

Đề kiểm tra thực hành Xác suất Thống kê

Ngày 25 tháng 12 năm 2023

Lưu ý:

- Thời gian làm bài: 60ph.
- Link nộp bài: <https://forms.gle/KDeFoaLNP9xhqFmV9>
- Sinh viên làm bài trên R.script, lưu lại với tên: "THXSTK_MSSV_HoTen_Test4.R"
- Bài làm cần trình bày như sau:

```
##
## Bai kiem tra thuc hanh XSTK
## Thu ... - Ca hoc ...
##
## Ho ten : ..... - MSSV : .....
##
## *****
```

Bài 01. (3đ) Cho cặp dữ liệu (x, y) trong bảng sau:

x	0.21	0.23	0.68	0.71	0.73	0.85	0.87
y	1.45	1.50	2.28	2.42	2.58	2.58	2.84

- (1.1) (0.5đ) Tạo vectơ x và vectơ y như trong bảng trên.
- (1.2) (0.5đ) Tạo ma trận M có kích thước 7×2 với cột đầu tiên là 7 số 1, cột thứ hai chứa vectơ x .
- (1.3) (1đ) Dùng các phép toán có sẵn trong R, tìm $\alpha, \beta \in \mathbb{R}$ sao cho

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (M^T M)^{-1} M^T y$$

- (1.4) (i) (0.5đ) Vẽ biểu đồ phân tán của bộ dữ liệu x, y
- (ii) (0.5đ) Trên cùng biểu đồ ở câu (i), vẽ đường thẳng $y = \alpha + \beta x$ với `col = 'red'` và `lwd = 2`. Các giá trị α, β là các giá trị tìm được ở câu (1.3)

Bài 02. (2đ) Chín người tình nguyện tham gia thí nghiệm xem việc tiếp xúc carbon monoxide (CO) có ảnh hưởng đến khả năng thở hay không. Dữ liệu được thu thập bởi các nhân viên ở khoa GDTC và Sức khỏe tại Học Viện Kỹ thuật Virginica và được phân tích ở Trung Tâm Tư Vấn Thống Kê tại Hokie Land.

Các đối tượng được cho tiếp xúc với các buồng thở, một trong số đó có nồng độ CO cao. Mỗi đối tượng ở mỗi buồng được đo tần số thở. Các đối tượng được cho tiếp xúc với các buồng thở theo một thứ tự ngẫu nhiên. Dữ liệu cho tần số thở, với đơn vị là số lần thở trong 1 phút, được cho ở bảng sau:

Người tình nguyện	Có CO	Không có CO
1	30	30
2	45	40
3	26	25
4	25	23
5	34	30
6	51	49
7	46	41
8	32	35
9	30	28

- (2.1) (0.5đ) Gọi μ_1, μ_2 lần lượt là tần số thở trung bình khi tiếp xúc với CO và khi không tiếp xúc với CO. Viết giả thuyết thống kê phù hợp cho bài toán.
- (2.2) (1.5đ) Dùng hàm có sẵn trong **R**, với mức ý nghĩa 1%, có thể nói tần số thở của của một người khi tiếp xúc CO sẽ *giảm* so với khi không tiếp xúc với CO hay không ?

Bài 03. (2đ) Quan sát tuổi thọ của một số bóng đèn của một xí nghiệp và ta được bảng sau:

Số giờ	1000	1100	1200	1300	1400	1500	1600	1700	1800
Số bóng đèn	10	14	16	17	18	16	16	12	9

- (3.1) (1đ) Tính trung bình mẫu và phương sai mẫu.
- (3.2) (1đ) Tính khoảng tin cậy 95% cho tuổi thọ trung bình của bóng đèn

Bài 04. (3đ) Phân phối Gumbel, kí hiệu $\text{Gumbel}(\mu, \beta)$, có hàm mật độ được cho bởi

$$f(x) = \frac{1}{\beta} \exp(-z - e^{-z}), \text{ với } z = \frac{x - \mu}{\beta} \forall x \in \mathbb{R}$$

Nếu $X \sim \text{Gumbel}(\mu, \beta)$ thì $\mathbb{E}(X) = \mu + \beta\gamma$ và $\text{Var}(X) = \frac{\pi^2}{6}\beta^2$.

Hệ số γ trong **R** được cho bởi lệnh `-digamma(1)`

- (4.1) (0.5đ) Viết hàm `f(x, mu, beta)` tính hàm f được cho như trên.
- (4.2) (0.5đ) Vẽ đồ thị hàm f trên $[-5, 5]$ với $\mu = 0, \beta = 1$
- (4.3) (1đ) Mô phỏng bằng thực nghiệm kết quả sau:

Định lí. Nếu $X \sim \text{Exp}(1)$ thì $-\ln X \sim \text{Gumbel}(0, 1)$

Thực hiện mô phỏng 5000 lần. Kết quả mô phỏng bao gồm biểu đồ tần suất thực nghiệm (`hist` với tham số `freq = FALSE`) và đường phân phối lý thuyết (`lwd = 2, col = 'red'`). Sinh viên chọn tham số `breaks` trong hàm `hist` hợp lý

- (4.4) (1đ) Sử dụng định lý giới hạn trung tâm, tính khoảng tin cậy 99% cho tham số μ , nếu biết $\beta = 1$, trung bình mẫu $\bar{x} \approx 0.565$, phương sai mẫu $s^2 \approx 1.642$ và cỡ mẫu $n = 200$

----- ■ -----