|  |  |  |
| --- | --- | --- |
| TRƯỜNG ĐẠI HỌC NHA TRANG  **KHOA CÔNG NGHỆ THÔNG TIN**  Người ra đề    Nguyễn Đức Thuần | ĐỀ THI KẾT THÚC MÔN HỌC  MÔN: **THỐNG KÊ MÁY TÍNH**  Thời gian: 75 phút  Lớp: 61 CNTT-1  *Không sử dụng tài liệu*  **ĐỀ I**  **Ngày thi: 07 /01/2022** | DUYỆT ĐỀ    Nguyễn Đức Thuần |

* **Đề này dành cho các sinh viên ký số cuối trong Mã số sinh viên là số lẻ.**

**Câu I**: (4.0đ) (**Mô tả dữ liệu**)

1. (1.0đ) Đọc File dữ liệu ex1.csv từ thư mục chứa đề bài hay trong <https://pluto.huji.ac.il/~msby/StatThink/index.html> vào dataframe có tên là ex1.
2. (1.0đ) Vẽ biểu đồ boxplot cho thuộc tính height thuộc file ex1.csv
3. (1.0đ) Phát hiện phần tử ngoại lệ của thuộc tính height bằng phương pháp số trung vị.
4. (1.0đ) Viết hàm hiển thị tổng 2 giá trị max, min của thuộc tính height.

**Câu II**: (3.0đ) (**Ước lượng**)

Để nghiên cứu nhu cầu của một loại hàng hóa ở một khu vực, người ta tiến hành khảo sát 800 gia đình chọn ngẫu nhiên trong khu vực. Kết quả thu được:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nhu cầu  (kg/tháng) | 35-40 | 40-45 | 45-50 | 50-55 | 55-60 | 60-65 |
| Số gia đình | 10 | 50 | 147 | 230 | 219 | 144 |

Gia đình có nhu cầu hơn 45kg/tháng là gia đình không nghèo.

1. Ước lượng tỷ lệ gia đình không nghèo trong khu vực với độ tin cậy 95%.
2. Ước lượng nhu cầu trung bình mỗi hộ gia đình trong khu vực với độ tin cậy 98%.

**Câu III**: (3.0đ) (**Kiểm định giả thuyết thống kê cho kỳ vọng không biết psai, mẫu lớn**)

Tại một nông trường, để điều tra khối lượng của một loại trái cây sau một đợt bón một loại phân mới. Người ta cân thử một số trái cây được chọn ngẫu nhiên và được kết quả sau:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Khối lượng (gam) | 45-50 | 50-55 | 55-60 | 60-65 | 65-70 | 70-75 | 75-80 | ≥80 |
| Số trái cây | 2 | 11 | 25 | 74 | 187 | 43 | 16 | 3 |

Trước kia chưa bón loại phân mới, khối lượng trung bình mỗi trái là 65gam. Hãy đánh giá xem việc bón phân có hiệu quả không với mức ý nghĩa α=0.01.

|  |  |  |
| --- | --- | --- |
| TRƯỜNG ĐẠI HỌC NHA TRANG  **KHOA CÔNG NGHỆ THÔNG TIN**  Người ra đề    Nguyễn Đức Thuần | ĐỀ THI KẾT THÚC MÔN HỌC  MÔN: **THỐNG KÊ MÁY TÍNH**  Thời gian: 75 phút  Lớp: 61 CNTT-1  *Không sử dụng tài liệu*  **ĐỀ II**  **Ngày thi: 07 /01/2022** | DUYỆT ĐỀ    Nguyễn Đức Thuần |

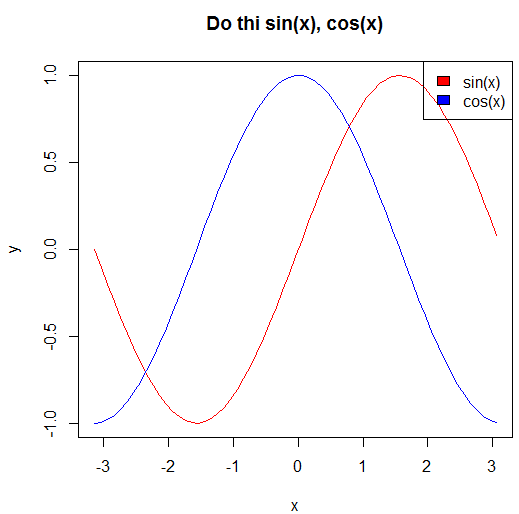
* **Đề này dành cho các sinh viên ký số cuối trong Mã số sinh viên là số chẵn.**

**Câu I**: (4.0đ) (**Mô tả dữ liệu**)

1. (1.0đ) Tạo một data frame **df** như sau:

df <- data.frame(col1 = c(1:3,99), col2 = c(2.5,4.2,99,3.2))

1. (1.0đ) Viết câu lệnh thay đổi tất cả giá trị 99 trong data frame df thành NA (giá trị thiếu).
2. (1.0đ) Viết các câu lệnh thay tất cả giá trị thiếu trong mỗi cột bằng giá trị trung bình của các giá trị đầy đủ trong cột đó.
3. (1.0đ) Vẽ hai đồ thị của 2 hàm sin(x), cos(x) trong cùng một cửa sổ màn hình có ghi chú như hình sau:



*Hướng dẫn*: Trong ngôn ngữ R có hằng **pi** đểtính độ đo các góc theo radian.

*Đồ thị trên là vẽ x∈(-pi,pi)*

**Câu II**: (3.0đ) (**Ước lượng**)

Kiểm tra năng suất của 1 số hecta lúa chọn ngẫu nhiên ở một vùng, người ta thu được kết quả:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Năng suất  (ta/ha) | 30-35 | 35-40 | 40-45 | 45-50 | 50-55 | 55-60 | 60-65 | 65-70 |
| Diện tích (ha) | 2 | 5 | 8 | 10 | 16 | 25 | 17 | 13 |

Những thửa ruộng có năng suất hơn 45 (tạ/ha) gọi là ruộng tốt.

1. Ước lượng tỷ lệ các ruộng tốt trong vùng với độ tin cậy 95%.
2. Ước lượng năng suất trung bình của các hecta ruộng với độ tin cậy 98%

**Câu III**: (3.0đ) (**Kiểm định giả thuyết thống kê cho kỳ vọng không biết psai, mẫu lớn**)

Với ý muốn làm tăng chỉ số mỡ sữa của loại giống bò A, một trang trại chăn nuôi cho lai bò giống A với một loại bò giống B. Đo chỉ số mỡ sữa của 130 con bò lai giống được chọn ngẫu nhiên trong đàn bò lai của trại, người ta có kết quả sau:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chỉ số mỡ sữa | 3.0 – 3.6 | 3.6-4.2 | 4.2-4.8 | 4.8-5.4 | 5.4-6.0 | 6.0-6.6 | 6.6-7.2 |
| Số bò lai | 2 | 8 | 35 | 43 | 22 | 15 | 5 |

Biết rằng chỉ số mỡ sữa trung bình của giống bò A thuần chủng là 4.95. Hãy cho kết luận về hiệu quả của việc lai giống với mức ý nghĩa α=0.01.

* Nộp bài:
* Chụp ảnh các câu trang bài làm, lưu các ảnh trang bài làm vào 1 file word có tên: TênSV\_Mã số SV, ví dụ: Tuan\_62123456 nộp lên trang Elearning.