



TRƯỜNG ĐẠI HỌC NHA TRANG KHOA CÔNG NGHỆ THÔNG TIN Người ra đề  Nguyễn Đức Thuận	ĐỀ THI KẾT THÚC MÔN HỌC MÔN: THỐNG KÊ MÁY TÍNH Thời gian: 75 phút Lớp: 61 CNTT-2 Không sử dụng tài liệu ĐỀ I Ngày thi: 05 /01/2022	DUYỆT ĐỀ  Nguyễn Đức Thuận
---	---	---

- **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)

Số liệu về điểm số các bài tập của một sinh viên A làm bài tập Thống kê Máy tính như sau (thang điểm 100) lưu trong vector **dl**:

33; 42; 49; 49; 53; 55; 55; 61; 63; 67; 68; 68; 69; 69; 72; 73; 74; 78; 80; 83; 88; 88; 88; 90; 92; 94; 94; 94; 94; 96; 100

Hãy thực hiện các yêu cầu thống kê liên quan điểm số của sinh viên A:

- (0.5đ) Xây dựng vector **dlu** gồm các giá trị điểm phân biệt của A (*dùng hàm unique*).
- (1.5đ) Viết hàm tính tần suất (số lần xuất hiện) giá trị D trong vector dl, ký hiệu **ts(D,dl)**.
- (1.0đ) Xây dựng vector **H** mà $H[i] = t(dlu[i],dl)$ # Vector H có cùng kích thước vector dlu.
- (1.0đ) Nếu làm tiếp, bài tập sinh viên A hy vọng điểm số là bao nhiêu thì hợp lý.

Câu II: (3.0đ) (Ước lượng)

Dữ liệu sau là kết quả của một cuộc khảo sát ngẫu nhiên 39 quốc kỳ từ các quốc gia khác nhau. Chúng ta quan tâm đến việc ước lượng khoảng tin cậy cho số màu trung bình trên lá cờ các quốc gia. Gọi X = số màu trên lá cờ quốc gia.

X	Số lần xuất hiện
1	1
2	7
3	18
4	7
5	6



- (1.5đ) Tính các đại lượng sau: Kích thước mẫu, Giá trị trung bình, Độ lệch chuẩn.
- (1.5đ) Với độ tin cậy 95%, hãy ước lượng khoảng tin cậy cho giá trị trung bình số màu trên mỗi lá cờ.

Câu III: (3.0đ) (Kiểm định giả thuyết thống kê)

Một tiệm bánh mì cung cấp những ổ bánh cho các siêu thị. Tiệm Bánh mì cam kết trọng lượng tối thiểu là $\mu = 2$ kg. Tuy nhiên, không phải gói hàng (ổ bánh) nào cũng nặng đúng 2 kg vì có sự thay đổi trong trọng lượng. Do đó, các siêu thị tìm hiểu xem trọng lượng trung bình của ổ bánh có nhỏ hơn đáng kể 2 kg hay không. Giả định, khối lượng X (tính bằng kg) của ổ bánh là giả định là phân phối chuẩn. Bằng kinh nghiệm, người ta xác định được phương sai $\sigma^2 = 0.12$. Một siêu thị lấy một mẫu gồm $n = 20$ ổ bánh mì và cân chúng, trọng lượng trung bình được tính là $\bar{x} = 1.97$ kg.

- (1.0đ) Phát biểu giả thuyết thống kê.
- (1.5đ) Kiểm định giả thuyết thống kê với mức ý nghĩa $\alpha = 0.05$.
- (0.5đ) Sử dụng P-value để kiểm định với mức ý nghĩa $\alpha = 0.05$.

- Nộp bài:
- Chụp ảnh các trang bài làm, tên các ảnh tương thích với thứ tự các trang. Ví dụ: Page1, Page2,...
- Lưu các ảnh trang bài làm vào 1 thư mục có tên: TênSV_Mã số SV, ví dụ: Tuan_62123456, nén lại và nộp lên trang Elearning.

