

# BÀI KIỂM TRA 03 - LÝ THUYẾT ƯỚC LƯỢNG

Ngày 21 tháng 5 năm 2024


Thời gian làm bài: 45 phút

- Làm bài trên [R script](#), lưu lại với tên có dạng: "LTTK\_MSSV\_HoTen\_Test3.R".
- Copy phần code bài làm sang một [file text .txt](#) để backup, lưu với tên có dạng "LTTK\_MSSV\_HoTen\_Test3.txt".
- Nộp bài cả file [R script](#) và [file text](#). Link nộp bài: [link google form](#).
- Tài liệu được phép sử dụng: giáo trình thực hành LTTK, file lý thuyết, các tài liệu tham khảo đã giới thiệu và ghi chú cá nhân của sinh viên.
- KHÔNG được dùng tài liệu là các phần code có sẵn.
- Bài làm cần trình bày như sau:

```
##
## Bai kiem tra 3 - Thuc hanh Ly thuyet Thong ke
## Lop: ST3 ca ...
##
## Ho ten: ..... - MSSV: .....
##
##*****
## Bai 1:
(phan bai lam cua Bai 1)
##-----
## Bai 2:
(phan bai lam cua Bai 2)
##-----
.....
##*****
## Ket thuc
```

**Chú ý:** Trong bài làm, sinh viên KHÔNG sử dụng các hàm `t.test`, `prob.test` và `binom.test` có sẵn trong .

## Bài 1 (4đ)

Tập dữ liệu `mtcars` (có sẵn trong ) chứa các thông tin của 32 mẫu xe ô tô.

- 1.1 (1đ)** Biến `qsec` trong dữ liệu là thời gian nhanh nhất (tính bằng giây) để một chiếc xe đi được 1/4 dặm từ trạng thái dừng. Hãy vẽ biểu đồ histogram cho biến này và nhận xét.

- 1.2** (1đ) Viết hàm `ci.mean(x, alpha)` để tính khoảng tin cậy cho kỳ vọng, với  $x$  là vector dữ liệu và  $(1-\alpha)$  là độ tin cậy.
- 1.3** (1đ) Áp dụng hàm `ci.mean(x, alpha)` để tìm khoảng tin cậy 95% và 99% cho biến `qsec`.
- 1.4** (1đ) Nếu muốn sai số ước lượng không vượt quá  $\varepsilon_0 = 0.4$  giây thì phải có thông tin của ít nhất bao nhiêu mẫu xe?

## Bài 2 (3đ)

Chiều cao (cm) của 100 cây giống được ghi lại trong bảng sau:

Chiều cao (cm)	90 – 95	95 – 100	100 – 105	105 – 110	110 – 115	115 – 120
Số cây	10	8	20	18	29	15

Một cây giống được coi là tốt nếu nó cao từ 110 cm trở lên.

- 2.1** (1đ) Chuyển bảng tần số dạng khoảng ở trên thành dữ liệu dạng vector cột và vẽ đồ thị histogram với tham số `breaks=seq(90,120,by=5)`.
- 2.2** (1đ) Viết hàm `ktc.tyle(data, data.p, alpha)` để tính khoảng tin cậy cho tỷ lệ, trong đó `data` là dữ liệu mẫu, `data.p` là phần dữ liệu thoả tính chất đang quan tâm và  $(1-\alpha)$  là độ tin cậy.
- 2.3** (1đ) Áp dụng hàm `ktc.tyle(data, data.p, alpha)` để tìm khoảng tin cậy 95% cho tỷ lệ cây giống tốt.

## Bài 3 (3đ)

Xét biến ngẫu nhiên  $X$  có hàm mật độ xác suất cho bởi


$$f(x | \theta) = \begin{cases} \theta x^{\theta-1} & \text{khi } 0 < x < 1, \\ 0 & \text{nơi khác,} \end{cases}$$


với  $\theta > 0$ . Giả sử có

$$\mathbf{x} = (0.0192, 0.4773, 0.3399, 0.4324, 0.8675, 0.2079, 0.3713, 0.2547, 0.3976, 0.2048)$$

là một mẫu ngẫu nhiên của  $X$ .

- 3.1** (1đ) Viết hàm `log_likelihoood(theta)` để tính log hàm hợp lý với dữ liệu trên.
- 3.2** (2đ) Tìm ước lượng hợp lý cực đại  $\hat{\theta}_{MLE}$  cho  $\theta$ , bằng hai cách:

**Cách 1:** giải tìm công thức cụ thể của  $\hat{\theta}_{MLE}$  rồi tính kết quả bằng .

**Cách 2:** sử dụng hàm `nlminb` có sẵn trong .