UNIVERSIDADE DE AVEIRO
DEPARTAMENTO DE ELECTRÓNICA, TELECOMUNICAÇÕES E INFORMÁTICA

MPEI - Métodos Probabilísticos para Engenharia Informática (2018/2019) **PL 02**

---

**Palavras chave**: probabilidade, probabilidade condicional, independência

Responda às seguintes questões através de simulações em Matlab/Octave e sempre que seja pedido compare com os valores teóricos:

1. Consideremos famílias com filhos:

   (a) Obtenha uma estimativa da probabilidade do acontecimento "ter pelo menos um filho rapaz" em famílias com 2 filhos. Considere que a probabilidade de nascimento de rapazes é igual à de nascimento de raparigas.

   (b) Compare com o valor obtido aplicando a teoria clássica. Que aproximações teve de efetuar? Os valores são iguais?

   (c) Suponha agora que para uma família (com os referidos 2 filhos) escolhida ao acaso sabemos que um dos filhos é rapaz. Qual a probabilidade do outro filho ser também rapaz? Obtenha o valor aplicando primeiro a definição frequencista e de seguida a teoria clássica, se aplicável.

   (d) Sabendo que o primeiro filho é rapaz, qual a probabilidade do segundo ser rapaz? O que se pode concluir do resultado obtido?

   (e) Considere uma família com mais filhos, por exemplo 5, e que sabemos que pelo menos um dos filhos é rapaz. Obtenha usando simulação uma estimativa para a probabilidade de um dos outros (e apenas um) ser também rapaz.

   (f) Repita a questão anterior, mas considerando a probabilidade de pelo menos um dos outros ser também rapaz.

2. Considere uma sequência de 10 lançamentos de uma moeda honesta em que saíram 10 caras:

   (a) Obtenha por simulação uma estimativa para a probabilidade de sair essa sequência de 10 caras?

   (b) Obtenha por simulação uma estimativa da probabilidade de sair cara no próximo lançamento, o décimo-primeiro?

   (c) Compare os valores anteriores com os valores teóricos.

3. Consideremos o problema de deteção de cancro da mama. O mamograma (como muitas outras análises clínicas) não é infalível. Estudos prolongados revelaram que:

   P[ "mamograma positivo se cancro da mama" ] = 0,9

   P[ "mamograma positivo se NÃO cancro da mama" ] = 0,1

   Calcule a probabilidade de uma mulher escolhida ao acaso na população portuguesa ter cancro sabendo que o seu mamograma deu positivo. Considere que a frequência de ocorrência de cancro da mama na população portuguesa feminina é de 1 em 10000.

   Estava à espera deste resultado?

   Calcule nova estimativa, mas agora considerando as mulheres que procuram a consulta específica e que para este subgrupo a ocorrência de cancro atinge 1 em 1000.

   O que sugere para aumentar esta probabilidade? Se melhorar esse valor em 10% qual o ganho na probabilidade de ser mesmo cancro quando o exame dá positivo?

4. Considere o seguinte “jogo”: lançamento com os olhos vendados de $n = 20$ dardos, um de cada vez, para $m = 100$ alvos, garantindo-se que cada dardo atinge sempre um alvo (e apenas 1).

   a) Qual a probabilidade de nenhum alvo ter sido atingido mais do que uma vez?

   b) Qual a probabilidade de pelo menos 1 alvo ter sido atingido 2 (ou mais) vezes?

   c) Faça gráficos da variação em função de $n$ do valor da probabilidade da alínea b). Considere para $m = 1000, 10000, 100000, 1000000$ e para cada valor de $m$ valores de n entre e 1 e 100, com incremento de 10. Os 4 gráficos devem ser sub-gráficos de uma mesma figura.

   d) Faça um gráfico similar em função de m, fixando n em 100.

5. ** TPC ** Considere um array de tamanho 1000 que serve de base à implementação de uma memória associativa (por exemplo em Java) e que se pode assumir que a função de hash devolve um valor entre 0 e 999 com igual probabilidade.

   (a) Qual a probabilidade de haver colisões (pelo menos 2 *keys* mapeadas pela função de hash para a mesma posição do array) em situações em que temos apenas 10 *keys* ?

