Probabilidade

Prof. Wagner Hugo Bonat

Curso de Especialização em
Data Science & Big Data
Universidade Federal do Paraná

17 de maio de 2018

Conteúdo

Conteúdo

1. Probabilidade
    1.1 Definições;
    1.2 Regra da adição de probabilidade;
    1.3 Probabilidade Condicional e Independência;
    1.4 Teorema de Bayes.
2. Variáveis aleatórias e Distribuições de Probabilidade
    2.1 Definição;
    2.2 Variáveis aleatórias discretas;
    2.3 Variáveis aleatórias contínuas;
    2.4 Distribuições conjuntas.
3. Esperança matemática
    3.1 Média de uma variável aleatória;
    3.2 Variância e Covariância de variáveis aleatórias.

Probabilidade

Definições básicas

- Fenômeno aleatório: situação ou acontecimento cujos resultados não podem ser previstos com certeza.
- Espaço amostral: conjunto de todos os resultados possíveis de um fenômeno aleatório, denotado por Ω.
- Eventos: subconjuntos de Ω, denotado por $A, B, \ldots$.
- Conjunto vazio: conjunto sem eventos, denotado por $\emptyset$.
- União $A \cup B$: ocorrência de pelo menos um dos eventos A ou B.
- Intersecção $A \cap B$: ocorrência simultânea de A e B.
- Eventos disjuntos ou mutuamente exclusivos: $A \cap B = \emptyset$.
- Eventos complementares: $A \cup A^c = \Omega$ e $A \cap A^c = \emptyset$.

Exemplos: Definições básicas

- Fenômeno aleatório: jogar um dado e verificar a face superior.
- Espaço amostral: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Eventos: $A =$ face é par, $B =$ face é ímpar, $C > 3$.
- União: $A \cup B = \Omega, A \cup C = \{2, 4, 6, 5\}$.
- Intersecção: $A \cap B = \emptyset, A \cap C = \{4, 6\}$.
- Eventos complementares: $A^c = B$.

Definição de probabilidade

- Probabilidade é uma função $P(\cdot)$ que atribui valores numéricos aos eventos do espaço amostral, de tal forma que
  i) $0 \leq P(A) \leq 1, \quad \forall A \in \Omega$;
  ii) $P(\Omega) = 1$;
  iii) $P(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} P(A_j)$, com os A_j's disjuntos.
- Exemplos triviais: lançamento de uma moeda e lançamento de um dado.
- Regra da adição de probabilidades. Sejam A e B eventos em Ω. Então

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

- Mostre que

$$P(A) = 1 - P(A^c).$$

Probabilidade condicional

- Definição: Dados dois eventos A e B, a probabilidade condicional de A dado que ocorreu B é representado por $P(A|B)$ e dada por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{para} \quad P(B) > 0.$$

- Caso $P(B) = 0$ definimos $P(A|B) = P(A)$.
- Regra do produto: Sejam A e B eventos em Ω, Então

$$P(A \cap B) = P(A|B)P(B), \quad \text{com} \quad P(B) > 0.$$

- Exemplo usando dado.

Independência de eventos

- Definição: Dois eventos A e B são independentes, se a informação da ocorrência ou não de B não altera a probabilidade de ocorrência de A. Isto é,

$$P(A|B) = P(A), \quad P(B) > 0,$$

ou ainda da seguinte forma

$$P(A \cap B) = P(A)P(B).$$

Exemplo 1: Probabilidade Condicional e Independência

Uma empresa produz peças em duas máquinas I e II, que podem apresentar desajustes com probabilidade 0.05 e 0.10, respectivamente. No início do dia de operação um teste é realizado e, caso a máquina esteja fora de ajuste, ela ficará sem operar nesse dia passando por revisão técnica. Para cumprir o nível mínimo de produção pelo menos uma das máquinas deve operar. Você diria que a empresa corre o risco de não cumprir com suas metas de produção?

Exemplo 2: Probabilidade Condicional e Independência

Suponha que a probabilidade de um avião decolar no horário é de $P(D) = 0.83$; a probabilidade do avião chegar no horário é de $P(A) = 0.82$; e a probabilidade de decolar e chegar no horário é de $P(D \cap A) = 0.78$. Encontre as probabilidades:

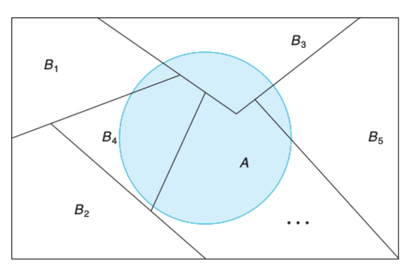
a) Chegue no horário dado que decolou no horário.

b) Decole no horário dado que chegou no horário.

