Resumo Probabilidade

Espaço amostral S, eventos A e B : $P(A) \geq 0$, $P(B) \geq 0$, $P(S) = 1$ $P(\cup A_i) = \sum (P(A_i))$ para A_i disjuntos

Probabilidade Condicional

$P(A|B) = P(A \cap B)/P(B)$

PROBABILIDADE TOTAL

$P(A) = \sum P(A|B_i).P(B_i)$ se B_i são partição de S.

TEOREMA DE BAYES

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

Eventos $A_1, A_2, \ldots A_n$ independentes se $P(A_{i_1} \cap \ldots A_{i_k}) = P(A_{i_1}) \ldots \times \ldots P(A_{i_k})$

Nº escolhas ordenadas de k elementos entre n possíveis: $\frac{n!}{(n-k)!}$

Nº de escolhas não ordenadas de k elementos dentre n $\frac{n!}{k!(n-k)!} = \binom{n}{k}$

$\frac{6!}{0!(6-0)!}(0{,}5)^6$

$\frac{6!}{6!} \cdot 0{,}5(0{,}5)^5$

$0{,}0937\,5$

Variável Aleatória

* Discreta (intervalos por finitos números) $P(X = x)$ $\quad E[X] = \sum_i x_i P(X = x_i)$

* Contínua, densidade de probabilidade $f_x(x) = \frac{d F_x(x)}{dx}$, onde $F_x(x) = P(X \leq x)$ é a função de distribuição cumulativa

$F_x(X)$ é não decrescente, se $X \to -\infty$ $F_x(x) \to 0$ e se $X \to +\infty$ $F_x(x) \to 1$

$P(a \leq x \leq b) = \int_a^b f_x(x)\,dx$, A esperança é: $E(X) = \int x f_x(x)\,dx$,

$$V(X) = E[(x - E(x))^2] = E(x^2) - (E(x))^2 = \sigma^2$$

$\sigma^2 = E[(X-\mu)^2] = \sum (X-\mu)^2 . p(x)$

momento de ordem K de x é $E(x^k)$

Se $h(x)$ é função de variável aleatória X

$E(h(x)) = \sum h(x_i) P(X = x_i)$ se X é discreto

ou

$E(h(x)) = \int h(x) f_x(x)\,dx$ se x é contínua $\rightarrow \int (x - E(x))^2 . f(x)\,dx$

ex: $X: f(x)$

f.d.p.

$g(X)$ qualquer

$E(g(x)) = \int g(x) f(x)\,dx$

Scanned by CamScanner

→ SUCESSO E FRACASSO

Binomial (n ensaios, p probabilidades de sucesso)

$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$, $E(x) = n.p$, $V(x) = np(1-p)$

Geométrica (p probabilidade de sucesso) → série ilimitada de experimentos

$P(X=x) = p(1-p)^x$ e $E(x) = (1-p)/p$ $V(x) = (1-p)/p^2$

Poisson (λ) → Binomial quando nº de ensaios → infinito | mas $np = \lambda$ constante

$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$ e $E(x) = V(x) = \lambda$

$\boxed{P(X<x) = P\left(z \leq \frac{x-\mu}{\sigma}\right) = F(z)}$

Normal (μ, σ^2)

$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{(2\sigma^2)}\right)$, $E(x) = \mu$, $V(x) = \sigma^2$

$\boxed{z = \frac{X-\mu}{\sigma}}$ é $N(0,1)$ sempre

se X é $N(\mu, \sigma^2)$ e $Y = aX+b$ com $a \neq 0$ então Y é $N(a\mu+b, (a\sigma)^2$

se X é $N(\mu_1, \sigma_1^2)$ e Y é $N(\mu_2, \sigma_2^2)$ e X e Y são independentes então $W = X+Y$

é $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ e $Z = X-Y$ é $(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)$

Exponencial (λ) : $f_x(x) = \lambda \exp(-\lambda x)$ | $E(x) = \frac{1}{\lambda}$ $V(x) = \frac{1}{\lambda^2}$

função de probabilidade acumulada : $1 - e^{-\lambda . x}$

Intuito: probabilidade para criar modelos.

