MPEI 2015-2016

Aula 12

Exercício(s) de aplicação

Exercício

- Para ajudar com os conceitos das aulas anteriores…
- Considere duas variáveis aleatórias com a seguinte distribuição conjunta:

X/Y	0	1	2
0	0,3	0,2	0
1	0,1	0,15	0,05
2	0	0,1	0,1

- Responda às seguintes questões efectuando os cálculos no papel e confirmando usando o Matlab

…

- Calcule as funções de probabilidade de massa marginais de X e Y

- Calcule a média e variância de X e Y

- Calcule a correlação, covariância e coeficiente de correlação entre X e Y

…

- Calcule os momentos de ordem 2 e 3 para as variáveis X e Y

- Calcule os momentos de ordem 2 e 3 para a variável $Z = \sqrt{X}$

Variáveis Aleatórias em Situações Limite

Desigualdades de Markov e Chebyshev

Lei dos Grandes Números

Teorema do Limite Central

Média e variância da Média

- Se criarmos a v.a. relativa à média das n variáveis IID X_i, $\boldsymbol{M_n = \frac{S_n}{n}}$, e assumindo $E[X_i] = \mu$ e $Var(X_i) = \sigma^2$, teremos :

- $\boldsymbol{E[M_n]} = E\left[\frac{S_n}{n}\right] = \frac{\sum_i E[X_i]}{n} = E[X_i] = \boldsymbol{\mu}$

- $\mathbf{Var}[\boldsymbol{M_n}] = Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\frac{\sum_i Var[X_i]}{1} = \frac{Var(X_i)}{n} = \boldsymbol{\frac{\sigma^2}{n}}$
- À medida que se aumenta o número de experiências vai diminuindo a variância da estimativa da média

Questões

- Interessa-nos, por exemplo, saber **o que acontece quando n tende para infinito** ao valor esperado e variância da Média de n medições

- Colocam-se questões como:
  - Quão provável é termos a média das amostras superior a um determinado valor (exemplo: o dobro do valor esperado) ?
  - Quão próximo a média obtida com as amostras fica do valor médio ?
  - Qual a distribuição da média para valores de n muito grandes ?

Desigualdades de Markov e Chebyshev

- Os dois teoremas que apresentaremos de seguida, sem muita preocupação com demonstrações, permitem estabelecer facilmente majorantes para probabilidades de certas classes de acontecimentos
  - partindo apenas do conhecimento da média e variância de uma variável aleatória
  - Mais informação, por exemplo, na secção 3.8 do livro de F. Vaz

Questão 1

- Quão provável termos valores superiores a um determinado valor ?

- Exemplo:
- A média das classificações numa turma é 15,2.
- Qual o limite superior de probabilidade de um dos alunos ter nota igual ou superior a 17 ?

Desigualdade de Markov

- Seja X uma variável aleatória
- Pela Desigualdade de Markov:
- $P(X \geq a) \leq \frac{E[X]}{a},\ \forall a > 0$

- Esta desigualdade dá-nos um limite superior para a probabilidade de a função X ser maior ou igual a um determinado valor
- Qual o valor de P com a = E[X] ?
  - E a < E(X) ?
  - E a > E(X) ?

Desigualdade de Markov

- Demonstração:
- $E[X] =?$
- $= \cancel{\int_0^a x f_X(x) dx} + \int_a^\infty x f_X(x) dx \geq$
- $\geq \int_a^\infty x f_X(x) dx \geq \int_a^\infty a f_X(x) dx =$
- $a\, P[X \geq a]$
- Logo: $P[X \geq a] \leq \frac{E[X]}{a}$

Exemplo (continuação)

- A média da altura de uma população é 1,65 m.
- Qual o limite superior de probabilidade de um indivíduo ultrapassar os 2 metros ?

- $P(X \geq 2) \leq \frac{1{,}65}{2} = 0{,}825$

- Limite não muito útil ou significativo !

Exemplo 2

- A média das classificações numa turma é 15,2.
- Qual o limite superior de probabilidade de um dos alunos ter nota igual ou superior a 17 ?

- $P(X \geq 17) \leq \frac{15,2}{17} = 0.8941$
- E superior a 19?

- $P(X \geq 19) \leq \frac{15,2}{19} = 0.8$

Questão 2

- Quão provável é a **diferença entre a variável e o seu valor esperado** ser superior/inferior a um determinado valor ?

- Isto é $P(|X - E[X]| \geq a) = ?$

  Ou $P(|X - E[X]| < \ a) = ?$

- Exemplo: Probabilidade de os valores diferirem da média mais que 2 desvios padrão ?