Partição do espaço amostral

- Definição: Os eventos $C_1, C_2, \ldots, C_k$ formam uma partição do espaço amostral, se eles não tem intersecção entre si e se sua união é igual ao espaço amostral. Isto é,

$$C_i \cap C_j = \emptyset \quad \text{para} \quad i \neq j \quad \text{e} \quad \bigcup_{i=1}^{k} C_i = \Omega.$$



Teorema de Bayes

- Suponha que os eventos $C_1, C_2, \ldots, C_k$ formem uma partição de Ω e que suas probabilidades sejam conhecidas. Suponha, ainda, que para um evento A, se conheçam as probabilidades $P(A|C_i)$ para todos $i = 1, 2, \ldots, k$. Então, para qualquer j,

$$P(C_j|A) = \frac{P(A|C_j)P(C_j)}{\sum_{i=1}^{k} P(A|C_i)P(C_i)}, \quad j = 1, 2, \ldots, k.$$

- Demonstração.

Exemplo 1: Teorema de Bayes

Uma montadora trabalha com 2 fornecedores (A e B) de uma determinada peça. As chances de que uma peça proveniente dos fornecedores A e B esteja fora das especificações são 10% e 5% respectivamente. A montadora recebe 30% das peças do fornecedor A e 70% de B. Se uma peça do estoque inteiro é escolhido ao acaso:

a) Calcule a probabilidade de que ela esteja fora das especificações.
b) Se uma peça escolhida ao acaso está fora das especificações, qual é a probabilidade que venha do fornecedor fornecedor A?

Exemplo 2: Teorema de Bayes

Uma empresa de manufatura emprega três planos para o desenvolvimento de um particupar produto. Por razões de custos, os três planos são usados aleatóriamente. Dados históricos mostram que os planos 1, 2 e 3 são usados para 30%, 20% e 50% dos produtos, respectivamente. A taxa de defeito é diferente para os três planos, sendo

$$P(D|P_1) = 0.01, P(D|P_2) = 0.03 \quad \text{e} \quad P(D|P_3) = 0.02.$$

Se um produto é observado aleatóriamente e verificado defeituoso, qual é o plano mais provável de ter sido usado em sua produção?

Variáveis aleatórias

Definição e exemplos

- **Variável aleatória** - Descrição numérica do resultado de um experimento.
- Notação: Y denota a variável aleatória, enquanto que y denota os valores realizados de uma variável aleatória.
- Exemplos:
  - Duas bolas são retiradas de uma urna sucessivamente sem reposição. Na urna tem-se 4 bolas vermelhas e 3 pretas. Define a v.a Y - número de bolas vermelhas.
  - Em uma linha de produção peças são avaliadas sobre sua adequação a uma dada norma. Defina a v.a.

$$Y = \begin{cases} 1, & \text{se o componente é defeituoso} \\ 0, & \text{se o componente é não defeitusoso} \end{cases}$$

  - Número de dias até a entrega de um produto.
  - Tamanho, peso, diâmetro de um componente elétrico.
  - etc ...

Tipos de dados, espaço amostral e v.a

- Tipos de dados
  1. Dados na reta real, $\Omega = \Re$.
  2. Dados estritamente positivos, $\Omega = \Re_+$.
  3. Dados positivos com zeros, $\Omega = \Re_0 = [0, \infty)$.
  4. Proporções, $\Omega = (0, 1)$.
  5. Direções, $\Omega = [0, 2\pi)$.
  6. Contagens, $\Omega = \mathbb{N}_0 = \{0, 1, 2, \ldots\}$.
  7. Binomial, $\Omega == \{0, 1, 2, \ldots, m\}$.
- Tipos de espaço amostral
  1. **Espaço amostral Discreto**: Contêm apenas um número finito ou contável de elementos.
  2. **Espaço amostral Contínuo**: Contêm um número infinito de elementos.
- Tipos de variáveis aleatórias
  1. Variável aleatória é **contínua** se seu espaço amostral é contínuo.
  2. Variável aleatória é **dicreta** se seu espaço amostral é discreto.

