Statistics for Business and Economics

7th Edition

Aula 10 - Capítulo 5

Variáveis Aleatórias Contínuas e Distribuições de Probabilidade

5.4 Aproximação de distribuição normal para distribuição binomial

- Lembre-se da distribuição binomial :
  - n ensaios independentes
  - probabilidade de sucesso em qualquer tentativa dada = P

- Variável aleatória X :
  - X_i =1 se a i^{esima} tentativa é “sucesso”
  - X_i =0 se a i^{esima} tentative é “fracasso”

$$\boxed{E(X) = \mu = nP}$$

$$\boxed{Var(X) = \sigma^2 = nP(1-P)}$$

Aproximação de distribuição normal para distribuição binomial

(cont.)

- A forma da distribuição binomial é aproximadamente normal se n for grande
- A normal é uma boa aproximação do binômio quando $nP(1 – P) > 5$
- Padronize para Z a partir de uma distribuição binomial:

$$Z = \frac{X - E(X)}{\sqrt{Var(X)}} = \frac{X - np}{\sqrt{nP(1-P)}}$$

Aproximação de distribuição normal para distribuição binomial

(cont.)

- Seja X o número de sucessos de n tentativas independentes, cada uma com probabilidade de sucesso P.
- Se $nP(1 - P) > 5$,

$$P(a < X < b) = P\left(\frac{a - nP}{\sqrt{nP(1-P)}} \leq Z \leq \frac{b - nP}{\sqrt{nP(1-P)}} \right)$$

