

TRƯỜNG ĐẠI HỌC VĂN LANG

Đơn vị: Khoa Khoa học cơ bản

**ĐỀ THI VÀ ĐÁP ÁN**  
**THI KẾT THÚC HỌC PHẦN**  
**Học kỳ 2, năm học 2023-2024**

**I. Thông tin chung**

Tên học phần:	Xác suất thống kê Y học		
Mã học phần:	71MATM10033	Số tin chỉ:	3
Mã nhóm lớp học phần:	232_71MATM10033_01,02		
Hình thức thi: <b>Trắc nghiệm kết hợp Tự luận</b>	Thời gian làm bài:	<b>75</b>	phút
<i>Thí sinh được tham khảo tài liệu:</i>	<input checked="" type="checkbox"/> Có	<input type="checkbox"/> Không	

**Cách thức nộp bài phần tự luận:**

Nộp bài làm riêng cho từng câu hỏi dưới 1 trong 2 dạng sau đây (Không cần chép lại đề bài)

1/ Upload file bài làm (word, excel, pdf, ...);

2/ Upload hình ảnh bài làm.

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

- Sinh viên tự chuẩn bị bảng tra cứu các phân vị thường dùng.

**II. Các yêu cầu của đề thi nhằm đáp ứng CLO**

Ký hiệu CLO	Nội dung CLO	Hình thức đánh giá	Trọng số CLO trong thành phần đánh giá (%)	Câu hỏi thi số	Điểm số tối đa	Lấy dữ liệu đo lường mức đạt PLO/PI
(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>CLO1</b>	Áp dụng kiến thức cơ bản về xác suất trong việc chẩn đoán và quyết định lâm sàng	Thi cuối kỳ	30%	1-2 TN, 3TL	3	PI1.1
<b>CLO2</b>	Vận dụng kiến thức về thống kê trong quá trình thu thập và phân tích dữ liệu y khoa	Thi cuối kỳ	40%	3-10 TN	4	PI1.1
<b>CLO3</b>	Sử dụng phần mềm thống kê hỗ trợ cho việc phân tích dữ liệu y khoa	Thi cuối kỳ	30%	1-2 TL	3	PI4.2

### III. Nội dung câu hỏi thi

#### PHẦN TRẮC NGHIỆM (10 câu hỏi, mỗi câu 0,5 điểm, tổng số điểm: 5 điểm)

Cho biết tỉ lệ bệnh nhân bị ung thư vú đáp ứng điều trị với CMF là 0,6. Điều trị 5 bệnh nhân ung thư vú bằng CMF. Tính xác suất có 4 người đáp ứng.

- A. 0,2592
- B. 0,9222
- C. 0,8000
- D. 0,3370

ANSWER: A

Một người phải có điểm bài kiểm tra IQ nằm trong 2% điểm cao nhất của tổng thể thì mới đủ điều kiện trở thành hội viên trong Mensa, một hội quốc tế IQ-cao. Nếu điểm IQ có phân phối chuẩn với trung bình là 100 và độ lệch chuẩn là 15 thì một người phải có điểm số tối thiểu là bao nhiêu để đủ điều kiện gia nhập Mensa? (kết quả làm tròn đến hàng đơn vị)

- A. 131
- B. 130
- C. 135
- D. 125

ANSWER: A

Theo Dubbo Osteoporosis Epidemiology Study, dữ liệu về mật độ xương ở cổ xương đùi của 50 nữ và 50 nam được ghi nhận như sau: (đơn vị tính:  $g/cm^3$ )

Dữ liệu về mật độ xương ở cổ xương đùi của 50 nữ:

0,74	0,63	0,77	0,7	0,82	0,67	0,58	0,85	0,72	0,71
0,37	0,64	0,78	0,75	0,51	0,65	0,74	0,82	0,69	1
0,69	0,72	0,56	0,61	0,65	0,77	0,61	0,64	0,73	0,68
0,47	0,71	0,66	0,86	0,82	0,69	0,58	0,63	0,83	0,75
0,7	0,82	0,7	0,68	0,74	0,73	0,77	0,84	0,81	0,71

Dữ liệu về mật độ xương ở cổ xương đùi của 50 nam:

1,08	1,01	0,81	1,01	1,16	0,85	0,94	0,96	0,83	0,99
1,17	1,19	1,14	1	1,1	0,71	0,93	0,77	0,71	1,04
1,06	0,97	0,83	0,98	1,02	1,06	0,66	0,92	0,86	0,74
1	0,98	0,84	0,83	0,77	0,97	1,08	1,06	0,72	0,9

0,9	0,81	0,97	0,86	0,79	0,9	1,06	1	0,71	0,86
-----	------	------	------	------	-----	------	---	------	------

Tính mật độ xương ở cổ xương đùi trung bình của 100 người được khảo sát trong mẫu trên.

- A. 0,8181 ( $\text{g}/\text{cm}^3$ )
- B. 0,7060 ( $\text{g}/\text{cm}^3$ )
- C. 0,1654 ( $\text{g}/\text{cm}^3$ )
- D. 0,9302 ( $\text{g}/\text{cm}^3$ )

ANSWER: A

Theo Dubbo Osteoporosis Epidemiology Study, dữ liệu về mật độ xương ở cổ xương đùi của 50 nữ và 50 nam được ghi nhận như sau: (đơn vị tính:  $\text{g}/\text{cm}^3$ )

Dữ liệu về mật độ xương ở cổ xương đùi của 50 nữ:

0,74	0,63	0,77	0,7	0,82	0,67	0,58	0,85	0,72	0,71
0,37	0,64	0,78	0,75	0,51	0,65	0,74	0,82	0,69	1
0,69	0,72	0,56	0,61	0,65	0,77	0,61	0,64	0,73	0,68
0,47	0,71	0,66	0,86	0,82	0,69	0,58	0,63	0,83	0,75
0,7	0,82	0,7	0,68	0,74	0,73	0,77	0,84	0,81	0,71

Dữ liệu về mật độ xương ở cổ xương đùi của 50 nam:

1,08	1,01	0,81	1,01	1,16	0,85	0,94	0,96	0,83	0,99
1,17	1,19	1,14	1	1,1	0,71	0,93	0,77	0,71	1,04
1,06	0,97	0,83	0,98	1,02	1,06	0,66	0,92	0,86	0,74
1	0,98	0,84	0,83	0,77	0,97	1,08	1,06	0,72	0,9
0,9	0,81	0,97	0,86	0,79	0,9	1,06	1	0,71	0,86

Tìm phương sai (hiệu chỉnh) của mật độ xương ở cổ xương đùi của 50 nam trong mẫu trên.

- A. 0,0181 ( $\text{g}/\text{cm}^3$ )<sup>2</sup>
- B. 0,0115 ( $\text{g}/\text{cm}^3$ )<sup>2</sup>
- C. 0,0273 ( $\text{g}/\text{cm}^3$ )<sup>2</sup>
- D. 0,1345 ( $\text{g}/\text{cm}^3$ )<sup>2</sup>

ANSWER: A

Giả sử hàm lượng nước  $X$  (%) trong một loại cam là biến ngẫu nhiên có phân phối chuẩn:  $X \sim N(\mu, \sigma^2)$ . Quan sát một mẫu gồm 25 quả cam thuộc loại trên ta thu được số liệu sau:

$X$ (%)	79	80	84,6	86	87,5	89	90
Số quả	2	3	5	7	4	3	1

Hãy tìm một ước lượng không chệch cho phương sai của hàm lượng nước của loại cam trên.

- A. 10,4917 (%<sup>2</sup>)
- B. 10,0720 (%<sup>2</sup>)
- C. 3,2391 (%<sup>2</sup>)
- D. 0,6478 (%<sup>2</sup>)

ANSWER: A

Theo dõi số chuột chết khi cho các lô chuột thí nghiệm sử dụng các liều thuốc có độ tăng dần, người ta thu được số liệu sau:

Liều thuốc (mg/kg)	0,015	0,02	0,025	0,03	0,035	0,04
Số chuột mỗi lô	20	69	95	78	44	20
Số chuột chết	0	11	50	61	37	20

Tìm khoảng tin cậy 99% cho tỉ lệ chuột chết ở liều thuốc 0,025 mg/kg.