O espaço amostral S é o cjto de todos os resultados possíveis de um experimento.

Relação com teoria de conjuntos

A ou B → $A \cup B$ A mas não B $A - B$

A e B → $A \cap B$

não A → A^c

Regras de Morgan (conjuntos) $\begin{cases} (A \cup B)^c = A^c \cap B^c \\ (A \cap B)^c = A^c \cup B^c \end{cases}$

Eventos disjuntos

↳ não possuem elementos em comum $A_1 \cap A_2 = \phi$

Partição → uma partição de S é uma coleção de conjuntos disjuntos que a união é S.

$S = A_1 \cup A_2 \cup A_3 \cup \ldots A_n$ se A_1, A_2, A_n disjuntos

útil: $A = A \cap S = A \cap (A_1 \cup A_2 \cup A_3 \ldots A_n) = (A \cap A_1) \cup (A \cap A_2) \ldots \cup (A \cap A_n)$

Axiomas da Probabilidade

1 Para qualquer evento A $P(A) \geq 0$

2 $P(S) = 1$

3 Se $A_1, A_2, A_3 \ldots$ grupo infinito de elementos disjuntos vale:

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = P(A_1 \cup A_2 \cup A_3 \ldots \cup A_k) = \sum_{k=1}^{n} P(A_k)$$

Scanned by CamScanner

Definição Clássica

S são N eventos

A é evento simples de S A são m

$$\boxed{P(A) = \frac{m}{N}}$$

para um evento A

propriedades usadas $P(A) + P(A^c) = 1$ então. $P(A) = 1 - P(A^c)$

para quaisquer 2 eventos A e B

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

para 3 eventos A, B e C

$$P(A \cup B \cup B) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Técnicas de Contagem

Arranjos → amostras ordenadas

→ duas amostras com os mesmos elementos em ordem distinta são diferentes

O número de arranjos (sem reposição) de tamanho n que pode ser criado a partir de N elementos é $(N)_n = \frac{N!}{(N-n)!}$

com reposição temos: N^n

Permutações → amostra ordenada de n elementos

o número de permutações de n elementos é $P_n = n!$

Combinações → subconjunto desordenado

O número de combinações (amostras ñ ordenadas sem reposição) de tamanho n que pode ser criado a partir de N elementos é igual a:

$$\frac{(N)_n}{P_n} = \binom{N}{n} = \frac{(N)_n}{n!} = \frac{N!}{n!(N-n)!} = C_{N,n}$$

Scanned by CamScanner

Probabilidade Condicional

Para eventos A e B com $P(B)>0$, a probabilidade de A dado que B ocorreu é

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \longrightarrow P(A \cap B) = P(A|B).P(B)$$

Teorema da Probabilidade Total

$A_1, \dots, A_K$ partições de S então para qualquer outro B de S

$$P(B) = \sum_{i=1}^{K} P(A_i \cap B) = P(A_1 \cap B) + \dots + P(A_K \cap B) = \sum_{i=1}^{K} P(B|A_i).P(A_i)$$

Teorema de Bayes (Corolário do P. Total)

Sejam $A_1, \dots, A_n$ partições de S (eventos disjuntos) com probabilidade a Priori $P(A_i)$ $i = 1, 2, \dots K$. Então, para qualquer evento B em S para o qual $P(B)>0$ a probabilidade a Posteriori de A_J dado que B ocorreu é:

$$P(A_J/B) = \frac{P(A_J \cap B)}{P(B)} = \frac{P(B|A_J).P(A_J)}{P(B)} \overset{\text{P.T.}}{=} \frac{P(B|A_J).P(A_J)}{\sum_{i=1}^{K} P(B|A_i).P(A_i)}$$