Distribuição de Probabilidade

- Definição: O conjunto de pares $(y, f(y))$ é uma função de probabilidade de uma v.a discreta Y se para cada possível valor $y, f(y)$ satisfaz
  1. $f(y) \geq 0$.
  2. $\sum_y f(y) = 1$.
  3. $P(Y = y) = f(y)$.
- A função de probabilidade atribui a cada possível valor de no espaço amostral de Y a probabilidade de sua ocorrência.
- Cada probabilidade deve estar entre 0 e 1.
- A soma de todas as possíveis as possibilidades deve somar 1.
- Exemplo: Continuar exemplo da urna.

Função de Distribuição acumulada

- Definição: A função de distribuição acumulada de uma v.a Y com função de probabilidade $f(y)$ é dada por

$$F(y) = P(Y \leq y) = \sum_{t \leq y} f(t), \quad \text{para} \quad -\infty < y < \infty.$$

- Exemplo: Continuar exemplo da urna.

Exemplo: Função distribuição de probabilidade e acumulada

- Uma agência de carro vende 50% de seus veículos equipado com airbags. Encontre a distribuição de probabilidade e distribuição acumulada do número de carros vendidos com airbags para os próximos 4 carros vendidos por esta agência.
- Obtendo todas as possíveis combinações de com airbag (1) e sem airbag (2).

```
require(gtools)
sample_space <- permutations(n= 2, r = 4, repeats.allowed = TRUE)
head(sample_space)

##      [,1] [,2] [,3] [,4]
## [1,]    1    1    1    1
## [2,]    1    1    1    2
## [3,]    1    1    2    1
## [4,]    1    1    2    2
## [5,]    1    2    1    1
## [6,]    1    2    1    2
```

Exemplo: Função distribuição de probabilidade e acumulada

- Contando quantos carros vendidos com airbag para cada possível ocorrência

```
count = apply(sample_space, 1, function(x) sum(x == 1))
count

##  [1] 4 3 3 2 3 2 2 1 3 2 2 1 2 1 1 0
```

- Note cada combinação de eventos tem a mesma chance de ocorrer, ou seja, $1/16$.
- Assim, podemos multiplicar o número de ocorrência de cada evento pela sua probabilidade e obter a função distribuição de probabilidade.

```
fp <- table(count)*(1/16)
fp

## count
##      0      1      2      3      4
## 0.0625 0.2500 0.3750 0.2500 0.0625
```

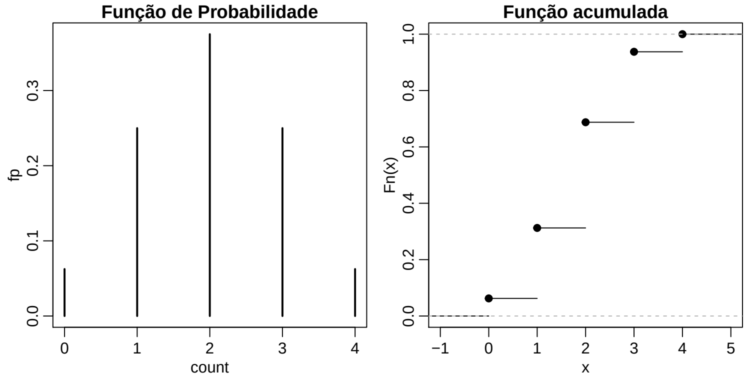
- Função de distribuição acumulada

```
cumsum(fp)

##      0      1      2      3      4
## 0.0625 0.3125 0.6875 0.9375 1.0000
```

