode ser escrita como $f_{X,Y}(x, y) = g(x)h(y)$ para alguma função g e h.

Densidades Condicionais

Caso discreto

Definição

Se X e Y são variáveis aleatórias discretas, definimos a *massa de probabilidade condicional* de X dado que $Y = y$ por

$$p_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)},$$

para todos os valores de y tais que $\mathbb{P}(Y = y) > 0$.

Densidades Condicionais

Caso discreto

Definição

Se X e Y são variáveis aleatórias discretas, definimos a *massa de probabilidade condicional* de X dado que $Y = y$ por

$$p_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)},$$

para todos os valores de y tais que $\mathbb{P}(Y = y) > 0$.

Note que, para um dado y, $g(x) = p_{X|Y}(x|y)$ é uma fmp no sentido em que a estudamos antes.

Densidades Condicionais

Caso contínuo

Definição

Se X e Y são variáveis aleatórias contínuas com densidade conjunta $f(x, y)$, definimos a *densidade de probabilidade condicional* de X dado $Y = y$ por

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

para todos os valores de y tais que $f_Y(y) > 0$.

Densidades Condicionais

Caso contínuo

Exemplo

Seja densidade conjunta de X e Y definida por

$$f(x,y) = \begin{cases} \frac{15}{2}x(2-x-y) & \text{se } 0 < x, y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

Calcule $X|Y = y$ quando $0 < y < 1$.

Densidades Condicionais

Esperança condicional - caso discreto

Definição

Esperança condicional

A esperança condicional de X dado $Y = y$, no caso discreto, é dada por:

$$\begin{aligned}
\mathbb{E}[X|Y = y] &= \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}|Y = y) \\
&= \sum_{x \in \mathcal{X}} \sum_{\{\omega : X(\omega) = x\}} X(\omega)\mathbb{P}(\{\omega\}|Y = y) \\
&= \sum_{x \in \mathcal{X}} x \sum_{\{\omega : X(\omega) = x\}} \mathbb{P}(\{\omega\}|Y = y) \\
&= \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}(X = x|Y = y) \\
&= \sum_{x \in \mathcal{X}} x \cdot p_{X|Y}(x|y)
\end{aligned}$$

Definição

Esperança condicional

A esperança condicional de X dado $Y = y$, no caso contínuo, é dada por:

$$\mathbb{E}[X|Y = y] = \int_{\mathbb{R}} x \cdot f_{X|Y}(x|y)dx$$

Densidades Condicionais

Esperança condicional - caso contínuo

Definição

Esperança condicional

A esperança condicional de X dado $Y = y$, no caso contínuo, é dada por:

$$\mathbb{E}[X|Y = y] = \int_{\mathbb{R}} x \cdot f_{X|Y}(x|y)dx$$

Atenção!

- $\mathbb{E}[X|Y]$ é uma variável aleatória,
- $\mathbb{E}[X|Y = y]$ é um número.

Lema

Sejam X e Y duas v.a.. Então,

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$$

Função Geradora de Momentos

Definição

(Função Geradora de Momentos) A função geradora de momentos de uma variável aleatória X é uma função $M_X : \mathbb{R} \to \mathbb{R}$ definida por:

$$M_X(t) = \mathbb{E}[e^{tX}] \tag{3}$$

Isto é, a função geradora de momentos é calculada através da esperança da função e^{tX}.

Referências I

Lima, E. L. (1982).
Curso de analise: volume 1, volume 1.
Instituto de Matematica Pura e Aplicada.

Stern, R. and Izbicki, R. (2016).
Introducao à Teoria das Probabilidades e Processos Aleatorios.
UFSCAR.