INDEPENDÊNCIA

A e B independem se $P(A|B) = P(A)$

$$P(A|B) = P(A) \Rightarrow P(B|A) = \frac{P(A|B).P(B)}{P(A)} = \frac{P(A).P(B)}{P(A)} = P(B)$$

A^C e B independem
A e B^C "
A^C e B^C "

Para dois eventos disjuntos se A ocorrer, B não pode ter ocorrido $P(B|A) = 0 \neq P(A)$ → não são independentes

se A e B são independentes

$P(A \cap B) = P(A).P(B)$ (vale a extensão para n eventos)

Scanned by CamScanner

Vaváel Aleatária

• X é qualquer coisa que associa um número com qualquer resultado s de S, ou seja. V.A. é uma função que atribui um X(s) a um número s de saída.

se V.A. possui números de valores 0 ou 1 → Vaváel de Bernalli

Distribuição de Probabilidade – discreta

definida para todo x do domínio

$P(X = x) = P(\text{todos } s \in S : X(x) = x)$

– condições para todas as distribuições de probabilidade

$P(X = x) \geq 0$

$\sum_{\text{todos } x} P(X = x) = 1$

Função de distribuição de uma V.A. discreta

$F(x)$

A função de distribuição acumulada de uma vaváel aleatória X com prob. P(X) é:

$F(X) = P(X \leq x) = \sum_{y : y \leq x} P(X = y)$ → probabilidade do máximo valor observado ser X

Para qualquer valor de y, F(y) será igual o valor de F no valor mais próximo possível à esquerda de y.

proposição: $P(a \leq X \leq b) = F(b) - F(a-)$

onde a- é o maior valor possível de X antes de a.

se quiser $P(X = 3)$ é igual a $F(3) - F(2)$ para 2 contendo 3.

Scanned by CamScanner

Variáveis aleatórias contínuas

distribuição

a função de densidade de probabilidade de uma V.A. contínua é uma $f(x)$ tal que

$f(x) \geq 0 \quad \forall x$

$\int_{-\infty}^{+\infty} f(x) = 1$

$P(a \leq x \leq b) = \int_a^b f(x)dx \text{ com } a \leq b$

são áreas do gráfico

Função de distribuição acumulada do VA contínua

$$F(X) = P(X \leq x) = \int_{-\infty}^{x} f(\lambda)d\lambda$$

* $P(a \leq X \leq b) = F(b) - F(a)$

Valor Esperado de V.A. discreta – média

X é V.A. discreta com N valores possíveis $x_1, x_2, \ldots, x_n$ e distribuição de probabilidade $P(X = x_i), i = 1, 2, \ldots, N$. O valor médio é:

$$E(x) = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

valor esperado de uma função $E(g(x))$

(o V.A.)

$$E(g(x)) = \sum_{i=1}^{n} g(x_i) \cdot P(X = x_i)$$

se $g(x)$ é linear com x. $g(x) = ax + b$ $\quad E(g(x)) = \bar{a} \cdot E(x) + b$

$E(X_1 + X_2) = E(X_1) + E(X_2)$

Scanned by CamScanner

Variância de X

↳ medida de variabilidade da distribuição

Fórmula:

$$\sigma^2(X) = \sum_{i=1}^{n} (x_i - E(X))^2 \cdot P(X = x_i) = E[(X - E(X))^2]$$

Desvio Padrão é $\sigma(X) = \sqrt{\sigma^2(X)}$

$$\sigma^2(X) = E(X^2) - E(X)^2$$

Valor esperado de V.A. contínua

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx \quad , \text{ para função } g(X) \quad E(g(X)) = \int_{-\infty}^{+\infty} g(x)\, f(x)\,dx$$