Exemplo: Função distribuição de probabilidade e acumulada

- Graficamente, tem-se



Função densidade probabilidade

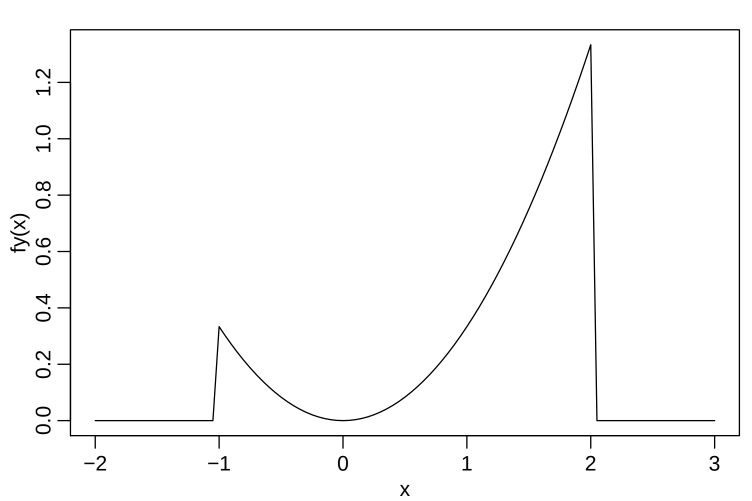
- Definição: A função $f(y)$ é a função densidade probabilidade de uma v.a contínua Y, definida sobre o conjunto dos números reais, se
  1. $f(y) \geq 0$, para todo $y \in \Re$.
  2. $\int_{-\infty}^{\infty} f(y)dy = 1$.
  3. $P(a < Y < b) = \int_a^b f(y)dy$.
- Exemplo: Suponha que o erro na temperatura em, graus Celsius, de um experimento laboratorial controlado é uma v.a contínua Y com função densidade probabilidade

$$f(y) = \begin{cases} \frac{y^2}{3}, & -1 < y < 2, \\ 0 & \text{caso contrário} \end{cases}$$

  a) Verifique que $f(y)$ é uma função de densidade.
  b) Encontre $P(0 < Y \leq 1)$.

Exemplo: Função densidade probabilidade

- Gráfico da densidade



Exemplo: Função densidade probabilidade

- Verificando se a função integra 1 no intervalo $(-1, 2)$.

```r
fy <- function(y) {
  out = (y^2)/3
  if(y < -1 | y > 2) {out <- 0}
  return(out)
}
# Integra 1?
integrate(fy, lower = -1, upper = 2) ## Ok

## 1 with absolute error < 1.1e-14
```

- Obtendo a probabilidade desejada.

```r
# P(0 < Y <= 1)
integrate(fy, lower = 0, upper = 1)

## 0.1111111 with absolute error < 1.2e-15
```

Função de distribuição acumulada

- Definição: A **função de distribuição acumulada** $F(y)$ de uma v.a. contínua Y com densidade $f(y)$ é

$$F(y) = P(Y \leq y) = \int_{-\infty}^{y} f(t)dt, \quad \text{para} \quad -\infty < y < \infty.$$

- De imediato, tem-se

$$P(a < Y < b) = F(b) - F(a) \quad \text{e} \quad f(y) = \frac{dF(y)}{dy},$$

desde que a derivada exista.

Exemplo: Função de distribuição acumulada

- Para a densidade do exemplo anterior, tem-se

$$F(y) = \int_{-\infty}^{y} f(t)dt = \int_{-1}^{y} \frac{t^2}{3} = \frac{t^3}{9}\Big|_{-1}^{y} = \frac{y^3+1}{9},$$

- Consequentemente, tem-se

$$F(y) = \begin{cases} 0, & y < -1, \\ \frac{y^3+1}{9}, & -1 < y < 2, \\ 1, & y \geq 2. \end{cases}$$

- Obtendo $P(0 < Y < 1)$

$$F(1) - F(0) = \frac{1^3+1}{9} - \frac{0^3+1}{9} = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}.$$

Distribuição de probabilidade conjunta

- Nosso estudo até aqui se restringiu a apenas uma v.a.
- Na prática, muitas vezes estamos interessados no comportamento **conjunto** de mais de uma v.a.
- Sendo Y e X duas v.a, tem-se as seguintes situações
    1. Y e X são discretas $\rightarrow$ conjunta será discreta.
    2. Y e X são contínuas $\rightarrow$ conjunta será contínua.
    3. Y é discreta e X é contínua $\rightarrow$ conjunta será mista.

Distribuição de probabilidade conjunta v.a discreta

- Definição: A função $f(x, y)$ é a função de distribuição conjunta das v.a. discretas X e Y, se
  1. $f(x, y) \geq 0 \quad \text{para todo} \quad (x, y)$.
  2. $\sum_x \sum_y f(x, y) = 1$.
  3. $P(X = x, Y = y) = f(x, y)$.

