


Giảng viên tổng hợp đề: (Chữ ký và Họ tên)	Ngày ra đề 10/05/2022	Người phê duyệt: (Chữ ký, Chức vụ và Họ tên)	Ngày duyệt đề:
--	-----------------------	--	----------------

(phần phía trên cần che đi khi in sao đề thi)

 TRƯỜNG ĐH BÁCH KHOA – ĐHQG-HCM KHOA KHOA HỌC ỨNG DỤNG	ĐỀ THI CUỐI KỲ		Học kỳ/năm học		2	2021-2022
			Ngày thi		13/05/2022	
	Môn học	XÁC SUẤT THỐNG KÊ				
	Mã môn học	MT2013				
	Thời lượng	100 phút	Mã đề	DT		
Ghi chú: <ul style="list-style-type: none">- Được sử dụng 02 tờ tài liệu A4 do sinh viên tự soạn (không sử dụng tài liệu được viết tay).- Được sử dụng các bảng tra số và máy tính bỏ túi.- Các số gần đúng lấy tròn 4 chữ số phần thập phân.- Nộp lại đề thi cùng với bài làm						

Câu hỏi 1 (L.O.2.1): (2 điểm)

Giả thiết rằng số lỗi X trên mỗi trang sách của một xưởng in là biến ngẫu nhiên tuân theo phân phối Poisson. Có 70% các trang sách không chứa một lỗi nào, và 30% các trang sách chứa ít nhất một lỗi.

- Tìm số lỗi trung bình trên một trang sách của xưởng in đó.
- Tìm tỉ lệ trang sách của xưởng in có từ 3 lỗi trở lên.
- Giả sử trong 1 quý, xưởng in đã hoàn thành 12000 trang sách in. Tìm xác suất của biến cố có ít nhất 9000 trang trong đó không có lỗi nào.

Câu hỏi 2 (L.O.2.1): (4 điểm)

Có một mẫu áo sơ-mi được gia công tại phân xưởng A và phân xưởng B của một nhà máy. Thời gian mỗi cái áo được hoàn thành ở từng phân xưởng là biến ngẫu nhiên tuân theo phân phối chuẩn.

Ở phân xưởng A, người ta chọn ngẫu nhiên 6 cái áo và ghi nhận được thời gian hoàn thành của mỗi áo như sau:

19 21 22 23 22 21 (đơn vị: phút)

- Hãy tìm khoảng ước lượng cho thời gian gia công trung bình một cái áo ở phân xưởng A, với độ tin cậy 95%.
- Ở phân xưởng B, người ta lấy ngẫu nhiên 10 áo và tính được thời gian hoàn thành trung bình của chúng là 21.8 (phút) và phương sai mẫu 2.4 (phút²). Với mức ý nghĩa 5%, có thể xem như phân xưởng A gia công mẫu áo này nhanh hơn so với phân xưởng B hay không?
- Tỉ lệ sản phẩm có lỗi ở nhà máy thông thường là 8%. Sau khi cải tiến kỹ thuật, người ta khảo sát ngẫu nhiên 90 sản phẩm thì thấy chỉ có 4 sản phẩm lỗi. Với mức ý nghĩa 5%, có thể nói rằng việc cải tiến là hiệu quả hay không?

Câu hỏi 3 (L.O.2.1): (2 điểm)

Khi đo một số cây thông ba lá trong rừng thông ở Lâm đồng, người ta ghi nhận lại số liệu trong bảng dưới đây. Gọi X(cm) là đường kính thân cây và Y(m) là chiều cao tương ứng.

X (cm)	6.5	8.5	10.5	12	13.5	15.5	17
Y (m)	3.5	7.5	8.5	9	10.5	11	11.5

Tìm hệ số tương quan mẫu, phương trình hồi quy tuyến tính mẫu Y theo X, và khoảng tin cậy 95% cho hệ số góc của đường hồi quy tuyến tính.

Câu hỏi 4) (L.O.2.1): (2 điểm)

Hãy sử dụng phương pháp Anova để so sánh chiều cao các cây thông ba lá ở 3 khu vực rừng khác nhau, các số liệu được lấy ở những cây thông đã 6 năm tuổi. Nêu kết luận với mức ý nghĩa 1%.

Địa điểm	Chiều cao của cây (m)			
Khu vực 1	7.2	7.1	6.9	6.9
Khu vực 2	7.1	7.2	7.2	7.3
Khu vực 3	7.1	6.9	6.9	6.8

Xem như các giả thiết của bài toán Anova được thỏa mãn.

--- HẾT ---



ĐÁP ÁN:

Câu hỏi 1 (L.O.2.1): (2 điểm)

a) $P(X=0) = e^{-\lambda} = 0.7 \Rightarrow \lambda = -\ln(0.7) = 0.3567$

b) $1 - P(0 \leq X \leq 2) = 1 - \sum_{x=0}^2 \frac{e^{\ln(0.7)} \times (-\ln 0.7)^x}{X!} = 0.0058$

c) $1 - \Phi\left(\frac{9000 - 0.5 - 8400}{\sqrt{2520}}\right) \approx 0$

Cách 1:

Cách 2: $\Phi\left(\frac{12000 + 0.5 - 8400}{\sqrt{2520}}\right) - \Phi\left(\frac{9000 - 0.5 - 8400}{\sqrt{2520}}\right) \approx 0$

Câu hỏi 2 (L.O.2.1): (4 điểm)

a) Hãy tìm khoảng ước lượng cho thời gian gia công trung bình một cái áo ở phân xưởng A, với độ tin cậy 95%.