- A. [39,41%; 65,85%]
- B. [40,69%; 64,57%]
- C. [47,80%; 62,02%]
- D. [48,49%; 61,33%]

ANSWER: A

Người ta đo nồng độ  $\text{Na}^+$  trên một số người và ghi nhận lại được kết quả như sau:

Nồng độ $\text{Na}^+$ (mmol/L)	129	132	140	138	143	133
Nồng độ $\text{Na}^+$ (tt) (mmol/L)	137	140	143	138	140	141

Nếu muốn sai số của bài toán ước lượng nồng độ  $\text{Na}^+$  trung bình không vượt quá 1 mmol/L với độ tin cậy 95% thì cần quan sát tối thiểu bao nhiêu người?

- A. 95
- B. 131
- C. 128
- D. 75

ANSWER: A

Theo dõi số chuột chết khi cho các lô chuột thí nghiệm sử dụng các liều thuốc có độc tăng dần, người ta thu được số liệu sau:

Liều thuốc (mg/kg)	0,015	0,02	0,025	0,03	0,035	0,04
Số chuột mỗi lô	10	10	10	10	10	10
Số chuột chết	0	2	5	8	9	10

Có ý kiến cho rằng tỉ lệ chuột chết (ở tất cả các liều thuốc) lớn hơn 45%. Với mức ý nghĩa 5%, hãy tính giá trị kiểm định và cho kết luận về ý kiến trên.

- A. Giá trị kiểm định: 1,8170. Chấp nhận ý kiến.  
 B. Giá trị kiểm định: 1,8170. Bác bỏ ý kiến.  
 C. Giá trị kiểm định: 1,3732. Chấp nhận ý kiến.  
 D. Giá trị kiểm định: 1,3732. Bác bỏ ý kiến.

ANSWER: A

Bốn loại thuốc chữa cùng một loại bệnh B, khảo sát trên bốn nhóm bệnh nhân (mỗi nhóm dùng một loại thuốc), thu được kết quả:

	Nhóm 1	Nhóm 2	Nhóm 3	Nhóm 4
Khỏi bệnh	123	95	152	132
Không khỏi bệnh	28	19	63	53

Hỏi hiệu quả của bốn loại thuốc có như nhau không? Yêu cầu tính giá trị kiểm định của phép kiểm định Khi - bình phương và cho kết luận với mức ý nghĩa 5%.

- A. Giá trị kiểm định: 11,0769. Hiệu quả của bốn loại thuốc khác nhau.  
 B. Giá trị kiểm định: 11,0769. Hiệu quả của bốn loại thuốc như nhau.  
 C. Giá trị kiểm định: 0,0113. Hiệu quả của bốn loại thuốc khác nhau.  
 D. Giá trị kiểm định: 0,0113. Hiệu quả của bốn loại thuốc như nhau.

ANSWER: A

Đo đường kính của viên thuốc (đơn vị tính: mm) do hai máy thuộc hai loại dập ra, ta thu được số liệu sau:

Máy I	5,54	5,69	5,62	5,80	5,67	5,52	5,77	5,65		
Máy II	5,64	5,42	5,58	5,52	5,29	5,50	5,67	5,48	5,32	5,44

Độ chính xác của hai máy có như nhau không? Hãy tính giá trị kiểm định và đưa ra kết luận với mức ý nghĩa 5%.

- A. Giá trị kiểm định: 0,6266. Độ chính xác của hai máy như nhau.
- B. Giá trị kiểm định: 0,6266. Độ chính xác của hai máy khác nhau.
- C. Giá trị kiểm định: 1,5958. Độ chính xác của hai máy như nhau.
- D. Giá trị kiểm định: 1,5958. Độ chính xác của hai máy khác nhau.

ANSWER: A

### PHẦN TỰ LUẬN (3 câu hỏi, tổng số điểm: 5 điểm)

#### Câu hỏi 1: (1,5 điểm)

**Yêu cầu:** Đối với bài toán kiểm định so sánh nhiều giá trị trung bình, yêu cầu phát biểu rõ giả thuyết, đối thuyết và dựa vào kết quả phân tích phương sai từ phần mềm Excel (sinh viên tự thực hành và trình bày bảng ANOVA vào bài làm), hãy nêu rõ kết luận. Khuyến khích sinh viên nộp bài làm bằng tập tin Excel.