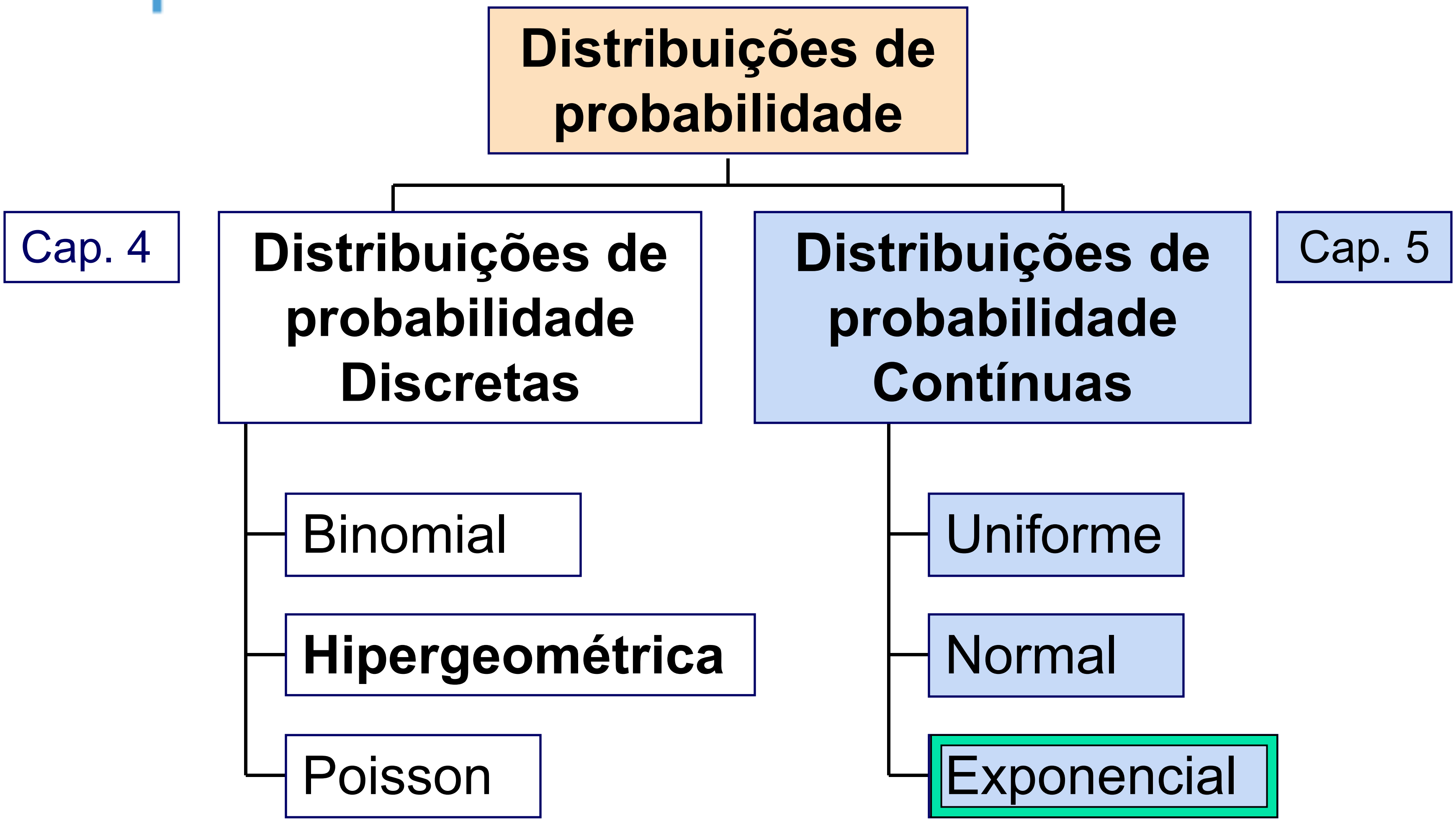
Exemplo de Aproximação Binomial

- 40% de todos os eleitores apóiam a proposição A. Qual é a probabilidade de que entre 76 e 80 eleitores indiquem apoio em uma amostra de n = 200?
  - $E(X) = \mu = nP = 200(0.40) = 80$
  - $Var(X) = \sigma^2 = nP(1 - P) = 200(0.40)(1 - 0.40) = 48$

    (note: $nP(1 - P) = 48 > 5$)

$$\begin{aligned}
P(76 < X < 80) &= P\left(\frac{76-80}{\sqrt{200(0.4)(1-0.4)}} \le Z \le \frac{80-80}{\sqrt{200(0.4)(1-0.4)}}\right) \\
&= P(-0.58 < Z < 0) \\
&= F(0) - F(-0.58) \\
&= 0.5000 - 0.2810 = 0.2190
\end{aligned}$$

A Distribuição Exponencial



A Distribuição Exponencial

- Usada para modelar o período de tempo entre duas ocorrências de um evento (o tempo entre as chegadas)
  - Exemplos:
    - Tempo entre caminhões chegando a uma doca de descarga
    - Tempo entre transações em um caixa eletrônico
    - Tempo entre chamadas telefónicas para o operador principal

A Distribuição Exponencial

(continued)

- A variável aleatória exponencial T (t>0) tem uma função de densidade de probabilidade

$$\boxed{f(t) = \lambda e^{-\lambda t} \quad \text{for } t > 0}$$

- Sendo que
  - λ é o número médio de ocorrências por unidade de tempo
  - t é o número de unidades de tempo até a próxima ocorrência
  - e = 2.71828
- T segue uma distribuição de probabilidade exponencial

A Distribuição Exponencial

- Definido por um único parâmetro, sua média λ (lambda)
- A função de distribuição cumulativa (a probabilidade de que um tempo de chegada seja menor que algum tempo especificado t) é

$$\boxed{F(t) = 1 - e^{-\lambda t}}$$

onde e = constante matemática aproximada por 2,71828

λ = número médio populacional de chegadas por unidade

t = qualquer valor da variável contínua onde $t > 0$

Exemplo de Distribuição Exponencial

Exemplo: Os clientes chegam ao balcão de atendimento à taxa de 15 por hora. Qual é a probabilidade de que o tempo de chegada entre clientes consecutivos seja inferior a três minutos?

