

R 软件作业 1

2016 年 3 月 4 日

要求将编写的 R 程序以作者姓名命名，上机课时会让大家在电脑上演示。切记：学好 R 的唯一途径就是 action!

1. (1) 生成100个均值为2，标准差为3的正态分布的随机数。
(2) 计算上述随机数的均值、方差、标准差、最小值、最大值、95%的分位数、十分位数。
(2) 查看函数 hist 的帮助，熟悉参数 freq 的用法，并作出上述随机数的频数 (高为频数) 直方图和密度 (高为频率/组距，直方图下面积为1) 直方图。
2. 对 R 中的数据集 cars 将数据集作如下分析：
(1) 将 cars 中的 speed 作为自变量 X ，dist 作为因变量 Y ，作散点图，查看是否有线性趋势。
(2) 建立一元线性回归模型 $Y = \beta_0 + \beta_1 + \varepsilon$ ，用 R 函数 lm 对上述模型进行回归分析，并导出相应的最小二乘估计。(参考例 2.3)
(3) 根据教材第六章第一节的理论 (或上学期线性模型的理论)，自己编写程序，计算上述估计结果。
(4) 作回归系数的显著性检验，即检验 $H_0 : \beta_1 = 0$ ，自己计算检验统计量的观测值和 p 值，在显著性水平 0.05 下，判断是否显著，并与 lm 的输出结果作比较。
(5) 自己计算相应的残差和标准化残差，并以标准化残差为纵坐标，以拟合值为横坐标，作残差图，判断数据是否满足高斯马尔科夫假设。
3. (1) 将文本文件 exam0203.txt 读入 R 软件，命名为 student。
(2) 对 student 中的 Height 取均值，求和。
(3) 查看函数 tapply 的用法，对 student 中的 Weight 按照第二列的 Sex 取均值。
(4) 对 student 中的 Weight 按照第二列的 Sex 和第三列的 Age 取均值。
4. (1) 生成向量 x

-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10

- (2) 按如下分段函数生成向 y

$$y = \begin{cases} x^2 + 1, & x \geq 0; \\ \sin(x), & x < 0. \end{cases}$$

5. 生成矩阵

$$P = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.4 & 0.1 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.4 & 0.1 \end{pmatrix}$$

- (1) 使用函数 `apply` 验证 P 的各行和为1.
- (2) 计算 P^n , $n = 2, 3, 5, 10$ 。有什么规律吗?