**Đề bài:** Một nghiên cứu về mối quan hệ giữa vị trí địa lý và chứng trầm cảm được các nhà xã hội học và bác sĩ tâm lý ở Trung tâm Y tế Goworth - ngoại ô New York thực hiện trên đối tượng người trên 65 tuổi. Một mẫu gồm 60 cá nhân, tất cả đều có sức khỏe khá tốt, được chọn: 20 người là cư dân của Florida, 20 người là cư dân của New York và 20 người là cư dân của Bắc Carolina. Mỗi cá nhân trong mẫu được chọn thực hiện một bài kiểm tra tiêu chuẩn để đo mức độ trầm cảm. Điểm kiểm tra cao hơn cho thấy mức độ trầm cảm cao hơn. Dữ liệu thu thập được như sau:

Hỏi vị trí địa lý có ảnh hưởng đến chứng trầm cảm không? Kết luận với mức ý nghĩa 5%.

	Florida	New York	North Carolina
	3	8	10
	7	11	7
	7	9	3
	3	7	5
	8	8	11
	8	7	8
	8	8	4
	5	4	3
	5	13	7
	2	10	8
	6	6	8
	2	8	7
	6	12	3
	6	8	9
	9	6	8
	7	8	12
	5	5	6
	4	7	3
	7	7	8
	3	8	11

#### Câu hỏi 2: (1,5 điểm)

**Yêu cầu:** Khuyến khích sinh viên nộp bài làm bằng tập tin Excel.

Theo dõi độ tuổi ( $X$ , đơn vị tính: tuổi) và nhịp tim trung bình ( $Y$ , đơn vị tính: lần/phút) của một số trẻ em được chọn ngẫu nhiên ở độ tuổi từ 9-15, người ta thu thập được số liệu sau:

Độ tuổi	9	10	11	12	13	14	15
Nhịp tim trung bình	72,8	72,5	73,6	69,8	69,2	68,6	70,2

Số trẻ	30	45	32	44	45	31	35
--------	----	----	----	----	----	----	----

- a) Tìm hệ số tương quan mẫu giữa độ tuổi và nhịp tim trung bình. Nêu mối liên hệ tuyến tính giữa độ tuổi và nhịp tim trung bình.
- b) Tìm phương trình hồi quy tuyến tính ước lượng dùng để dự đoán giá trị của nhịp tim trung bình theo độ tuổi. Hãy kiểm định ý nghĩa của mô hình hồi quy vừa tìm được với mức ý nghĩa 5%.

### Câu hỏi 3: (2 điểm)

Một xét nghiệm dùng chẩn đoán bệnh. Xác suất dương tính của xét nghiệm bằng 0,2. Giá trị tiên đoán dương của xét nghiệm bằng 0,2. Biết tỉ lệ bị bệnh trong nhóm xét nghiệm âm tính bằng 0,0125. Tìm độ nhạy, độ chuyên của xét nghiệm.

----- HẾT -----

## ĐÁP ÁN PHẦN TỰ LUẬN VÀ THANG ĐIỂM

### Đáp án câu hỏi 1 (1,5 điểm)

\* Kết quả phân tích phương sai từ phần mềm Excel:

(0,5 điểm)

Anova: Single Factor  
SUMMARY

Groups	Count	Sum	Average	Variance
Column 1	20	111	5.55	4.5763
Column 2	20	160	8	4.8421
Column 3	20	141	7.05	8.05

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	61.0333	2	30.5167	5.2409	0.0081	3.1588
Within Groups	331.9	57	5.8228			
Total	392.9333	59				

\* Kiểm định sự khác biệt về ảnh hưởng của vị trí địa lý đối với chứng trầm cảm:

Giả thuyết  $H_0$ : Vị trí địa lý không ảnh hưởng đến chứng trầm cảm

Đối thuyết  $H_1$ : Vị trí địa lý ảnh hưởng đến chứng trầm cảm

(0,25 điểm)

Giá trị của tiêu chuẩn kiểm định:  $F - stat = 5,2409$

Phân vị Fisher:  $F - crit = 3,1588$

(0,25 điểm)

Do  $F - stat > F - crit$  nên ta bác bỏ giả thuyết  $H_0$ , chấp nhận đối thuyết  $H_1$ .