- O número médio de chegadas por hora é de 15, então $\lambda = 15$
- Três minutos são 0,05 horas
- P(tempo de chegada < .05) $= 1 – e^{-\lambda X} = 1 – e^{-(15)(.05)} =$ $\boxed{0.5276}$
- Portanto, há uma probabilidade de 52,76% de que o tempo de chegada entre clientes sucessivos seja inferior a três minutos

5.6 Funções de Distribuição Acumulada Conjunta

- Seja $X_1, X_2, \ldots X_k$ variáveis aleatórias contínuas
- Sua função de distribuição acumulada conjunta,

  $F(x_1, x_2, \ldots x_k)$

  define a probabilidade que simultaneamente X_1 seja menor que x_1, X_2 menor que x_2, e assim por diante; ou

$$\boxed{F(x_1, x_2, \ldots, x_k) = P(X_1 < x_1 \cap X_2 < x_2 \cap \cdots X_k < x_k)}$$

Funções de Distribuição Acumulada Conjunta

(cont.)

- As funções de distribuição acumuladas

  $$F(x_1), F(x_2), \ldots, F(x_k)$$

  das variáveis aleatórias individuais são chamadas de suas funções de distribuição marginal

- As variáveis aleatórias são independentes se e somente se

$$\boxed{F(x_1, x_2, \ldots, x_k) = F(x_1)F(x_2)\cdots F(x_k)}$$

Covariância

- Sejam X e Y variáveis aleatórias contínuas, com médias μ_x e μ_y

- O valor esperado de $(X - \mu_x)(Y - \mu_y)$ é chamado de covariância entre X e Y

$$\boxed{\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]}$$

- Uma expressão alternativa, mas equivalente, é

$$\boxed{\text{Cov}(X, Y) = E(XY) - \mu_x \mu_y}$$

- Se as variáveis aleatórias X e Y forem independentes, então a covariância entre elas é 0. No entanto, o inverso não é verdadeiro.

Correlação

- Sejam X e Y variáveis aleatórias distribuídas conjuntamente.

- A correlação entre X e Y é

$$\boxed{\rho = \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}}$$

Somas de Variáveis Aleatórias

Sejam $X_1, X_2, \ldots X_k$ k variáveis aleatórias com médias $\mu_1, \mu_2, \ldots \mu_k$ e variâncias $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$. Então:

- A média de sua soma é a soma de suas médias

$$\boxed{E(X_1 + X_2 + \cdots + X_k) = \mu_1 + \mu_2 + \cdots + \mu_k}$$

Somas de Variáveis Aleatórias

(cont.)

Sejam $X_1, X_2, \dots X_k$ k variáveis aleatórias com médias $\mu_1, \mu_2, \dots \mu_k$ e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$. Então :

- Se a covariância entre cada par dessas variáveis aleatórias for 0, então a variância de sua soma é a soma de suas variâncias

$$\boxed{\mathrm{Var}(X_1 + X_2 + \cdots + X_k) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2}$$

- No entanto, se as covariâncias entre pares de variáveis aleatórias não forem 0, a variância de sua

$$\boxed{\mathrm{Var}(X_1 + X_2 + \cdots + X_k) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2 + 2\sum_{i=1}^{K-1}\sum_{j=i+1}^{K}\mathrm{Cov}(X_i, X_j)}$$

Diferenças entre duas variáveis aleatórias

Para duas variáveis aleatórias, X e Y

A média de suas diferenças é a diferença de suas médias; aquilo é

$$E(X - Y) = \mu_X - \mu_Y$$

Se a covariância entre X e Y for 0, então a variância de sua diferença é

$$Var(X - Y) = \sigma_X^2 + \sigma_Y^2$$

Se a covariância entre X e Y não for 0, então a variância de sua diferença é

$$Var(X - Y) = \sigma_X^2 + \sigma_Y^2 - 2Cov(X, Y)$$

Combinações Lineares de Variáveis Aleatórias

- Uma combinação linear de duas variáveis aleatórias, X e Y, (onde a e b são constantes) é

$$\boxed{W = aX + bY}$$

- A média de W é

$$\boxed{\mu_W = E[W] = E[aX + bY] = a\mu_X + b\mu_Y}$$

Combinações Lineares de Variáveis Aleatórias

(cont.)