TRƯỜNG ĐẠI HỌC NHA TRANG KHOA CÔNG NGHỆ THÔNG TIN Người ra đề  Nguyễn Đức Thuận	ĐỀ THI KẾT THÚC MÔN HỌC MÔN: THỐNG KÊ MÁY TÍNH Thời gian: 75 phút Lớp: 61 CNTT-2 <i>Không sử dụng tài liệu</i> ĐỀ II Ngày thi: 05 /01/2022	DUYỆT ĐỀ  Nguyễn Đức Thuận
---	---	---

- 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)

Một đề tài A có 32 người tham gia, số ngày công tham gia của mỗi thành viên được ghi nhận như sau:

32, 35, 45, 45, 45, 48, 50, 50, 50, 54, 55, 57, 57, 60, 60, 60,
60, 62, 62, 65, 68, 68, 68, 70, 72, 72, 72, 75, 75, 75, 75, 75

Hãy thực hiện các yêu cầu thống kê liên quan số ngày công của các thành viên đề tài A:

- (0.5đ) Xây dựng vector dlu gồm các giá trị điểm phân biệt của A (*dùng hàm unique*).
- (1.5đ) Viết hàm tính tần suất (số lần xuất hiện) giá trị D trong vector dl , ký hiệu $ts(D, dl)$.
- (1.0đ) Xây dựng vector H mà $H[i] = t(dlu[i], dl)$
- (1.0đ) Cho biết số ngày công hy vọng mỗi thành viên có thể đạt được.

Câu II: (3.0đ) (Ước lượng)

Tại trường phổ thông trung học X. Trong một lớp học đàn dành cho Khối 10, người ta nhận thấy có 64 em nữ và 16 em nam. Ban giám hiệu quan tâm đến tỷ lệ các em nữ tham gia học đàn. Giả sử các em Khối 10 đăng ký ngẫu nhiên vào lớp đàn.

- (1.5đ) Với độ tin cậy 95%, hãy ước lượng khoảng tin cậy cho tỷ lệ các em nữ Khối 10 tham gia học đàn.
- (1.5đ) Với số liệu trên, nếu muốn sai số ước lượng tỉ lệ các em nữ Khối 10 tham gia học đàn là $\varepsilon = 0.09$, thì độ tin cậy là bao nhiêu?

Câu III: (3.0đ) (Kiểm định giả thuyết thống kê, 2 mẫu không độc lập)

Một tiệm bánh nhỏ bán bánh quy trong các gói 500 g. Các chiếc bánh là được làm thủ công và việc đóng gói do chính người thợ làm bánh hoặc vợ của anh ta làm. Một số khách hàng phỏng đoán rằng người vợ hào phóng hơn người thợ làm bánh. Một khách hàng thực hiện một thử nghiệm: anh ta mua những gói bánh quy do người thợ làm bánh và vợ anh ta đóng gói vào 8 ngày khác nhau và cân các gói. Anh ta nhận được hai mẫu sau (một cho người làm bánh, một cho vợ).

Bánh (Người vợ làm)	512	530	498	540	521	528	505	523
Bánh (Ng Làm bánh)	499	500	510	495	515	503	490	511

Chúng ta muốn kiểm tra xem phỏng đoán của khách hàng có chính đáng hay không.

- (1.0đ) Phát biểu giả thuyết thống kê.
- (1.5đ) Kiểm định giả thuyết thống kê với mức ý nghĩa $\alpha = 0.05$.
- (0.5đ) Sử dụng P-value để kiểm định với mức ý nghĩa $\alpha = 0.05$.

Gợi ý: Kiểm định trường hợp 2 mẫu không độc lập.

- Nộp bài:
- Chụp ảnh các trang bài làm, tên các ảnh tương thích với thứ tự các trang. Ví dụ: Page1, Page2,...
- Lưu các ảnh trang bài làm vào 1 thư mục có tên: TênSV_Mã số SV, ví dụ: Tuan_62123456, nén lại và nộp lên trang Elearning.