  Para qualquer região A no plano x e y,
  $P[(X, Y) \in A] = \sum \sum_A f(x, y)$.
- Definição: A distribuição marginal de X e Y são dadas por

$$f(x) = \sum_y f(x, y) \quad \text{e} \quad f(y) = \sum_x f(x, y).$$

Exemplo: Distribuição conjunta

- Exemplo: Considere o experimento de retirar duas bolas aleatóriamente de uma caixa contendo 3 bolas azuis, 2 bolas vermelhas e 3 bolas verdes. Define as variáveis X número de bolas azuis e Y número de bolas vermelhas selecionadas, encontre
  a) A função de distribuição conjunta $f(x, y)$.
  b) Obtenha as distribuições marginais de X e Y.
  c) $P[(X, Y) \in A]$, onde A é a região $\{(x, y) | x + y \leq 1\}$.

Exemplo: Distribuição conjunta

- Temos 8 bolas na urna das quais vamos selecionar 2.
- Total de formas de seleção é $\binom{8}{2} = 28$.
- Emulando o experimento

```
Urna <- c(rep("Azul", 3), rep("Vermelho",2), rep("Verde", 3))
Urna[sample(1:8, size = 2, replace = FALSE)]

## [1] "Verde" "Azul"

Urna[sample(1:8, size = 2, replace = FALSE)]

## [1] "Vermelho" "Azul"
```

- Obtendo todas as combinações possíveis

```
require(gtools)
idx <- combn(8, 2)
resul <- list()
for(i in 1:28){resul[[i]] <- Urna[idx[,i]]}
resul <- do.call(rbind, resul)
head(resul, 3)

##      [,1]   [,2]
## [1,] "Azul" "Azul"
## [2,] "Azul" "Azul"
## [3,] "Azul" "Vermelho"
```

Exemplo: Distribuição conjunta

- Para cada uma das possíveis combinações vamos obter o número de bolas azuis X e o número de bolas vermelhas Y.

```
Azul <- c()
Vermelha <- c()
for(i in 1:28) {
  Azul[i] <- sum(resul[i,] == "Azul")
  Vermelha[i] <- sum(resul[i,] == "Vermelho")
}
```

- Distribuição conjunta.

```
joint = MASS::as.fractions(table(Vermelha, Azul)/28)
joint

##          Azul
## Vermelha 0    1    2
##        0 3/28 9/28 3/28
##        1 3/14 3/14    0
##        2 1/28    0    0
```

- Distribuições marginais.

```
# Marginal AZUL
MASS::as.fractions(colSums(joint))

##     0     1     2
##  5/14 15/28  3/28

# Marginal VERMELHA
MASS::as.fractions(rowSums(joint))

##     0     1     2
## 15/28   3/7  1/28
```

Função de densidade conjunta

- Definição: A função $f(x, y)$ é a **função de densidade conjunta** das v.a. contínuas X e Y se
  1. $f(x, y) \geq 0, \quad \text{para todo} \quad (x, y).$
  2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) = 1.$
  3. $P[(X, Y) \in A] = \int \int_A f(x, y) dxdy$, para qualquer região A no plano xy.
- Distribuições marginais

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{e} \quad f(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Exemplo: Função de densidade conjunta

- Uma empresa atua com dois métodos para produzir um certo produto. Para um certo dia de operação, sejam X e Y, as proporções de produtos produzidos pelos métodos A e B, respectivamente. Suponha, que a distribuição conjunta de X e Y é dada por

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0 & \text{caso contrário.} \end{cases}$$

1. Verifique se $f(x,y)$ é uma distribuição de probabilidade.
2. Encontre as distribuições marginais de X e Y.
3. Encontre a probabilidade $P[(X,Y) \in A]$, onde $A = \{(x,y) | 0 < x < 0.5, 0.25 < y < 0.5\}$.

Exemplo: Função de densidade conjunta

- Verificar (1) é trivial.
- Verificando (2).

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{2}{5}(2x+3y)dxdy = \int_0^1\int_0^1\left(\frac{4x}{5}+\frac{6y}{5}\right)dxdy$$

$$\int_0^1\frac{6y}{5}+\int_0^1\frac{4x}{5}dxdy = \int_0^1\frac{6y}{5}+\frac{4}{5}\left[\frac{x^2}{2}\right]_0^1 dy = \int_0^1\frac{6y}{5}+\frac{4}{10}dy$$

$$\frac{6}{5}\int_0^1 ydy + \frac{4}{10}\int_0^1 1dy = \frac{6}{5}\left[\frac{y^2}{2}\right]_0^1 + \frac{4}{10}[y]_0^1 = \frac{6}{10}+\frac{4}{10} = 1$$

Exemplo: Função de densidade marginal

- Obtendo as marginais

$$f(x) = \int_0^1 \frac{4x}{5} + \frac{6y}{5} dy = \left[\frac{4xy}{5} + \frac{6y^2}{10}\right]_0^1 = \frac{4x+3}{5},$$

para $0 \leq x \leq 1$, e 0 caso contrário.

$$f(y) = \int_0^1 \frac{4x}{5} + \frac{6y}{5} dx = \frac{2+6y}{5}.$$

para $0 \leq y \leq 1$, e 0 caso contrário.

