Ước Lượng khoảng cho kì vọng μ: $(\bar{x} - \varepsilon;\ \bar{x} + \varepsilon)$

ϭ quần thể đã biết:	$\varepsilon = z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}$	ϭ quần thể chưa biết:	$\varepsilon = t_{\alpha/2}^{n-1} * \frac{s}{\sqrt{n}}$

Do covid, khảo sát 25 người già, thọ TB 70 tuổi. Với độ tin cậy 90%, UL tuổi biết tuổi thọ có PPC, độ lệch chuẩn 8.
Ϭ=8; n=25; α=10%=0.1; α/2=0.05; $z_{\alpha/2}$=Φ(1-α/2)=Φ(0.95)=1.645
Từ đó tính ε=1.645*8/sqrt(25)=2,632. KUL (70-2.632;70+2.632)

Do covid, khảo sát 25 người già, thọ TB 70 tuổi, độ lệch chuẩn hiệu chỉnh 4. Với độ tin cậy 90%, UL tuổi biết tuổi thọ có PPC.
n=25; α=10%=0.1; α/2=0.05; $t_{(0.05;24)}$=1.711
ε = 1.711*4/sqrt(25)=k; KUL (70-k 70+k)

Nếu đề ko cho Xtb và ko cho độ lệch chuẩn mẫu S mà cho bảng số liệu tuổi của 25 người thì ta đi tính Xtb, S. 2 loại bảng: bảng chi tiết và bảng tần số (lớp).
Với bảng chi tiết: Xtb=$\sum X_i$/n=(70+71+..+90)/25

Bảng chi tiết: 70,71,80,74,95,60,55,…90 (25 số nhé)
Bảng tần số:

Tuổi	60-62	63-65	67-70	71-77	80-90
Số lượng người	5	7	10	2	1

Với bảng tần số: Xtb=$\sum n_i*X_i$/n=(5*61+64*7+68.5*10+74*2+85*1)/25
Với bảng chi tiết: tìm S qua $S^2=\sum(X_i - Xtb)^2/(n-1)$ = [(70-Xtb)^2 + (71-Xtb)^2 +.. +(90-Xtb)^2]/24. Rồi suy ra S
Với bảng tần số: tìm S qua $S^2=\sum n_i*(X_i - Xtb)^2/(n-1)$ =[5*(61-Xtb)^2 + 7*(64-Xtb)^2 +.. 1*(85-Xtb)^2]/24. Rồi suy ra S

Ước Lượng khoảng cho tỉ lệ p: $(f - \varepsilon;\ f + \varepsilon)$

$$\varepsilon = z_{\alpha/2} * \sqrt{\frac{f(1-f)}{n}}$$

Do covid, khảo sát 480sv, có 12 thất nghiệp. UL với độ tin cậy 95%.
f=12/480=2.5%=0.025; n=480; α=5%=0.05; α/2=0.025; $z_{\alpha/2}$= Φ(1- α/2) =Φ(0.975) = 1.96
ε =1.96*sqrt(0.25*0.975/480)=1.396%. Vậy khoảng UL (1.1%; 3.9%)

Kiểm định giá trị trung bình μ_0

Chú ý: μ_0 là giả thuyết trung bình tổng thể, Xtb là tb của mẫu.

PHƯƠNG SAI CỦA TỔNG THỂ CHUNG σ^2 ĐÃ BIẾT

Tiêu chuẩn kiểm định là thống kê z: $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

H_0 đúng, z phân phối theo quy luật chuẩn hóa N(0,1).

Kiểm định phía phải: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$ Nếu z > Z_α, bác bỏ giả thuyết H_0.

Kiểm định phía trái: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$ Nếu z < -Z_α, bác bỏ giả thuyết H_0.

Kiểm định hai phía: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$ Nếu |z| > $Z_{\alpha/2}$, bác bỏ giả thuyết H_0.

PHƯƠNG SAI CỦA TỔNG THỂ CHUNG σ^2 CHƯA BIẾT

Tiêu chuẩn kiểm định là thống kê t: $t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$

H_0 đúng, t sẽ phân phối theo quy luật Student với (n - 1) bậc tự do.

Kiểm định phía phải: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases}$ Nếu t > $t_{\alpha,(n-1)}$, bác bỏ giả thuyết H_0.