ĐÁP ÁN (61 CNTT 2)

ĐỀ I

Câu I:

```
> dl
[1] 33 42 49 49 53 55 55 61 63 67 68 68 69 69 72 73 74 78 80
[20] 83 88 88 88 90 92 94 94 94 94 96 100
> # Cau I.a
> dlu <- unique(dl)
> # Cau I.b
> ts<-function(D,dl){
+ dem<-0
+ n<-length(dl)
+ for(i in 1:n) if (D==dl[i]) dem<-dem+1
+ return (dem)}
> # Cau I.c
> H<-rep(1,length(dlu))
> n<-length(dlu)
> for(i in 1:n) H[i]<-ts(dl[i],dl)
> # Cau I.d
> Xbar<-mean(X)
> XS<-H/length(dl)
> E<- dlu*XS
> EX<- sum(E)
> EX
[1] 80.48387
```

Câu II:

```
> # Cau II.a
> X<-1:5
> ts<-c(1,7,18,7,6)
> n<-sum(ts)
> Xbar<- sum(X*ts)/n
> xm2<- sum((X-Xbar)^2*ts)/(n-1)
> SD<-sqrt(xm2)
> n; Xbar; SD
[1] 39
[1] 3.25641
[1] 1.018718
> # Cau II.b
> alpha<-0.05
> zstar<-qnorm(1-alpha/2)
> er<-zstar*SD/sqrt(n)
> round(Xbar-er);round(Xbar+er)
[1] 3
[1] 4
```

Câu III:

```
> #Cau II.a
> # H0: mu >= 2; H1: mu <2
> #Cau III.b
> n<-20 # Mau nho
> # Biet phuong sai
> xma2<- 0.12
> Xbar<- 1.97
> alpha<-0.05
> xma<-sqrt(xma2)
> mu0<-2
> Z0<- (Xbar-mu0)/(S/sqrt(n))
Error: object 'S' not found
> Z0<- (Xbar-mu0)/(xma/sqrt(n))
> Zstar<-qnorm(1-alpha)
> Z0 < -Zstar
[1] FALSE
> # Chap nhan H0: Banh mi khong nho hon dang ke
> #Cau III.c
> P_value<-pnorm(Z0)
> P_value < alpha
[1] FALSE
> # Chap nhan H0.
```

ĐỀ II

Câu I:

```
[20] 12 12 13 13 13 13 13
> #Cau I.a
> dlu<-unique(dl)
> #Cau I.b
> ts<-function(D,dl){
+ n<- length(dl)
+ dem<-0
+ for(i in 1:n) if (D==dl[i]) dem<-dem+1
+ return (dem)}
> #Cau I.c
> H<-rep(1,length(dlu))
> n<-length(dlu)
> for(i in 1:n) H[i]<-ts(dlu[i],dl)
> H
[1] 1 1 3 1 3 1 1 2 4 2 1 3 1 3 5
> XS<-H/length(dl)
> #Cau I.c
> XS
[1] 0.03125 0.03125 0.09375 0.03125 0.09375 0.03125 0.03125 0.06250 0.12500
[10] 0.06250 0.03125 0.09375 0.03125 0.09375 0.15625
> E<-sum(dlu*XS)
> E
[1] 59.90625
```

Câu II

Câu III

```
> # Cau II.a
> nu<-64; nam<-16
> n<- nu+nam
> pmu<-nu/n
> alpha<-0.05
> Zstar<-qnorm(1-alpha/2)
> e1<-sqrt((pmu*(1-pmu)/n))
> er<-Zstar*e1
> pmu-er;pmu+er
[1] 0.7123477
[1] 0.8876523
> #Cau II.b
> # Neu er=0.1
> er1<-0.1
> Zstar<-er1/e1
> alpha1<-2*(1-pnorm(Zstar))
> alpha1
[1] 0.02534732
> # Do tin cay
> (1-alpha1)*100
[1] 97.46527
```

```
> # Cau III
> X<-c(512,530,498,540,521,528,505,523)
> Y<-c(499,500,510,495,515,503,490,511)
> # Cau III.a
> # a.H0 :  $\mu_X \leq \mu_Y$ ; H1 :  $\mu_X > \mu_Y$ 
> # H0:  $\mu_D > 0$ ; H1:  $\mu_D > 0$  ;  $\mu_D = \mu_X - \mu_Y$ 
> alpha<-0.05
> D<-X-Y
> Dbar<-mean(D)
> S2<-var(D)
> S<-sqrt(S2)
> n<-length(D)
> T0<-(Dbar-0)/(S/sqrt(n))
> ta<-qt(1-alpha,n-1)
> T0>ta
[1] TRUE
> # Bac bo H0: Nguoi vo hao phong hon
> # Cau III.b
> p<-1-pt(T0,n-1)
> p
[1] 0.01352229
> p < alpha
[1] TRUE
> # Bac bo H0
```