ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
BỘ MÔN TOÁN – LÝ

ĐÁP ÁN ĐỀ THI CUỐI KỲ
MÔN XÁC SUẤT THỐNG KÊ
Học kỳ 1, năm học 2021-2022.
Ngày thi: .../.../ 2022
Thời gian làm bài: 90 phút.
Được sử dụng tài liệu giấy.

Câu 1: *(2 điểm)* Giả sử X và Y là hai biến ngẫu nhiên có phân phối xác suất đồng thời như sau:

$P(X,Y)$		X		
		1	2	3
Y	1	0.01	0.02	0.08
	2	0.01	0.02	0.08
	3	0.07	0.08	0.63

a. Tính $P(X \leq 2, Y \leq 3)$.
b. *X, Y* có độc lập không? Vì sao?
c. Tính $P(X > 1|Y = 3)$.
d. Tính giá trị trung bình của X.

Đáp án tham khảo:

a. $P(X \leq 2, Y \leq 3) = P(1,1) + P(1,2) + P(1,3) + P(2,1) + P(2,2) + P(2,3)$
$= 0.01 + 0.01 + 0.07 + 0.02 + 0.02 + 0.08 = 0.21$

b. $P(X = 1) = 0.9, P(Y = 1) = 0.11$
Vì $P(X = 1, Y = 1) = 0.01 \neq P(X = 1).P(Y = 1) = 0.099$
nên X, Y không độc lập với nhau.

c. $P(X > 1|Y = 3) = \dfrac{P(X > 1|Y = 3)}{P(Y = 3)} = \dfrac{0.08 + 0.63}{0.07 + 0.08 + 0.63} = \dfrac{71}{78}$

d.

X	1	2	3
P	0.09	0.12	0.79

$E(X) = 0{,}09.1 + 0{,}12.2 + 0{,}79.3 = 2{,}7$

Câu 2: *(2 điểm)* Cho (X, Y) là vecto ngẫu nhiên liên tục có hàm mật độ đồng thời là:

$$f(x,y) = \begin{cases} \dfrac{k}{x^2 y} & \text{nếu } 1 \leq x \leq 2 \text{ và } 1 \leq y \leq e \\ 0 & \text{nếu } (x,y) \text{ khác.} \end{cases}$$

a. Tìm *k.*

b. Tìm hàm mật độ thành phần của X. Tính $P\left(X < \dfrac{3}{2}\right)$.

c. Tìm $P\left(X < \dfrac{3}{2}, Y > \sqrt{e}\right)$. Tính $P\left(Y > \sqrt{e} \middle| X < \dfrac{3}{2}\right)$.

Đáp án tham khảo:

a. $\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1 \Leftrightarrow \int_1^e \int_1^2 \frac{k}{x^2y}\,dx\,dy = 1$

$$\Leftrightarrow \int_1^e \left(\frac{-k}{xy}\Big|_1^2\right) dy = 1 \Leftrightarrow \int_1^e \left(\frac{-k}{2y} - \frac{-k}{y}\right) dy = 1$$

$$\Leftrightarrow \left[\frac{-k}{2} ln|e| - (-kln|e|)\right] = 1 \Rightarrow k = 2$$

b. $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \begin{cases} \int_1^e \frac{2}{x^2y}dy = \frac{2}{x^2} & \forall x \in [1,2] \\ 0 & \forall x \notin [1,2] \end{cases}$

$c. P\left(X < \frac{3}{2}, Y > \sqrt{e}\right) = \int_{\sqrt{e}}^{e} \int_1^{\frac{3}{2}} \frac{2}{x^2y} dxdy = \int_{\sqrt{e}}^{e} \left[\frac{-2}{\frac{3}{2}y} - \left(\frac{-2}{y}\right)\right] dy = \frac{1}{3}$

$$P\left(Y > \sqrt{e} \middle| X < \frac{3}{2}\right) = \frac{P\left(Y > \sqrt{e}, X < \frac{3}{2}\right)}{P\left(X < \frac{3}{2}\right)} = \frac{\frac{1}{3}}{\int_1^e \int_1^{\frac{3}{2}} \frac{2}{x^2y} dxdy} = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$$

Câu 3: *(1 điểm)* Dữ liệu dưới đây thể hiện các khoản đầu tư X (tính bằng 1000 đô la) vào việc phát triển phần mềm mới của một công ty máy tính theo năm Y.

Y	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
X	17	23	31	29	33	39	39	40	41	44	47

a. Tính hệ số tương quan và nhận xét về tính tuyến tính của X và Y (mạnh hay yếu, nghịch biến hay đồng biến)

b. Tìm phương trình hồi qui tuyến tính của Y theo X. Dự đoán giá trị đầu tư của công ty vào năm 2023.

Đáp án tham khảo:

a. Hệ số tương quan mẫu: r = 0,9597 > 0,9 → X, Y tương quan mạnh và đồng biến.

b. Phương trình hồi qui tuyến tính của Y theo X:

Y = a + bX= 2003,92 + 0,3469X (*)

Thay Y = 2023 vào (*), ta được giá trị đầu tư X là 55,16 (đô la)

Câu 4: *(1.5 điểm)* Một công ty đưa ra tuyên bố rằng một loại sản phẩm mới mà họ đã phân phối trong toàn bộ hệ thống đại lý bán lẻ của họ sau một năm đã có hơn 90% sản phẩm được khách hàng quan tâm và chọn mua. Kiểm tra ngẫu nhiên 300 đại lý thấy có đến 40 đại lý không bán được một sản phẩm loại này, trong khi số đại lý còn lại gần như đã bán hết sản phẩm. Với mức ý nghĩa 5%, hãy kiểm định xem tuyên bố trên của công ty có quá sự thật không?