Kiểm định phía trái: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu < \mu_0 \end{cases}$ Nếu t < -$t_{\alpha,(n-1)}$, bác bỏ giả thuyết H_0.

Kiểm định hai phía: $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$ Nếu |t| > $t_{\alpha/2,(n-1)}$, bác bỏ giả thuyết H_0.

Dạng: Kiểm định 2 phía, chưa biết phương sai tổng thể
Giá tăng đất dự đoán 580k; Kiểm tra 50 lô: mức tăng 599.5k với std là 84.84k; Với α =2%=0.02 (độ tin cậy 98%), kiểm dự đoán?

Giải:

Gọi μ là giá đất dự đoán tăng
Ta cần kiểm định giả thuyết: H_0: μ = μ_0 = 580k; H_1: μ ≠ 580k
Tiêu chuẩn kiểm định t=(Xtb-μ_0)*sqrt(n)/S=
=(599.5-587)*sqrt(50)/84.84=1.042
Tra bảng tìm giá trị tới hạn $t_{\alpha/2,(n-1)}$ =$t_{0.01;\,49}$= 2.403
|t|=1.042< $t_{0.01;\,49}$ nên ko thể bác bỏ H_0. Chấp nhận dự đoán.

Cũng dạng, nhưng không cho 599.5 và 84.48 thì ta phải đi tìm thông qua mẫu. Vi dụ:
Mẫu 1 liệt kê: 5, 7, 9, 4, 6, 7, 3, 9.
Suy ra Xtb=(5+..+9)/8=..
$S^2=\sum(X_i\text{-}Xtb)^2/(n-1)$ rồi suy ra S
Mẫu 2: thống kê

X	5	6	8	9	10
n	3	4	6	6	5

n=3+4+..+5=..; Xtb=(3*5+6*4+..+10*5)/n=..
S^2= ∑ n_i*(X_i-Xtb)/(n-1) rồi suy ra S

KIỂM ĐỊNH GIẢ THIẾT VỀ TỶ LỆ CỦA MỘT TỔNG THỂ CHUNG

Với n đủ lớn ($n.p_0 \geq 5$ và $n(1 - p_0) \geq 5$), tiêu chuẩn kiểm định là thống kê Z:

$$z = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Chú ý: p_0 là giả thuyết tỉ lệ trung bình tổng thể, f=n/N là tỉ lệ tb của mẫu đang nghiên cứu

H_0 đúng, thì z phân phối theo quy luật chuẩn hóa N(0,1).

Kiểm định phía phải: $\begin{cases} H_0: p = p_0 \\ H_1: p > p_0 \end{cases}$ Nếu z > Z_α, bác bỏ giả thuyết H_0.

Kiểm định phía trái: $\begin{cases} H_0: p = p_0 \\ H_1: p < p_0 \end{cases}$ Nếu z < -Z_α, bác bỏ giả thuyết H_0.

Kiểm định hai phía: $\begin{cases} H_0: p = p_0 \\ H_1: p \neq p_0 \end{cases}$ Nếu |z| > $Z_{\alpha/2}$, bác bỏ giả thuyết H_0.

Dạng: Kiểm định (tỉ lệ) phía phải
Tỉ lệ thất nghiệp công bố: 2%=0.02; Khảo sát 480sv, có 12 thất nghiệp. Đoàn k/s nghi ngờ số liệu bị làm đẹp. Vậy 2% đúng ko với α=5%=0.05; **Giải**
Ta có N=480; n=12; Vậy f=n/N=0.025>0.2
Ta cần kiểm định giả thuyết: H_0: p = p_0 = 0.02; H_1: p > 0.02
Tiêu chuẩn kiểm định z=(f-p_0)*sqrt(n)/sqrt(p_0*(1-p_0))
=(0.025-0.02)*sqrt(480)/sqrt(0.02*0.98)=0.782461
Tra bảng tìm giá trị tới hạn Φ(z_α)=1-α = 0.95; z_α=1.65
z=0.78< z_α=1.65 nên ko thể bác bỏ H_0. Chấp nhận con số 2%.

Dạng 2: Kiểm định phía trái
Tỉ lệ tốt công bố 90%. Kiểm tra: 261/300 tốt. α =5%, 90%???

