

Đáp án đề thi: KẾT THÚC HỌC PHẦN

Mã đề: 02

Môn thi: **Xác suất thống kê Y học**

Khóa/Lớp: **K26Y-DD01** - Học kỳ: **1** - Năm học: **2020-2021**

Thời gian thi: 45 phút ; 60 phút ; **75 phút** ; 90 phút ; 120 phút

Hình thức thi: **Tự luận**  Trắc nghiệm  Tự luận + Trắc nghiệm

Được sử dụng tài liệu

Không sử dụng tài liệu

Người biên soạn đề  
(Ký và ghi rõ họ tên)

Ths. Đỗ Lư Công Minh

Trưởng Khoa/Bộ môn duyệt  
(Ký và ghi rõ họ tên)

Ths. Đinh Tiên Liêm

*Lưu ý: - Các đáp án làm tròn 4 chữ số thập phân sau dấu phẩy.*

## I. PHẦN XÁC SUẤT

### Câu 1: (1 điểm)

Một bệnh viện phụ sản, số sản phụ sinh trong 1 tuần có phân phối Poisson với trung bình là 28. Tính xác suất trong 1 ngày có:

- Không sản phụ nào đến sinh.
- Đúng 1 sản phụ đến sinh.
- Nhiều hơn 1 sản phụ đến sinh.

**Giải**

Gọi  $X$  là số sản phụ sinh trong 1 ngày tại bệnh viện, ta có:  $X \sim P(4)$  (0,25 điểm)

a) Xác suất trong 1 ngày không có sản phụ nào đến sinh:

$$P(X=0) = \frac{e^{-4} \times 4^0}{0!} \approx 0,0183 \quad (0,25 \text{ điểm})$$

b) Xác suất trong 1 ngày có đúng 1 sản phụ đến sinh:

$$P(X=1) = \frac{e^{-4} \times 4^1}{1!} \approx 0,0733 \quad (0,25 \text{ điểm})$$

c) Xác suất trong 1 ngày có nhiều hơn 1 sản phụ đến sinh:

$$P(X > 1) = 1 - [P(X=0) + P(X=1)] = 1 - (0,0183 + 0,0733) \approx 0,9084 \quad (0,25 \text{ điểm})$$

### Câu 2: (2 điểm)

Một xét nghiệm chẩn đoán có độ nhạy 95% và độ chuyên 50%. Biết giá trị tiên đoán dương là 0,7; hãy tìm giá trị tiên đoán âm.

**Giải**

Giả sử  $P(B^+) = p, p \in (0, 1)$ . Xét bảng  $2 \times 2$  đối với xét nghiệm  $T$ : (0,5 điểm)

	$B^+$	$B^-$	
$T^+$	$95p$	$50(1-p)$	$50+45p$
$T^-$	$5p$	$50(1-p)$	$50-45p$
	$100p$	$100(1-p)$	$N=100$

Theo giả thiết, giá trị tiên đoán dương là 0,7:

$$PV^+ = P(B^+ | T^+) = \frac{95p}{50+45p} = 0,7 \quad (0,5 \text{ điểm})$$

$$\Leftrightarrow 95p = 0,7(50+45p) \Leftrightarrow 95p = 35 + 31,5p \Leftrightarrow 63,5p = 35 \Leftrightarrow p = \frac{70}{127} \quad (0,5 \text{ điểm})$$

Suy ra giá trị tiên đoán âm:

$$PV^- = P(B^- | T^-) = \frac{50(1-p)}{50-45p} = \frac{50 \times \left(1 - \frac{70}{127}\right)}{50 - 45 \times \frac{70}{127}} = \frac{57}{64} = 0,890625 \quad (0,5 \text{ điểm})$$

## II. PHẦN THỐNG KÊ

Câu 3: (3 điểm)

Năng suất cà phê ở vùng A là biến ngẫu nhiên có phân phối xấp xỉ chuẩn. Ở vùng A, người ta thu hoạch ngẫu nhiên một số hecta cà phê và có được các số liệu sau:

Năng suất (tạ/ha)	25	27	29	31	33	37
Số hecta	12	18	24	28	10	8

a) Tìm cỡ mẫu, trung bình mẫu, độ lệch chuẩn mẫu (hiệu chỉnh) và phương sai mẫu (hiệu chỉnh) của năng suất cà phê vùng A.

b) Với độ tin cậy 96%, hãy ước lượng năng suất trung bình của cà phê vùng A.

c) Biết năng suất từ 30 tạ/ha trở lên được gọi là năng suất cao. Hãy ước lượng tỉ lệ hecta cà phê có năng suất cao trong vùng A với độ tin cậy 97%.