- ***Đáp án tham khảo:***

  Sử dụng phương pháp kiểm định tỉ lệ

  $H_O: p = 0.9$ (hoặc $H_O: p \geq 0.9$) , $H_1: p < 0.9$

  $$k = 40, \quad n = 300, \quad f = 1 - \frac{40}{300} = \frac{13}{15}, \quad \alpha = 5\% = 0.05$$

  $$z = \frac{f - p_o}{\sqrt{p_o(1 - p_o)}}.\sqrt{n} = \frac{\frac{13}{15} - 0{,}9}{\sqrt{0{,}9(1 - 0{,}9)}}.\sqrt{300} = -1{,}92$$

  $$\phi(z_\alpha) = 1 - \alpha = 0{,}95 \xrightarrow{\text{tra bảng A4}} -z_\alpha = -1{,}645$$

  Vì $-z_\alpha > z \Longrightarrow$ bác bỏ H_o, chấp nhận H_1

  Vậy tuyên bố của công ty là quá sự thật với mức ý nghĩa 5%.

Câu 5: *(1.5 điểm)* Một giám đốc một công ty nói rằng mức lương trung bình của các kỹ sư IT của công ty ông là 24,6 triệu đồng/tháng. Giả sử mức lương hàng tháng của các kỹ sư IT có phân phối chuẩn với độ lệch chuẩn 2,52 triệu đồng. Chọn ngẫu nhiên 25 kỹ sư IT trong công ty, người ta thấy rằng lương trung bình của nhóm kỹ sư này là 25,33 triệu đồng/tháng. Hỏi có đủ căn cứ để chấp nhận tuyên bố của giám đốc này với mức ý nghĩa 1% không?

- ***Đáp án tham khảo:***

  Sử dụng phương pháp kiểm định giá trị trung bình (trường hợp 3)

  $H_O: \mu = 24.6$, $H_1: \mu \neq 24.6$

  $$\bar{x} = 25.33, \quad \sigma = 2.52, \quad n = 25, \quad \alpha = 1\% = 0.01$$

  $$z = \frac{\bar{x} - \mu_o}{\sigma}.\sqrt{n} = \frac{25{,}33 - 24{,}6}{2{,}52}.\sqrt{25} = 1{,}448$$

  $$\phi\left(z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} = 1 - \frac{0{,}01}{2} = 0{,}995 \xrightarrow{\text{tra bảng A4}} z_{\frac{\alpha}{2}} = 2{,}575$$

  Vì $z_{\frac{\alpha}{2}} > |z| \Rightarrow$ chấp nhận H_o

  Vậy tuyên bố của giám đốc công ty là đủ căn cứ để chấp nhận.

Câu 6: *(2 điểm)* Để ước lượng tỷ lệ phế phẩm của một dây chuyền sản xuất, người ta chọn ngẫu nhiên 300 sản phẩm và kiểm tra thấy có 20 phế phẩm.

a. Với độ tin cậy 99%, hãy tìm khoảng ước lượng cho tỷ lệ phế phẩm của dây chuyền trên.

b. Để sai số của việc ước lượng với độ tin cậy 95% không vượt quá 0.01 thì ta phải kiểm tra tối thiểu bao nhiêu sản phẩm.

- ***Đáp án tham khảo:***

  a. *Sử dụng khoảng tin cậy cho tỉ lệ (ước lượng đối xứng)*

$n = 300, \qquad k = 20, \qquad f = \frac{k}{n} = \frac{1}{15} =, \qquad 1 - \alpha = 0.99 \Rightarrow \alpha = 0.01$

$\phi\left(z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} = 1 - \frac{0,01}{2} = 0,995 \xrightarrow{\text{tra bảng A4}} z_{\frac{\alpha}{2}} = 2,575$

Độ tin cậy: $\varepsilon = z_{\frac{\alpha}{2}}.\sqrt{\frac{f(1-f)}{n}} = 2,575.\sqrt{\frac{\frac{1}{15}\left(1-\frac{1}{15}\right)}{300}} = 0,0371$

Khoảng ước lượng cho tỷ lệ phế phẩm là: $(0,0297; 0,1037)$

b. Tương tự với độ tin cậy 95% ta có: $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

$\phi\left(z_{\frac{\alpha}{2}}\right) = 0,975 \xrightarrow{\text{tra bảng A4}} z_{\frac{\alpha}{2}} = 1,96$

$\Rightarrow \varepsilon = 1,96.\sqrt{\frac{f(1-f)}{n}} \leq 1,96.\frac{1}{2\sqrt{n}} \quad \forall f \in [0,1]$

$\Rightarrow n \geq \left(\frac{1,96}{2\varepsilon}\right)^2 \geq \left(\frac{1,96}{2.0,01}\right)^2 = 9604$

Vậy cần kiểm tra tối thiểu 9604 sản phẩm để đạt sai số như yêu cầu.