- Obtendo a probabilidade

$$P[(X,Y) \in A] = \int_{\frac{1}{4}}^{\frac{1}{2}} \int_{\frac{1}{4}}^{\frac{1}{2}} \frac{4x}{5} + \frac{6y}{5} dxdy = \frac{13}{160}.$$

Distribuição condicional

- Sejam X e Y duas v.a. discretas ou contínuas. A **distribuição condicional** de uma v.a. Y dado $X = x$ é

$$f(y|x) = \frac{f(x,y)}{f(x)}, \quad \text{dado que} \quad f(x) > 0.$$

$$f(x|y) = \frac{f(x,y)}{f(y)}, \quad \text{dado que} \quad f(y) > 0.$$

- Probabilidades condicionais

$$P(a < X < b|Y = y) = \sum_{a<x<b} f(x|y).$$

$$P(a < X < b|Y = y) = \int_a^b f(x|y) dx.$$

Independência estatística

- Sejam X e Y duas v.a. discretas ou contínuas, com distribuição de probabilidade conjunta $f(x, y)$ e distribuições marginais $f(x)$ e $f(y)$. As v.a. X e Y são ditas **estatisticamente independentes** se e somente se

$$f(x, y) = f(x)f(y)$$

para todo par (x, y) in seus respectivos domínios.

Independência estatística - Generalização

- Sejam $Y_1, Y_2, \ldots, Y_n$ n v.a. discretas ou contínuas, com função de distribuição conjunta $f(y_1, y_2, \ldots, y_n)$ e distribuições marginais $f(y_1), f(y_2), \ldots, f_(y_n)$. As v.a. $Y_1, Y_2, \ldots, Y_n$ são ditas **mutuamente independentes** se e somente se

$$f(y_1, y_2, \ldots, y_n) = f(y_1)f(y_2)\ldots f(y_n),$$

para todo $(y_1, y_2, \ldots, y_n)$ em seus respectivos domínios.

Exemplo: Independência estatística

- Suponha que o tempo de prateleira, em anos, de um certo produto perecível é uma variável aleatória com distribuição de probabilidade dada por

$$f(y) = \begin{cases} e^{-y}, & y > 0 \\ 0 & \text{caso contrário.} \end{cases}$$

Sejam Y_1, Y_2 e Y_3 o tempo de prateleira de três destes produtos selecionados independentemente. Encontre $P(X1 < 2, 1 < X_2 < 3, X_3 > 2)$.

Exemplo: Independência estatística

- Usando a independência, tem-se

$$f(y_1, y_2, y_3) = f(y_1)f(y_2)f(y_3) = e^{-y_1}e^{-y_2}e^{-y_3} = e^{-y_1-y_2-y_3}$$

$$P(Y_1 < 2, 1 < Y_2 < 3, Y_3 > 2) = \int_2^{\infty}\int_1^{3}\int_0^{2} e^{-y_1-y_2-y_3} dy_1 dy_2 dy_3 = 0.0372.$$

Pontos importantes

- Como construímos a distribuição de probabilidade para um fenômeno?
  1. Por observação do fenômeno por longos períodos.
  2. Natureza do fenômeno sugere uma certa forma de distribuição.
  3. Em geral não sabemos qual distribuição de probabilidade pode ser uma boa escolha para um dado fenômeno aleatório.
  4. Existem uma grande quantidade de distribuições de probabilidade que podem ser avaliadas como uma boa aproximação para o fenômeno em estudo.
  5. Vamos ver as mais importantes:
     - 5.1 Bernoulli, binomial e Poisson (caso discreto).
     - 5.2 Normal, exponencial e Gama (caso contínuo).
- Mas antes precisamos entender como resumir a informação gerada pela distribuição de probabilidades.