(0,25 điểm)

Kết luận: Vị trí địa lý có ảnh hưởng đến chứng trầm cảm.

(0,25 điểm)

**Ghi chú:** Có thể dùng phương pháp P-value: P-value = 0,0081 < 5%, ...

### Đáp án câu hỏi 2 (1,5 điểm)

a) Hệ số tương quan mẫu giữa  $X$  và  $Y$ :  $r_{XY} = -0,7792$  **(0,25 điểm)**

Do  $r_{XY} < 0$ ,  $|r_{XY}| > 0,7$  nên ta có thể dự đoán: giữa  $X$  và  $Y$  có mối tương quan tuyến tính nghịch, chặt chẽ. **(0,25 điểm)**

b) Phương trình hồi quy tuyến tính của  $Y$  theo  $X$ :  $\bar{y}_x = 79,5009 - 0,7184x$  **(0,25 điểm)**

Mô hình hồi quy:  $Y = A + BX + \varepsilon$

Giả thuyết  $H_0$ : Mô hình hồi quy không có ý nghĩa thống kê **(0,25 điểm)**

Đối thuyết  $H_1$ : Mô hình hồi quy có ý nghĩa thống kê

\* **Kết quả phân tích hồi quy tuyến tính từ phần mềm Excel:** **(0,25 điểm)**

SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R		0.7792				
R Square		0.6072				
Adjusted R Square		0.6057				
Standard Error		1.1138				
Observations		262				
<i>ANOVA</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	498.5880	498.5880	401.9440	1.1E-54	
Residual	260	322.5148	1.2404			
Total	261	821.1028				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	79.5009	0.4355	182.5582	1.7E-276	78.6434	80.3584
X Variable 1	-0.7184	0.0358	-20.0485	1.1E-54	-0.7890	-0.6479

Do  $P\text{-value} = 10^{-54} \ll 5\%$  nên ta bác bỏ giả thuyết  $H_0$  và chấp nhận đối thuyết  $H_1$ .

Nghĩa là mô hình hồi quy vừa tìm được có ý nghĩa thống kê. **(0,25 điểm)**

### **Đáp án câu hỏi 3 (2 điểm)**

Xác suất dương tính của xét nghiệm:  $P(T^+) = 0,2$  **(0,25 điểm)**

Giá trị tiên đoán dương của xét nghiệm:  $P(B^+ | T^+) = 0,2$

Tỉ lệ bị bệnh trong nhóm xét nghiệm âm tính:  $P(B^+ | T^-) = 0,0125$  **(0,25 điểm)**

Tỉ lệ bệnh lưu hành trong dân số:  $P(B^+) = P(T^+) \cdot P(B^+ | T^+) + P(T^-) \cdot P(B^+ | T^-)$

$P(B^+) = 0,2 \times 0,2 + (1 - 0,2) \times 0,0125 = 0,05$  **(0,5 điểm)**

Độ nhạy của xét nghiệm:



$$P(T^+ | B^+) = \frac{P(T^+) \cdot P(B^+ | T^+)}{P(B^+)} = \frac{0,2 \times 0,2}{0,05} = \frac{4}{5} = 0,8$$

**(0,5 điểm)**

Độ chuyên của xét nghiệm:

$$P(T^- | B^-) = \frac{P(T^-) \cdot P(B^- | T^-)}{P(B^-)} = \frac{(1-0,2) \times (1-0,0125)}{(1-0,05)} = \frac{79}{95} \approx 0,8316$$

**(0,5 điểm)****Ghi chú:** Sinh viên có thể dùng phương pháp bảng  $2 \times 2$  hoặc tỉ số khả năng.----- **hết** -----

TP. Hồ Chí Minh, ngày 19 tháng 04 năm 2024

**Người duyệt đề**

**Đinh Tiến Liêm****Giảng viên ra đề**

**Đỗ Lưu Công Minh**