Variância de V.A. contínua com fdp $f(x)$ e média μ

$$\sigma^2(x) = \int_{-\infty}^{+\infty} [x - E(x)]^2 f(x)\,dx = E[(X - E(x))^2]$$

$$\sigma(x) = \sqrt{\sigma^2(x)}$$

Scanned by CamScanner

Probabilidade

Modelos Probabilísticos Discretos

Binomial (SUCESSO/FRACASSO)

- n ensaios pré-definidos independentes
- probabilidade de sucesso para cada ensaio é igual a p $P(S) = p$
- probabilidade de fracasso $P(F) = 1 - P(S) = 1 - p$

⇒ Dos o amostragem for feito SEM REPOSIÇÃO → não é binomial

$X \sim Bin(n, p)$ → quantidade de sucessos em n ensaios de probabilidade p.

$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \dots, n \\ 0, & \text{caso contrário} \end{cases}$$

probabilidade "de no máximo" k $P(X \leq K) = P(X = K-1) \dots (P(X = K))$

probabilidade de "pelo menos" K $P(X \geq K) = P(X = K) + P(X = K+1) + P(X = K+2) \dots$

Logo $F(x) = P(X \leq x) = \sum_{y=0}^{x} P(X = y), \quad x = 0, 1, 2, \dots n$ função de distribuição acumulada

Para variável discreta vale $P(a < X \leq b) = F(b) - F(a)$ $a \leq b$

média = $E(X) = np$

variância = $\sigma^2(X) = np(1-p)$

Scanned by CamScanner

Distribuição Geométrica

- número ilimitado de ensaios { p sucesso / 1-p falha
- ensaios ocorrem até o primeiro sucesso
- X dá o número de falhas que antecedem o primeiro sucesso

$$P(X=j) = (1-p)^j p$$

média: $\frac{p(1-p)}{p^2} = \frac{(1-p)}{p}$ variância $= \sigma^2(X) = \frac{1-p}{p^2}$

"probabilidade de o primeiro ser o 10º"

↳ Probabilidade de os 9 primeiros serem falha

Distribuição de Poisson (parâmetro λ)

$$P(X=K) = \frac{e^{-\lambda} \cdot \lambda^K}{K!}, \quad K = 0,1,2,\ldots$$

média $= E(X) = \lambda$ variância $= \sigma^2(X) = \lambda$

* limite do Binomial $n \to \infty$, $p \to 0$ mas $np \to \lambda > 0$

Ou seja, num experimento binomial com n muito grande ($n > 50$)

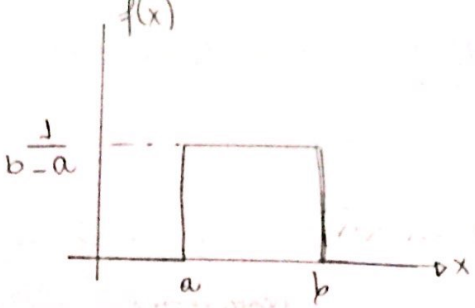
e p pequeno ($p < 5/n$)

$P(\text{Binomial}) \approx P(\text{Poisson})$ com $\lambda = n \cdot p$

Scanned by CamScanner

Distribuição Uniforme

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{caso contrário} \end{cases}$$



$$F(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , x \geq b \end{cases}$$