Esperança matemática

Esperança matemática

- Seja Y uma v.a. com distribuição de probabilidade $f(y)$. A **média** ou **valor esperado** de X é dado por
  1. Caso discreto: $\mu = E(Y) = \sum_y yf(y)$.
  2. Caso contínuo: $\mu = E(Y) = \int_{-\infty}^{\infty} yf(y)dy$.
- Exemplo: Um vendedor tem duas reunião de vendas em um dado dia. Na primeira reunião ele acredita ter 70% de chance de fazer uma venda que lhe renderá $R\$1000$. Na segunda ele acredita ter 40% de chance de fazer uma venda que se realizada lhe renderá $R\$1500$. Assuma que as vendas são independentes. Quanto de comissão ele esperava ganhar neste dia?
- Solução: Defina Y como sendo a v.a. comissão.
- O espaço amostral de Y é $\Omega = \{0, 1000, 1500, 2500\}$.

Esperança matemática

- A distribuição de probabilidade de Y é

y	0	1000	1500	2500
$P(Y = y)$	$(1-0.7)(1-0.4) = 0.18$	$0.7(1-0.4) = 0.42$	$(1-0.7)0.4 = 0.12$	$0.70.4 = 0.28$

- Usando a definição o valor esperado é

$$E(Y) = 0.18*0+0.42*1000+0.12*1500+0.28*2500 = 1300.$$

- Exemplo 2: Seja Y uma v.a. que mede o tempo de vida de um dispositivo eletrônico em horas. Considere a seguinte fdp para Y

$$f(y) = \begin{cases} 20000/y^3 & y > 100, \\ 0, & \text{caso contrário.} \end{cases}$$

Qual o tempo de vida esperado deste dispositivo?

$$\mu = E(Y) = \int_{100}^{\infty} y \frac{20000}{y^3} dy = 200.$$

Esperança matemática

- Sejam X e Y v.a. com função de distribuição conjunta $f(x, y)$. O valor esperado da variável $g(X, Y)$ é
  1. Caso discreto:
  $$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_x \sum_y f(x, y).$$
  2. Caso contínuo:
  $$\mu_{g(X,Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy.$$

Exemplo: Esperança matemática

- Sejam as v.a. discretas X e Y com fp dada por

```
joint

##          Azul
## Vermelha 0    1    2
##        0 3/28 9/28 3/28
##        1 3/14 3/14    0
##        2 1/28    0    0
```

Encontre a $E(XY)$.

$$E(XY) = \sum_x \sum_y xyf(x,y) = \frac{3}{14}$$

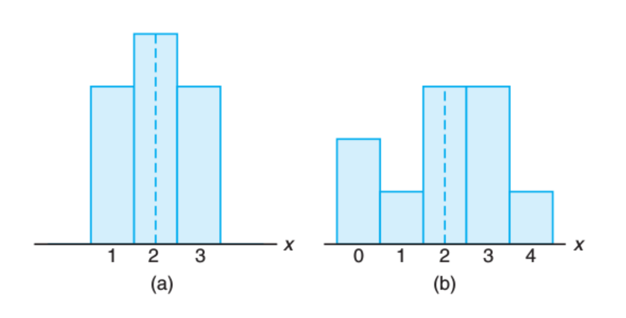
Exemplo: Esperança matemática

- Fazendo as contas

X	Y	XY	P(X,Y)	XY P(X,Y)
0	0	0	3/28	0
1	0	0	9/28	0
2	0	0	3/28	0
0	1	0	3/14	0
1	1	1	3/14	3/14
2	1	2	0	0
0	2	0	1/28	0
1	2	2	0	0
2	2	4	0	0

Variância de v.a.

- Esperança descreve o centro de massa da distribuição de probabilidade.
- Esperança não informa como a v.a se distribui em torno da esperança.
- Precisamos de uma medida de dispersão.



Variância de v.a.

- Definição: Seja Y uma v.a. com função de probabilidade $f(y)$ e esperança μ. A variância de Y é dada por
  - Caso discreto:
$$\sigma^2 = E[(Y-\mu)^2] = \sum_y (y-\mu)^2 f(y).$$
  - Caso contínuo:
$$\sigma^2 = E[(Y-\mu)^2] = \int_{-\infty}^{\infty} (y-\mu)^2 f(y).$$
- A raiz quadrada positiva da variância é chamada de **desvio padrão** de Y.
- Forma alternativa:
$$\sigma^2 = E(Y^2) - \mu^2.$$

Exemplo: Variância de v.a.