- A variância de W é

$$\boxed{\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\mathrm{Cov}(X, Y)}$$

- Ou usando a correlação,

$$\boxed{\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\mathrm{Corr}(X, Y)\sigma_X\sigma_Y}$$

- Se X e Y são variáveis aleatórias conjuntas normalmente distribuídas, então a combinação linear, W, também é normalmente distribuída

Exemplo

- Duas tarefas devem ser executadas pelo mesmo trabalhador.
  - X = minutos para completar a tarefa 1; $\mu_x = 20$, $\sigma_x = 5$
  - Y = minutos para completar a tarefa 2; $\mu_y = 30$, $\sigma_y = 8$
  - X e Y são normalmente distribuídos e independentes

Qual é a média e o desvio padrão do tempo para completar ambas as tarefas?

Exemplo

(cont.)

- X = minutos para completar a tarefa 1; $\mu_x = 20$, $\sigma_x = 5$
- Y = minutos para completar a tarefa 2; $\mu_y = 30$, $\sigma_y = 8$

- Qual é a média e o desvio padrão do tempo para completar ambas as tarefas?

$$\boxed{W = X + Y}$$

$$\boxed{\mu_W = \mu_X + \mu_Y = 20 + 30 = 50}$$

- Como X e Y são independentes, Cov(X,Y) = 0, então

$$\boxed{\sigma_W^2 = \sigma_X^2 + \sigma_Y^2 + 2\mathrm{Cov}(X,Y) = (5)^2 + (8)^2 = 89}$$

- O desvio padrão é

$$\boxed{\sigma_W = \sqrt{89} = 9.434}$$

Analise de portfólio

- Uma carteira financeira pode ser vista como uma combinação linear de instrumentos financeiros separados

$$\begin{pmatrix}\text{Return on}\\ \text{portfolio}\end{pmatrix} = \begin{pmatrix}\text{Proportion of}\\ \text{portfolio value}\\ \text{in stock1}\end{pmatrix} \times \begin{pmatrix}\text{Stock 1}\\ \text{return}\end{pmatrix} + \begin{pmatrix}\text{Proportion of}\\ \text{portfolio value}\\ \text{in stock2}\end{pmatrix} \times \begin{pmatrix}\text{Stock 2}\\ \text{return}\end{pmatrix}$$

$$\cdots + \begin{pmatrix}\text{Proportion of}\\ \text{portfolio value}\\ \text{in stock N}\end{pmatrix} \times \begin{pmatrix}\text{Stock N}\\ \text{return}\end{pmatrix}$$

Exemplo de Análise de Portfólio

- Considere duas ações, A e B
  - O preço da ação A tem distribuição normal com média 12 e variância 4
  - O preço da ação B tem distribuição normal com média 20 e variância 16
  - Os preços das ações têm uma correlação positiva, $\rho_{AB} = .50$
- Suponha que você possua
  - 10 ações do estoque A
  - 30 ações da Ação B

Exemplo de Análise de Portfólio

(continued)

- A média e a variância desta carteira de ações são: (Deixe W denotar a distribuição do valor da carteira)

$$\mu_W = 10\mu_A + 20\mu_B = (10)(12) + (30)(20) = 720$$

$$\begin{aligned}\sigma_W^2 &= 10^2\sigma_A^2 + 30^2\sigma_B^2 + (2)(10)(30)\text{Corr}(A,B)\sigma_A\sigma_B \\ &= 10^2\,(4)^2 + 30^2(16)^2 + (2)(10)(30)(.50)(4)(16) \\ &= 251{,}200\end{aligned}$$

Exemplo de Análise de Portfólio

(continued)

- Qual é a probabilidade de que o valor de sua carteira seja inferior a US$ 500?

$$\mu_W = 720$$

$$\sigma_W = \sqrt{251{,}200} = 501.20$$

- O Z valor para 500 é $$Z = \frac{500 - 720}{501.20} = -0.44$$

- P(Z < -0.44) = 0.3300
  - Portanto, a probabilidade é de 0,33 de que o valor do seu portfólio seja inferior a US$ 500..