Media $= E(x) = \frac{a+b}{2}$

$Var(X) = \sigma^2(x) = \frac{(b-a)^2}{12}$

$B \sim N((E(B), (\sigma^2(B)))$

$n \quad p \qquad np(1-p)$

$B(950, \frac{1}{3})$

$B \sim N(E(B), \sigma^2(B))$

$N(E(B), \sigma^2(B))$

Distribuição Normal

parametros μ e σ^2 $-\infty < \mu < +\infty$ e $\sigma > 0$

fdp $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < X < +\infty$

$N(\mu, \sigma^2)$ $N((150, 100))$

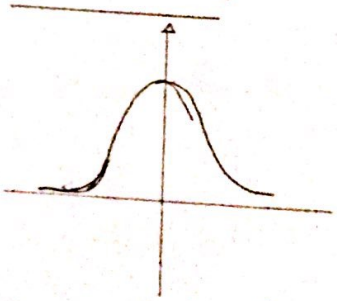
$1000 \quad 10^2$

$E(x) = \mu$ e $Var(x) = \sigma^2$

$z = \frac{X-\mu}{\sigma}$

$\frac{1200 - 1000}{10} = \frac{200}{10} = 20$

Normal Padrão



para $\mu = 0$ e $\sigma^2 = 1 \Rightarrow$ TABELA

contem valores do tipo $P(0 < z < zc) = \int_0^{zc} f(z)dz$

$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

para distribuições normais não padrão $X \sim N(\mu, \sigma^2) \rightarrow$ FAZER mudança

$z = \frac{X-\mu}{\sigma}$ então $P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq z \leq \frac{b-\mu}{\sigma}\right)$

$P(X \leq a) = P\left(z \leq \frac{a-\mu}{\sigma}\right)$, $P(X \geq b) = 1 - P\left(z - \frac{b-\mu}{\sigma}\right)$

Scanned by CamScanner

Função de variável aleatória (X variável aleatória)

$E(x)$ e $\sigma^2(x)$

$h(X) = aX+b$ logo $E(h(X)) = aE(x)+b$ e $Var(h(X)) = a^2\sigma^2(X)$

supondo $h^2(X)$, $E(h^2(x)) = a^2E(X^2) + 2abE(X) + b^2$

supondo $h(x^2)$ $E(h(x^2)) = a^2E(x)^2 + 2abE(X) + b^2$

Confiabilidade de um sistema

X é V.A. que representa um tempo de vida, F(X) é distribuição do tempo de vida

A confiabilidade é $R(X) = 1 - F(X)$

a taxa de falha é $h(x) = \dfrac{f(x)}{1-F(x)}$ $\approx$ Prob. de que ocorra falha num Δt pequeno em relação a t, após o instante t dado que não houve falha até t.

$P(X \le 10) = 1 - F(10)$ contínuo — $F(10)$ → $P(0<X<10)$

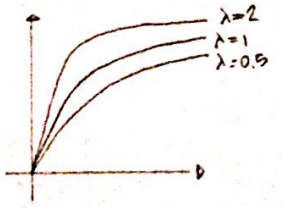
Distribuição Exponencial contínua, parâmetro λ ($\lambda > 0$)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{caso contrário} \end{cases}$$

$$E(x) = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda} = \text{desvio padrão}$$

$$V(x) = \sigma^2 = \frac{1}{\lambda^2}$$

$$F(x) = \begin{cases} 0, & x<0 \\ 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

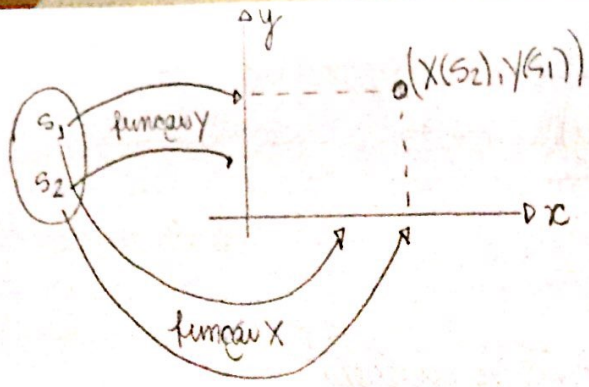


confiabilidade da exponencial $R(x) = 1 - F(x) = e^{-\lambda x}$, $h(x) = \dfrac{f(x)}{1-F(x)} = \dfrac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$ ← taxa de falha

Como a taxa de falha é constante, vemos que

a probabilidade de um componente falhar dado que um outro Δt já passou é o mesmo. (Falha independe do tempo decorrido).