Giải

Ta có N=300; n=261; Vậy f=n/N=0.87
Ta cần kiểm định giả thuyết: H_0: p = p_0 = 0.9; H_1: p < 0.9
Tiêu chuẩn kiểm định z=(f-p_0)*sqrt(n)/sqrt(p_0*(1-p_0))
=(0.87-0.9)*sqrt(300)/sqrt(0.9*0.1)=-1.73
Tra bảng tìm giá trị tới hạn Φ(z_α)=1-α = 0.95; z_α=1.65
z=-1.73< -z_α=-1.65 nên bác bỏ H_0. Không chấp nhận con số quảng cáo 90%.

Đề đã từng thi

Chú ý: ở đây Th viết ko theo kí hiệu toán học, còn vô thi các bạn trình bày theo kiểu toán học (như phân số, ∑ phải ghi trên dưới đầy đủ...) và trình bày rõ chứ ko ghi kết quả quá nhanhsẽ mất điểm.

Câu 1: *(2.0 điểm)*

Tìm hàm hồi quy tuyến tính của Y theo_ X và tính hệ số tương quan đối với mẫu như sau:

X	2	3	7	6	5
Y	6	6	15	13	10

Đ/s

a=1,918605, b= 1,174419

Kết luận phương trình:

Y=1,9186X+1,1744

r = 0,9794

Tính hệ số a, b dựa vào X, Y. Có a, b ta viết được phương trình.

a=(n*∑(XiYi) – ∑Xi*∑Yi)/(n*∑ $(Xi)^2$ – $(∑Xi)^2$); b=TrungBinhY -a*TrungBinhX

n=5; ∑(XiYi) = 2*6 + .. +5*10=263; ∑Xi = 2+..+5=23; ∑Yi = 6 + ..+10=50;

$∑(Xi)^2 = 2^2 +.. +5^2=123$; $(∑Xi)^2 =23^2=529$; TrungBinhX =(2+..+5)/5 =4,6; TrungBinhY=(6+..+10)/5=10;

Tính hệ số tương quan r:

r = [n*∑(XiYi) – ∑Xi∑Yi]/sqrt[(n*$∑(Xi)^2$ – $(∑Xi)^2$)(n*$∑(Yi)^2$ – $(∑Yi)^2$)], $∑(Yi)^2 = 6^2 + … + 10^2 = 566$; $(∑Yi)^2 = 2500$

Câu 2: *(4.0 điểm)*

Trong một công ty sản xuất bóng đèn, người quản lý lấy mẫu ngẫu nhiên 13 bóng đèn trong kho và bật bóng đèn cho đến khi nó hỏng để đo tuổi thọ mỗi bóng (đơn vị: giờ). Sau khi thực hiện, người quản lý thu được bảng số liệu dưới đây:

STT	1	2	3	4	5	6	7	8	9	10	11	12	13
Tuổi thọ (giờ)	342	426	317	545	264	451	631	512	266	492	562	298	X

Giả sử rằng tuổi thọ bóng đèn trong công ty là biến ngẫu nhiên có phân phối chuẩn với độ lệch chuẩn là 120 giờ. Bạn hãy:

a. Ước lượng tuổi thọ bóng đèn của công ty với độ tin cậy 90%.
b. Giải thích khái niệm độ tin cậy 90% theo cách hiểu của bạn.
c. Khoảng ước lượng sẽ thay đổi như thế nào khi độ tin cậy thay đổi. Theo bạn, người quản lý có nên ước lượng với độ tin cậy là 100% không?
d. Với độ tin cậy 90%, nếu người quản lý muốn tăng kích thước mẫu lên gấp 4 lần, thì độ rộng của khoảng tin cậy thay đổi như thế nào?

(Với X: là 3 chữ số áp cuối của mã sinh viên)

Thầy giả sử X là 300 nhé

Đ/s:

a/ 361,0972 < a < 470,595

b/ Khoảng ước lượng đúng

c/ Độ tin cậy càng cao thì KUL càng rộng (xem hình vẽ mấy đường nét đứng xanh/ hồng trong LAB 4, muốn tang số lượng đường màu xanh, tức là tang độ tin cậy thì phải cho các đường nét đứng này dài ra, tức KUL rộng ra)

Không nên ULK với độ tin cậy 100%, vì nếu ĐTC 100% thì phải mở rộng KUL tối đa, khi đó KUL chẳng còn ý nghĩa.

d/ Các em dựa vào công thức KUL tại B2 câu a. Ta có μ, xα, Ϭ không đổi, n đổi (tăng lên 4 lần), tức sqrt(n) tăng giảm 2 lần; Vậy độ rộng KUL giảm 1 nữa.