**Giải**

a) Theo công thức tính cỡ mẫu  $n$ , trung bình mẫu  $\bar{x}$  và phương sai mẫu  $s_x^2$ , ta có:

$$n = 100; \bar{x} = 29,76; s_x^2 = 10,0024 \Rightarrow s_x = 3,1627 \quad (0,25 \text{ điểm} \times 4)$$

b) Các đặc trưng mẫu:  $n = 100; \bar{x} = 29,76; s_x = 3,1627 \quad (0,25 \text{ điểm})$

$$\text{Độ tin cậy: } 1 - \alpha = 0,96 \Rightarrow \frac{1 - \alpha}{2} = 0,48 \Rightarrow z_{\alpha/2} = 2,05 \quad (0,25 \text{ điểm})$$

$$\text{Độ chính xác: } \varepsilon = 2,05 \times \frac{3,1627}{\sqrt{100}} \approx 0,6484 \quad (0,25 \text{ điểm})$$

Khoảng tin cậy 96% cho năng suất cà phê trung bình ở vùng A:



$$(\mu_1, \mu_2) = (29,76 - 0,6484; 29,76 + 0,6484) = (29,1116; 30,4084) \text{ (tạ/ha)} \quad (0,25 \text{ điểm})$$

c) Tỷ lệ mẫu:  $f_n = \frac{m}{n} = \frac{46}{100} = 0,46 \quad (0,25 \text{ điểm})$

Độ tin cậy:  $1 - \alpha = 0,97 \Rightarrow \frac{1 - \alpha}{2} = 0,485 \Rightarrow z_{\alpha/2} = 2,17 \quad (0,25 \text{ điểm})$

Độ chính xác:  $\varepsilon = 2,17 \times \sqrt{\frac{0,46 \times (1 - 0,46)}{100}} \approx 0,1082 \quad (0,25 \text{ điểm})$

Khoảng tin cậy 97% cho tỷ lệ hecta cà phê có năng suất cao trong vùng A:

$$(p_1, p_2) = (0,46 - 0,1082; 0,46 + 0,1082) = (0,3518; 0,5682) \quad (0,25 \text{ điểm})$$

**Câu 4: (2 điểm)**

a) Khảo sát ngẫu nhiên 150 người thấy có 12 người mắc bệnh K phổi. Hỏi quan sát này có phù hợp với tỷ lệ bệnh K phổi là 7% trong cộng đồng? Yêu cầu tính giá trị kiểm định và đưa ra kết luận với mức ý nghĩa 1%.

b) Một máy phân tích huyết học được gọi là không đạt yêu cầu, khi máy chạy hết công suất, trung bình máy phân tích 99 mẫu máu/ngày. Chọn một máy phân tích huyết học, cho chạy thử 1 tuần hết công suất, kết quả như sau:

Thứ	2	3	4	5	6	7	Chủ nhật
Số mẫu máu/ngày	104	93	97	101	105	95	105

Hỏi máy phân tích huyết học trên đạt yêu cầu chưa? Yêu cầu tính giá trị kiểm định và đưa ra kết luận với mức ý nghĩa 4%.

**Giải**

a) Gọi  $p$  là tỷ lệ bệnh nhân bị K phổi.