Desigualdade de Chebyshev

- Pela Desigualdade de Chebyshev temos:
- $P(|X - E[X]| \geq a) \leq \frac{Var(X)}{a^2}$

- Ou, em alternativa:
- $P(|X - E[X]| < a) \geq 1 - \frac{Var(X)}{a^2}$

Desigualdade de Chebyshev

- Demonstração:

- Define-se $D^2 = (X - E[X])^2$
- É óbvio que $D^2 \geq 0$ e $D^2 \geq a^2 \Leftrightarrow \ |D| \geq a$
- Aplicando a Desigualdade de Markov
- $P(|D| \geq a) = P(D^2 \geq a^2) \leq \frac{E\left[(X-E[X])^2\right]}{a^2} = \frac{Var(X)}{a^2}$

- Assume-se E[X] e Var(X) finitos

Desigualdade de Chebyshev

- Se expressarmos a em função do desvio padrão, fazendo $a = h\sigma$, teremos:
- $P(|X - E[X]| \geq h\sigma) \leq \frac{\sigma^2}{(h\sigma)^2} = \frac{1}{h^2}$
- Ou seja: a probabilidade de obter um valor que dista da média de h desvios padrão ou mais é menor ou igual a $\frac{1}{h^2}$
  - Exemplos:
    - h=1 => P <= 1
    - **h=2 => P <= ¼**   Valores pouco precisos

Voltando à média de n variáveis aleatórias ...

- Como vimos, **a variância da média das estimativas** tende para 0 à medida que n aumenta
  - o que se pode interpretar como a probabilidade da média das amostras se aproximar do valor médio ser cada vez maior, aproximando-se de 1
- Recorrendo à Desigualdade de Chebyshev temos:
- $P(|M_n - E[M_n]| \geq \epsilon) \leq \frac{Var(M_n)}{\epsilon^2}$
- $P(|M_n - E[M_n]| \geq \epsilon) \leq \frac{\frac{\sigma^2}{n}}{\epsilon^2}$
- $P(|M_n - E[M_n]| < \epsilon) \geq 1 - \frac{\sigma^2}{n\,\epsilon^2}$

Lei fraca dos grandes números

- Passando ao limite teremos:
- $\lim_{n \to \infty} P(|M_n - E[M_n]| < \epsilon) = 1$

- Resultado que é conhecido por **Lei Fraca dos Grandes Números**

Leis dos grandes números

- Existe um segundo enunciado (fora dos objectivos de MPEI), a **lei forte dos grandes números**, que afirma:

$$P(\lim_{n\to\infty} M_n = \mu) = 1$$

- A **Lei Fraca dos Grandes Números** afirma que para um valor de n suficientemente elevado a média das amostras estará muito próxima do valor esperado
- Enquanto que a lei forte garante que é certo que o limite para que tende a média (das amostras) é o valor esperado

L. G. N. e definição frequencista

- Consideremos uma sequência de experiências aleatórias independentes e repetidas
- e seja I_j indicadora da ocorrência de A na experiência de ordem j

  [1 significa que A ocorreu]

- O número total de ocorrências de A nas n experiências será:

$$N_n = I_1 + I_2 + \cdots + I_n$$

…

- Como a frequência relativa de A é

$$f_A(n) = \frac{(I_1 + I_2 + \cdots + I_n)}{n}$$

- f_A é a média das amostras das variáveis aleatórias I_i

- Então (pelas duas leis dos grandes números):

$$\lim_{n \to \infty} P(|f_A(n) - p(A)| < \epsilon) = 1$$

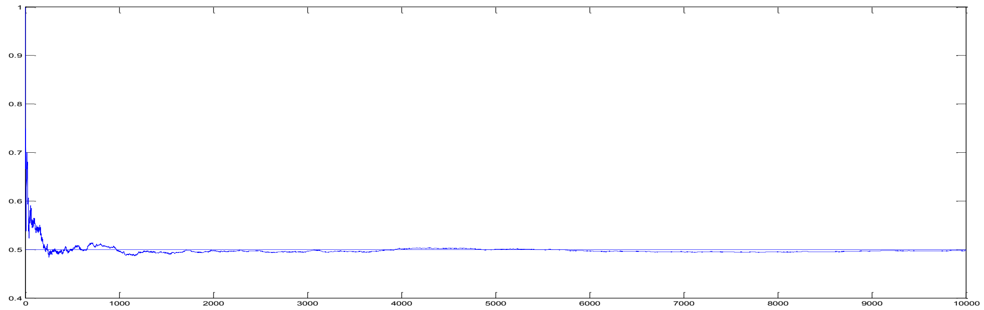
e

$$P[\lim_{n \to \infty} f_A(n) = p(A)] = 1$$

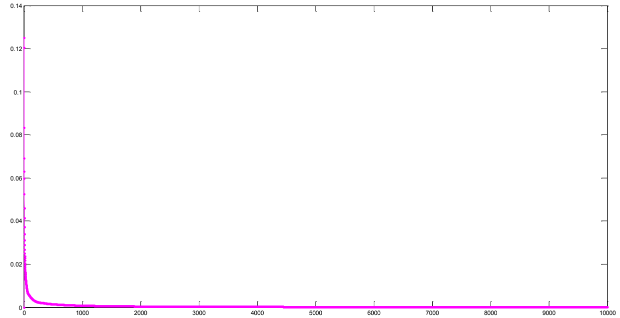
- Permitindo-nos dizer que a frequência relativa é uma boa estimativa da probabilidade

Ilustração

- Evolução da média amostral



- E da variância dessa média



Um pouco de História (para terminar esta parte)

- 1713: Lei fraca descrita por Jacob **Bernoulli**
- 1835: **Poisson** chama-lhe “La Loi des Grands Nombres”
  – Lei dos Grandes Números em Francês
- 1909: Émile Borel desenvolve a Lei forte para variáveis de Bernoulli
- 1928: Andrei Nikolaevich **Kolmogorov** prova a Lei forte no caso geral