- A demanda semanal de água potável, em milhares de litros, de uma cadeia de lojas é uma v.a. contínua com fdp dada por

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

- Encontra $E(Y)$ e $V(Y)$.

$$E(Y) = \int_1^2 y f(y) dy = 2\int_1^2 y(y-1)dy = \frac{5}{3}.$$

$$E(Y^2) = \int_1^2 y^2 f(y) dy = 2\int_1^2 y^2(y-1)dy = \frac{17}{6}.$$

$$V(Y) = E(Y^2) - E^2(Y) = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}.$$

Covariância

- Definição: Sejam X e Y v.a. com função de probabilidade conjunta $f(x, y)$. A **covariância** entre X e Y é dada por
  1. Caso discreto:
  $$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y).$$
  2. Caso contínuo:
  $$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dxdy.$$
- Covariância é uma medida da associação entre as v.a. X e Y.
- Forma alternativa

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

Correlação

- Sejam X e Y v.a. com covariância σ_{XY} e desvio padrão σ_X e σ_Y. A **correlação** entre X e Y é dada por

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}, \quad -1 \leq \rho_{XY} \leq 1.$$

- A correlação é uma medida livre das unidades de X e Y.
- Cuidado!! Correlação 0 não significa necessariamente independência estatística.

Exemplo: Covariância e Correlação

- Considere a seguinte distribuição conjunta.

		x			
f(x,y)		0	1	2	f(y)
	0	3/28	9/28	3/28	15/28
y	1	3/14	3/14	0	3/7
	2	1/28	0	0	1/28
	f(x)	5/14	15/28	3/28	1

- Obtenha σ_{xy} e ρ_{xy}.

Exemplo: Covariância e Correlação

- Esperanças marginais

$$E(X) = \mu_X = \sum_{x=0}^{2} xf(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{3}{4}.$$

$$E(Y) = \mu_Y = \sum_{y=0}^{2} yf(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) = \frac{1}{2}.$$

- Esperanças do quadrado

$$E(X^2) = \sum_{x=0}^{2} x^2 f(x) = (0^2)\left(\frac{5}{14}\right) + (1^2)\left(\frac{15}{28}\right) + (2^2)\left(\frac{3}{28}\right) = \frac{27}{28}.$$

$$E(Y^2) = \sum_{y=0}^{2} y^2 f(y) = (0^2)\left(\frac{15}{28}\right) + (1^2)\left(\frac{3}{7}\right) + (2^2)\left(\frac{1}{28}\right) = \frac{4}{7}.$$

Exemplo: Covariância e Correlação

- Esperança do produto (calculamos antes slide 49).

$$E(XY) = \frac{3}{14}.$$

- Variâncias marginais

$$V(X) = E(X^2) - E^2(X) = \frac{27}{28} - \left(\frac{3}{4}\right)^2 = \frac{45}{112}.$$

$$V(Y) = E(Y^2) - E^2(Y) = \frac{4}{7} - \left(\frac{1}{2}\right)^2 = \frac{9}{28}.$$

Exemplo: Covariância e Correlação

- Covariância

$$\sigma_{XY} = E(XY) - E^2(X)E^2(Y) = \frac{3}{14} - \left(\frac{3}{4}\right)^2 \left(\frac{1}{2}\right)^2 = -\frac{9}{56}.$$

- Correlação

$$\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2}\sqrt{\sigma_Y^2}} = \frac{-9/56}{\sqrt{(45/112)(9/28)}} = -\frac{1}{\sqrt{5}}.$$

Propriedades da Esperança e Variância de combinação de v.a.

- Sendo a e b constantes, X e Y v.a. e $g(\cdot)$ e $h(\cdot)$ funções lineares são válidas:
  1. $E[aX + b] = aE[X] + b$.
  2. $E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$.
  3. $E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)]$.
  4. $E[XY] = E[X]E[Y]$ se e somente se X e y são independentes.
  5. $V[aX + bY + c] = a^2V[X] + b^2V[Y] + 2abCOV(X, Y)$.
  6. $V[aX - bY] = a^2V[X] + b^2V[Y] - 2abCOV(X, Y)$.
- Sendo $g(\cdot)$ uma função não linear de X, tem-se as aproximações

$$E[g(X)] \approx g(\mu_X) + \frac{\partial^2 g(x)}{\partial x^2}|_{x=\mu_X} \frac{\sigma_X^2}{2}.$$

$$V[g(X)] \approx \left[\frac{\partial g(x)}{\partial x}\right]^2 |_{x=\mu_X} \sigma_X^2.$$