Giả thuyết  $H_0: "p = 0,07"$ , đối thuyết  $H_1: "p \neq 0,07"$  (0,25 điểm)

Tỷ lệ mẫu:  $f_n = \frac{m}{n} = \frac{12}{150} = 0,08$

Giá trị của tiêu chuẩn kiểm định:

$$z = \frac{|f_n - p_0|}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{|0,08 - 0,07|}{\sqrt{\frac{0,07(1 - 0,07)}{150}}} \approx 0,48 \quad (0,25 \text{ điểm})$$

Phân vị chuẩn mức ý nghĩa 0,5%:  $\alpha = 1\% \Rightarrow \frac{1 - \alpha}{2} = 0,495 \Rightarrow z_{\alpha/2} = 2,58 \quad (0,25 \text{ điểm})$

Do  $z < z_{\alpha/2}$  nên ta chưa đủ cơ sở để bác bỏ giả thuyết  $H_0$ .

Kết luận: Kết quả theo quan sát phù hợp với tỷ lệ bệnh K phổi trong cộng đồng. (0,25 điểm)

b) Gọi  $\mu$  là số mẫu máu phân tích được trung bình trong 1 ngày.

Giả thuyết  $H_0$ : " $\mu = 99$ ", đối thuyết  $H_1$ : " $\mu > 99$ " (0,25 điểm)

Các đặc trưng mẫu:  $n = 7$ ;  $\bar{x} = 100$ ;  $s_x = 5$

Giá trị của tiêu chuẩn kiểm định:

$$t = \frac{|\bar{x}_n - \mu_0|}{\frac{s}{\sqrt{n}}} = \frac{|100 - 99|}{\frac{5}{\sqrt{7}}} \approx 0,5292 \quad (0,25 \text{ điểm})$$

Phân vị Student mức ý nghĩa 4%, bậc tự do 6:

$$\alpha = 4\% \Rightarrow t_{(0,04; 6)} = 2,1043 \quad (0,25 \text{ điểm})$$

Do  $t < t_{(0,04; 6)}$  nên ta chưa đủ cơ sở để bác bỏ giả thuyết  $H_0$ .

Kết luận: Máy phân tích huyết học không đạt yêu cầu. (0,25 điểm)

### Câu 5: (1 điểm)

Nghiên cứu thải trừ một số loại thuốc, ta có số liệu như sau:

$X$	0,5	0,5	1	1	2	2	3	3	4	4
$Y$	0,42	0,45	0,35	0,33	0,25	0,22	0,20	0,20	0,15	0,17

( $X$  là thời gian thải trừ, đơn vị tính: giờ;  $Y$  là nồng độ thuốc, đơn vị  $\mu\text{g/ml}$ )

a) Tính hệ số tương quan tuyến tính mẫu giữa  $X$  và  $Y$ . Hỏi  $X$  và  $Y$  có mối tương quan tuyến tính nghịch không?

b) Tìm phương trình hồi quy tuyến tính mẫu của  $Y$  theo  $X$ .

c) Nồng độ thuốc trung bình dự đoán được sau 2 giờ thải trừ là bao nhiêu? (ước lượng điểm)

### Giải

a) Hệ số tương quan tuyến tính mẫu:  $r_{XY} = -0,9432$  (0,25 điểm)

Do  $r_{XY} < 0$  nên ta có thể dự đoán: giữa  $X$  và  $Y$  có mối tương quan tuyến tính nghịch

(0,25 điểm)

b) Phương trình hồi quy tuyến tính của  $Y$  theo  $X$ :  $\bar{y}_x = 0,4301 - 0,0743x$  (0,25 điểm)

c) Ta có:  $\bar{y}_2 = 0,4301 - 0,0743 \times 2 = 0,2815$  ( $\mu\text{g/ml}$ ) (0,25 điểm)

---

**Cho biết:** Giá trị của hàm tích phân Laplace  $\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$  như sau:

$$\varphi(1,96) = 0,475; \varphi(2,05) = 0,48; \varphi(2,17) = 0,485; \varphi(2,33) = 0,49; \varphi(2,58) = 0,495$$

**Bảng tra phân vị Student**  $P[T > t_{(\alpha, n)}] = \alpha$  với  $T \sim T(n)$

$$t_{(0,04; 6)} = 2,1043$$