$n = 6$ $\bar{x}_{tb} = 21.3333$ $s = 1.3663$ $s^2 = 1.8667$

$t_{0.025}(5) = 2.571$ $\varepsilon = (\text{ghi công thức}) = 1.4340$

KUL cần tìm: (ghi công thức) $= (19.8993; 22.7674)$

b) $n_1 = 6$ $\bar{x}_{1_tb} = 21.3333$ $s_1^2 = 1.8667$

$n_2 = 10$ $\bar{x}_{2_tb} = 21.8$ $s_2^2 = 2.4$

Đây là bài toán kiểm tra so sánh trung bình 2 mẫu độc lập, các tổng thể tuân theo phân phối chuẩn, phương sai tổng thể chưa biết nhưng được giả định bằng nhau

(do $s_1^2 / s_2^2 = 0.7778 \in [0.5; 2]$)

Phương sai mẫu gộp là $s_p^2 \equiv s^2 = \dots (\text{ghi ct}) \dots = 2.2095$

Gọi μ_1 là thời gian trung bình để gia công 1 chiếc áo ở phân xưởng 1.

μ_2 là thời gian trung bình để gia công 1 chiếc áo ở phân xưởng 2.

Gt không H_0 : $\mu_1 = \mu_2$ (hay $\mu_1 \geq \mu_2$; hay $\mu_1 - \mu_2 \geq 0$)

Gt đối H_1 : $\mu_1 < \mu_2$

Miền bác bỏ $RR = (-\infty; -t_{0.05}(14)) = (-\infty; -1.761)$

Tính giá trị thống kê kiểm định: $t_{qs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = -1.2809$

Do $t_{qs} \notin RR$ nên chưa đủ cơ sở để bác bỏ $H_0 \dots$

c) Gọi p là tỉ lệ sản phẩm có lỗi sau khi cải tiến. $n = 90$; $f = 4/90$

Gt không H_0 : $p = 8\%$ (hay $p \geq 8\%$)

Gt đối H_1 : $p < 8\%$

Miền bác bỏ $RR = (-\infty; -z_{0.05}) = (-\infty; -1.645)$

Tính giá trị thống kê kiểm định: $z_{qs} = \frac{f - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{n} = -1.2433$

Do $z_{qs} \notin RR$ nên chưa đủ cơ sở để bác bỏ $H_0 \dots$

Câu hỏi 3 (L.O.2.1): (2 điểm)

- Các giá trị đặc trưng mẫu (Các giá trị nào không dùng để tính đáp số thì không cần ghi)

$$\begin{array}{llll} n=7 & \bar{x}=11.9286 & \sigma_x=3.4685 & s_x=3.7464 \\ & \bar{y}=8.7857 & \sigma_y=2.5334 & s_y=2.7364 \end{array} \quad \Sigma xy=791.5\dots$$

$$S_{xx} = (\text{ghi ct tính đơn giản}) = 84.2123$$

$$(\text{Ví dụ: } S_{xx} = \sum_{i=1}^7 (x_i - \bar{x})^2 = 84.2123 \quad \text{hoặc } S_{xx} = (n-1) \cdot s_x^2 = 84.2123)$$

$$S_{xy} = (\text{ghi ct tính đơn giản hoặc ghi các biểu thức số}) = 57.8929$$

$$S_{yy} = (\text{ghi ct tính}) = 44.9286$$

$$SSR = (\text{ghi ct tính}) = 39.7983$$

$$SSE = (\text{ghi ct tính}) = 5.1303$$

- Đáp số: $r_{xy} = (\text{ghi ct tính}) = 0.9412$.
 $a = (\text{ghi ct tính}) = 0.5854$ (còn kí hiệu là $\hat{\beta}_0$)
 $b = (\text{ghi ct tính}) = 0.6874$ (còn kí hiệu là $\hat{\beta}_1$)
 \Rightarrow pthqtt: $y^{\wedge} = 0.5854 + 0.6874x$

Khoảng ước lượng 95% cho hệ số góc (ghi ct tính đơn giản và các giá trị trung gian)...
(0.4037; 0.9712)

Câu hỏi 4 (L.O.2.1): (2 điểm)

- Gọi $\mu_1; \mu_2; \mu_3$ lần lượt là chiều cao trung bình của các cây thông 6 năm tuổi ở 3 khu vực trên.

$$\text{Gt không } H_0: \mu_1 = \mu_2 = \mu_3$$

$$\text{Gt đối } H_1: \exists \mu_i \neq \mu_j \quad \text{với } i \neq j$$

$$k = 3; n_i = 4; N = 12$$

- Miền bác bỏ: $RR = (f_{0.01}(2; 9); +\infty) = (8.02; +\infty)$

$$\bar{x}_1 = 7.025 \quad \bar{x}_2 = 7.2 \quad \bar{x}_3 = 6.925 \quad \bar{x} = 7.05$$

$$SSB = (\text{ghi ct tính}) = 0.155 \quad df_B = (\text{ghi ct tính}) = 2 \quad MSB = (\text{ghi ct}) = 0.0775$$

$$SSW = (\text{ghi ct tính}) = 0.135 \quad df_W = (\text{ghi ct tính}) = 9 \quad MSW = (\text{ghi ct}) = 0.015$$

Tính giá trị thống kê kiểm định (hay là tiêu chuẩn kiểm định):

$$F_{qs} = (\text{ghi ct tính}) = 5.1667$$

- Kết luận: Do $F_{qs} \notin RR$ nên chưa đủ cơ sở để bác bỏ giả thiết không H_0 .
Chưa bác bỏ được giả thiết chiều cao trung bình của cây thông ở 3 khu vực là bằng nhau.