Scanned by CamScanner



X e Y são duas V.A. discretas. A distribuição de probabilidade conjunta:

$p(x,y) = P(X=x, Y=y)$ → função de massa de probabilidade conjunta

- $p(x,y) \geq 0$ e $\sum_x \sum_y p(x,y) = 1$
- se $A = \{(x,y) : x+y=5\}$ então $P[(x,y) \in A]$ é obtido como a soma da distribuição de probabilidade conjunta com os pares em A: $P[(x,y) \in A] = \sum_{(x,y) \in A} \sum p(x,y)$

Distribuições Marginais

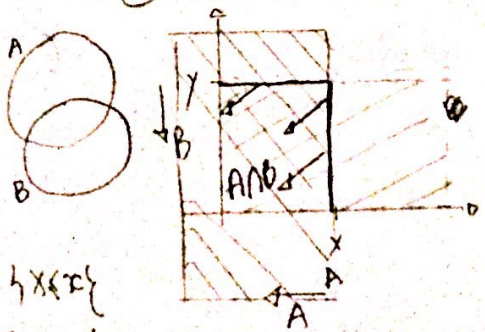
X: É dado por $P(X=x) = \sum_y p(x,y)$ para cada valor possível de X

Y: $P(Y=y) = \sum_x p(x,y)$ para cada possível valor de Y.

} conta as P(x) para todos os y e as P(y) para todos os x

=> só de X, só de Y

Distribuição conjunta



$A = \{X \leq x\}$

$B = \{Y \leq y\}$

$A \cap B = \{X \leq x, Y \leq y\}$

define-se como f de distribuição conjunta de X e Y

$F(x,y) = P[X \leq x, Y \leq y] = P(A \cap B)$

X e Y são independentes se:

$P[X=x, Y=y] = P[X=x] \cdot P[Y=y]$

CONJUNTA — MARGINAIS

Scanned by CamScanner

Alem disso, para variáveis aleatórias independentes

$$E(XY) = E(X).E(Y) \quad , \quad F(x,y) = F_X(x).F_Y(y)$$

Covariância e coef. de correlação

↳ se (X,Y). V.A. bidimensional

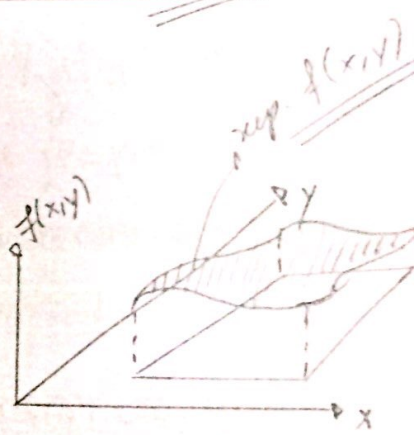
$$Cov(X,Y) = E(XY) - E(X).E(Y)$$

se X e Y independem

$$cov(XY) = 0$$

Coeficiente de correlação

$$\rho(XY) = \frac{cov(XY)}{\sigma(X)\sigma(Y)} \leq 1$$



Para Variáveis aleatórias contínuas } mesmas regras

$$\iint_S f(x,y)\,dS = 1$$

$$f(x,y) \geq 0$$

função de densidade de prob. conjunto $f(x,y)$

$$P(X,Y \in A) = \iint_A f(x,y)\underbrace{dy\,dx}_{dA}$$

se $\underbrace{a \leq x \leq b \text{ e } c \leq y \leq d}_{\bar{\bar{A}}}$ $P(X,Y \in A) = \int_a^b \int_c^d f(x,y)\,dy\,dx$

Variáveis aleatórias independentes → conceito para várias variáveis

↳ são independentes se

$$p(x,y) = p_X(x) \cdot p_Y(y)$$ quando x e y são discretas

$$f(x,y) = f_X(x) \cdot f_Y(y)$$ quando x e y são contínuas

se essas condições não forem satisfeitas, x e y são dependentes

⇒ Quando dito que uma V.A. não afeta a outra, podemos concluir que a distribuição conjunta é o PRODUTO das marginais.