a/Đây là bài ULK Kỳ Vọng μ, ko phải ULK tỉ lệ. Với Ϭ quần thể đã biết =120 nên ta làm 2 bước:

B1: với ĐTC 90% =0,9, α =1-0,9 = 0,1; Tính $Φ(z_{α/2})$ = 1 -α/2 =1 - 0,1/2 = 1 – 0,05 = 0,95. $z_{α/2}$ = 1,645

B2: KUL: Xtb - $z_{α/2}$*Ϭ/sqrt(n)<μ< Xtb + $z_{α/2}$*Ϭ/sqrt(n).

n=13; Xtb = (342+ …+ 300)/13 = 5406/13=415,8461; Vậy 415,8461 – 1,645*120/sqrt(13) <μ< 415,8461 – 1,645*120

b/ Khoảng ước lượng đúng

c/ Độ tin cậy càng cao thì KUL càng rộng (xem hình vẽ mấy đường nét đứng xanh g/ hồng p trong LAB 5, muốn tăng số lượng đường màu xanh g, tức là tăng độ tin cậy thì phải cho các đường nét đứng này dài ra, tức KUL rộng ra)

Không nên ULK với độ tin cậy 100%, vì nếu ĐTC 100% thì phải mở rộng KUL tối đa, khi đó KUL chẳng còn ý nghĩa.

d/ Theo KUL tại B2 câu a. Có Xtb, $z_{α/2}$, Ϭ không đổi, n đổi (tăng 4 lần), tức sqrt(n) tăng 2 lần; Vậy độ rộng KUL giảm 1 nữa.

Câu 3: *(4.0 điểm)*

Quan sát số hoa hồng bán ra trong ngày của một cửa hàng bán hoa sau một thời gian, người ta thu được bảng số liệu như sau:

Số hoa hồng	12	13	15	16	17	18	19
Số ngày	Y	2	7	7	3	2	1

a. Do tình hình giãn cách xã hội, số lượng hoa bán bị sụt giảm Chủ cửa hàng hoa nói rằng nếu một ngày không bán được ít nhất 15 bông hoa thì chẳng thà đóng cửa còn hơn. Bạn hãy đặt giả thuyết và sử dụng bảng số liệu thu nhập được để đưa ra quyết định giúp chủ cửa hàng ở mức ý nghĩa 5%
b. Giả sử những ngày bán được 13 – 17 bông là những ngày "bình thường". Với mức ý nghĩa 10%, bạn hãy cho biết về ý kiến: "Tỷ lệ những ngày bình thường của cửa hàng là 50%"

Th giả sử Y=3

a/ H0: μ≥ $μ_0$=15; H1: μ <15; (nhớ nhé: kiểm định kỳ vọng 1 phía, α chứ ko phải α/2)

Chú ý nhé: KĐ kì vọng; 1 phía; Ϭ chưa biết; dùng phân phối t, n= Y+2+..+1=25

B1:Tìm $t_α$; B2:Tính t. B3: Kết luận t theo $t_α$

B1: $t_α = t_{(α;\ n-1)} = t_{(0.05;24)} = 1.711$

B2: $S^2=∑(Xi-Xtb)^2/(n-1)$= 44

S=sqrt(44)=6,63. t = (Xtb-μ)/(S*sqrt(n))

t=(15,4-15)/(6,63*sqrt(25))=0,01

B4: t>-$t_α$. Vậy không thể bác bỏ H0. Tức mỗi ngày trung bình ít nhất 15 hoa.

b/ Nhớ nhé kiểm định tỉ lệ, số ngày bình thường = 2+7+7+3=19; tỉ lệ f=19/25. Ở đây H_0: p=p_0=50%=0.5; H_1: p ≠ 0.5

Chú ý: kiểm định 2 phía (dùng α/2). B1: Tính $z_{α/2}$ qua $Φ(z_{α/2})$=1 – α/2 = 0,95 (vì α =10%=0,1); $z_{α/2}$=1.645;

B2: z=(19/25-50%)*sqrt(25)/sqrt(50%*(1-50%))=2,6;

B3: |z|>$z_{α/2}$=1.645. Tức bác bỏ H0. Vậy những ngày bình thường (13-17 hoa) chiếm 50% chưa tin cậy.