   (b) Faça um gráfico da probabilidade de colisões em função do número de keys.

   (c) Para um número fixo de keys, por exemplo 50, represente graficamente a variação da probabilidade de não haver qualquer colisão em função do tamanho do array.

6. Consideremos uma festa em que estão presentes um determinado número de pessoas, que designamos por $n$.

   (a) Qual deve ser o menor valor de $n$ para o qual a probabilidade de duas ou mais pessoas terem a mesma data de aniversário (mês e dia) é superior a 0,5?

   (b) Qual deve ser o valor de $n$ para que a probabilidade da alínea anterior passe a ser superior a 0,9?

7. You have a fair five-sided die. The sides of the die are numbered from 1 to 5. Each die roll is independent of all others, and all faces are equally likely to come out on top when the die is rolled. Suppose you roll the die twice.

   (a) Let event A to be “the total of two rolls is 10”, event B be “at least one roll resulted in 5”, and event C be “at least one roll resulted in 1”.

   i. Is event A independent of event B?

   ii. Is event A independent of event C?

   (b) Let event D be “the total of two rolls is 7”, event E be “the difference between the two roll outcomes is exactly 1”, and event F be “the second roll resulted in a higher number than the first roll”.

   i. Are events E and F independent?

8. Considere uma linguagem com apenas 3 palavras {“um”, “dois”, “três”} e que permite sequências de 2 palavras. Se todas as frases forem equiprováveis e as duas palavras na frase puderem ser iguais:

   (a) Qual a probabilidade da sequência “um dois”?

   (b) Qual a probabilidade de “um” aparecer pelo menos uma vez?

   (c) Qual a probabilidade de ocorrer “um” ou “dois”?

   (d) Qual o valor de P[“sequência incluir a palavra um” | “sequência inclui palavra dois”]?

   (e) Resolva a questão anterior para o caso de termos 10 palavras diferentes e sequências de 5 palavras com ajuda de um programa que calcule exaustivamente todas as possibilidades. Sugestão: use números de 1 a 10 para representar as palavras e use Matlab/Octave. Tente criar uma função com os parâmetros de entrada que considere adequados.

   (f) Repita a questão anterior para 11, 12, ... 20 palavras diferentes e represente a variação da probabilidade num gráfico. Nota: Devido à memória necessária não se pode ter valores muito elevados.

   (g) Adicione ao gráfico anterior, para comparação, a probabilidade se a linguagem não permitir repetições das palavras.

9. Considere que uma empresa tem 3 programadores (André, Bruno e Carlos) e que a probabilidade de um programa de cada um deles ter problemas ("bugs") e o número de programas desenvolvidos assumem os valores apresentados na tabela seguinte.

| Programador | Prob("erro num programa") | programas |
|---|---:|---:|
| André | 0.01 | 20 |
| Bruno | 0.05 | 30 |
| Carlos | 0.001 | 50 |

O Diretor da empresa seleciona de forma aleatória um dos 100 programas produzidos pelos seus 3 programadores e descobre que este contém um erro sério.

(a) Qual é a probabilidade de o programa ser do Carlos?

(b) De quem é mais provável ser o programa?

Responda à questão seguinte usando uma abordagem teórica:

10 Most mornings, Victor checks the weather report before deciding whether to carry an umbrella. If the forecast is "rain" the probability of actually having rain that day is 80%. On the other hand, if the forecast is "no rain", the probability of it actually raining is equal to 10%. During fall and winter the forecast is "rain" 70% of the time and during summer and spring it is 20%.

(a) One day, Victor missed the forecast and it rained. What is the probability that the forecast was "rain" if it was during the winter? What is the probability that the forecast was "rain" if it was during the summer?

(b) The probability of Victor missing the morning forecast is equal to 20 % on any day in the year. If he misses the forecast, Victor will flip a fair coin to decide whether to carry an umbrella. On any day of a given season he sees the forecast, if it says "rain" he will always carry an umbrella, and if it says "no rain", he will not carry an umbrella. Are the events "Victor is carrying an umbrella", and "The forecast is no rain" independent? Does your answer depend on the season?

(c) Victor is carrying an umbrella and it is not raining. What is the probability that he saw the forecast? Does it depend on the season?