$$P(a \leq X \leq b,\ c \leq Y \leq d) = P(a \leq X \leq b) \cdot P(c \leq Y \leq d)$$

Scanned by CamScanner

As V.A. do esquema pelo menos uma variável podem ser de quaisquer dos tipos já citados até o momento.

p/ mais de duas

$$p(x_1, x_2, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

e contínuas

$$P(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \dots a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, x_2, \dots x_n)\, dx_n \cdots dx_1$$

Se n variáveis são independentes fmp ou fdp conjunta é igual ao produto das marginais

* Para se determinar a distribuição conjunta, muitas vezes o procedimento é entender as marginais e usar o conceito de independência.

Distribuições Condicionais (para X_i com $i > 1$)

sejam X e Y duas vas aleatórias contínuas com fdp conjunta e fdp marginal de X é $f_X(x)$. Então, para qualquer valor x de X para o qual $f_X(x) > 0$, a função densidade de probabilidade condicional de Y DADO QUE $X = x$ é:

$$f_{Y/X}(y/x) = \frac{f(x, y)}{f_X(x)} \quad -\infty < y < +\infty$$

e $E(Y/X=x) = \int_{-\infty}^{+\infty} y \cdot f(y/x)\, dy$

Para vas discretas basta substituir fdp por fmp

"(EXEMPLO 5.12 DEVORE) - MUITO BOM"

seção 5.1 (1-21) EXERCÍCIO

→ usando uma condicional, resulto do problema uma função densidade de Y para quando $X = x$.

Scanned by CamScanner

Valores Esperados, Covariância e Correlação

Vimos que $E(h(x))$ é uma média ponderada dos valores de $h(x)$ em que o função do peso foi a fmp $p(x)$ ou a fdp $f(x)$ de X

Para duas variáveis é semelhante;

X, Y vas de dist. conjunta com fmp $p(x,y)$ ou fdp $f(x,y)$ conforme discretas ou contínuas

$$E[h(x,y)] = \begin{cases} \sum_x \sum_y h(x,y) \cdot p(x,y) & \text{se } X \text{ e } Y \text{ discretas} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) \cdot f(x,y)\, dx dy & \text{se } X \text{ e } Y \text{ contínuas} \end{cases}$$

Covariância → avalia o grau de dependência de duas variáveis aleatórias que não são independentes:

A covariância entre X e Y é

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = \begin{cases} \sum_x \sum_y (x-\mu_X)(y-\mu_Y)(p(x,y)) & X, Y \text{ discretas} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\mu_X)(y-\mu_Y) f(x,y) & X, Y \text{ contínuas} \end{cases}$$

supondo que $(x-\mu_X)$ e $(y-\mu_Y)$ são positivos ou negativos (ambos) teremos uma covariância positiva; caso possuem sinal trocado, teremos uma covariância negativa; a média ponderada pelos $p(x,y)$ dá a relação entre x e y.

se X e Y forem independentes $\underline{cov(X,Y) = 0}$

podemos usar, para facilitar cálculos:

$$Cov(X,Y) = E(XY) - \mu_X \cdot \mu_Y$$

A covariância depende para a sua análise dos módulos dos dados, ou seja, é afetada diretamente pelas unidades, por isso define-se a escala da covariância

Coeficiente de Correlação

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \, \sigma_Y}$$

σ_X e σ_Y são os desvios padrão

$$\sigma_X = \sqrt{\sigma_X^2}$$

↳ variância

→ exemplos de cálculo pags.

Scanned by CamScanner

Proposição para coeficiente de correlação:

1) Se a e c são ambos positivos ou negativos
$\text{corr ou } \rho(aX+b, cY+d) = \text{corr}(X,Y)$

o coeficiente de correlação não [illegible] por mudança linear de unidades

2) Para quaisquer duas V.A. X e Y $\boxed{-1 \leq \text{Corr}(X,Y) \leq 1}$

3) Se X e Y são independentes então $\rho = 0$, mas $\rho = 0$ não implica independência

4) $\rho = 1$ ou $\rho = -1$ apenas quando $Y = aX+b$, ou seja, quando X e Y forem relações lineares diretas

5) Quando $\rho = 0$, X e Y não são independentes, apenas há total ausência de relação linear entre as variáveis. As variáveis ainda podem possuir uma relação não linear fortíssima. Ou seja ρ apenas mensura a relação linear entre as V.A.

MAIS DE DUAS VARIÁVEIS ALEATÓRIAS

$X_1, X_2, \ldots, X_n$ discretas $p(X_1, X_2, \ldots, X_n) = P(X_1 = x_1, X_2 = x_2, X_n = x_n)$

$X_1, X_2, \ldots X_n$ contínuas

$$P((a_1 \leq X_1 \leq b_1), \ldots, (a_n \leq x_n \leq b_n)) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

DISTRIBUIÇÃO MULTINOMIAL

experimento de n ensaios independentes

cada ensaio resulta em uma das r possibilidades.

$p_i = P(\text{resultado } i \text{ em qualquer ensaio})$

X_i = número de ensaios que resultam i, $i = 1, \ldots, r$

A função massa é:

$$p(x_1, \ldots, x_n) = \begin{cases} \dfrac{n!}{(x_1!)(x_2!)\ldots(x_n)!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r} & x_i = 0, 1, \ldots \\ 0 & \end{cases}$$

ex: $p_1 = P(AA)$, $p_2 = P(Aa)$, $p_3 = P(aa)$ } $n = 10$ ensaios

$0{,}25$ $\quad = 0{,}50 \quad 0{,}25$

X_1 = nº de AA $\quad X_2$ = nº de Aa $\quad X_3$ = nº de aa

$$p(2,5,3) = \frac{10!}{2!\,5!\,3!} (0{,}25)^2 (0{,}5)^5 (0{,}25)^3 = \underline{0{,}0769}$$

Scanned by CamScanner

Combinação linear de V.A.'s independentes

X e Y duas V.A. independentes

com médias $E(X)$ e $E(Y)$ e variâncias $\sigma^2(X)$ e $\sigma^2(Y)$

$W = aX + bY + c$

$E(W) = a\,E(X) + b(E(Y)) + c$

$\sigma^2(W) = a^2\sigma^2(X) + b^2\sigma^2(Y)$

Combinação Linear de Gaussianas

$X_1, X_2, \dots, X_n$ Gaussianas, com médias $\mu_1, \mu_2, \dots, \mu_n$

e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$

se $Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$, Y também é gaussiana

$E(Y) = \mu_Y = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n$

$\sigma^2(Y) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2$

$Y(E(Y), \sigma^2(Y))$ → a partir disso conseguimos todas as informações sobre Y

$$\boxed{Z = \frac{X - \mu}{\sqrt{\sigma_Y^2}}}$$

para o conjunto de duas gaussianas temos:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}} \cdot e^{\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho_{xy}\frac{(x-\mu_x)}{\sigma_x}\frac{(y-\mu_y)}{\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right)\right]}$$

max quando em (μ_x, μ_y) $= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$

* se x e y independem $cov(x,y) = 0$ ($\rho_{xy} = 0$)
* Apenas para gaussianas vale a recíproca, ou seja,
  se $cov(x,y) = 0$ e $\rho_{xy} = 0$ → INDEPENDENTES!!

Teorema do Limite Central

com diversas V.A. independentes $X_1, \dots, X_n$ onde $Y = X_1 + \dots + X_n$,

isso Y se aproxima de uma gaussiana com média e variância já discutidas

(APLICAÇÃO?) Lista 11 e 12

Scanned by CamScanner