

Giảng viên ra đề:	Ngày	Người phê duyệt:	Ngày
	19/05/2022		19/05/2022

 TRƯỜNG ĐẠI HỌC BÁCH KHOA - VNUHCM Khoa Khoa Học Ứng Dụng	Kiểm tra cuối kỳ		Năm học	2021-2022	Học kỳ	2
			Ngày thi	21/05/2022 (ca 2)		
	Môn học	Xác suất thống kê				
	Mã môn học	MT2013				
	Thời gian	100 phút	Mã đề			2127

### Ghi chú

- Sinh viên được sử dụng máy tính bỏ túi, bảng tra và tài liệu giấy là 2 tờ A4 được in hoặc photo (ghi rõ họ tên sinh viên)
- Sinh viên không làm tròn kết quả trung gian, làm tròn đáp án cuối cùng đến 4 chữ số thập phân.
- Sinh viên phải điền tên và MSSV vào đề thi. Đề thi gồm 17 câu hỏi trên 4 trang A4.
- Với các câu hỏi trắc nghiệm, sinh viên phải điền thông tin và tô phiếu trả lời trắc nghiệm theo đúng hướng dẫn, nếu không có thể bị mất điểm. Sinh viên phải nộp lại phiếu trả lời trắc nghiệm sau 60 phút tính từ lúc bắt đầu làm bài.
- Với các câu hỏi tự luận, sinh viên phải trình bày đầy đủ các bước, nếu không có thể bị trừ điểm.

Họ và tên sinh viên: .....

MSSV: ..... Giám thị 1: .....

Nhóm: ..... Giám thị 2: .....

### Phần I: Trắc nghiệm (6 điểm, 60 phút)

#### Từ câu 1 đến câu 5, sử dụng dữ kiện sau:

Giả sử có hai loại lỗi in được phát hiện trong một cuốn sách: lỗi chính tả và lỗi mực in (thiếu nét, mờ, nhoè). Giả thiết rằng số lỗi chính tả trên mỗi trang là một biến ngẫu nhiên ( $X$ ) có phân phối Poisson và 70% các trang sách không có lỗi chính tả nào. Số lỗi mực in ở mỗi trang sách cũng được giả thiết là một biến ngẫu nhiên Poisson ( $Y$ ) với trung bình 0.25 lỗi/trang.

1. Tìm số lỗi chính tả trung bình trên một trang sách.

- (A)  $E(X) = 0.2163$     (B)  $E(X) = 0.3567$     (C) Các câu còn lại đều sai.    (D)  $E(X) = 0.6658$   
 (E)  $E(X) = 0.2964$

2. Tìm tỉ lệ trang sách có từ 3 lỗi chính tả trở lên.

- (A) 0.0076    (B) 0.0058    (C) Các câu còn lại đều sai.    (D) 0.0074    (E) 0.0044

3. Tính xác suất để có đúng tổng cộng 2 lỗi mực in trên 4 trang sách.

- (A) 0.1839    (B) 0.1696    (C) 0.1637    (D) Các câu còn lại đều sai.    (E) 0.1002

4. Tính phương sai cho tổng số lỗi in (cả lỗi chính tả và lỗi mực in) trên một trang sách.

- (A) Các câu còn lại đều sai.    (B) 0.4577    (C) 0.7117    (D) 0.6067    (E) 0.3728

5. Tính xác suất để tổng số lỗi in (cả lỗi chính tả và lỗi mực in) trên một trang sách không quá 3 lỗi/trang.

- (A) 0.0196    (B) 0.0227    (C) Các câu còn lại đều sai.    (D) 0.0261    (E) 0.0238

#### Từ câu 6 đến câu 10, sử dụng dữ kiện sau:

Trong một hệ thống các trường tiểu học ở một thành phố A người ta ghi nhận rằng có 33% số học sinh đến từ các gia đình có một con. Tuy nhiên, khi thực hiện khảo sát trong các trường chuyên, thì người ta ghi nhận rằng có 33 trong tổng số 108 học sinh được khảo sát đến từ các gia đình có một con. Dữ liệu trên có đủ cơ sở để ta kết luận rằng tỷ lệ học sinh là con một ở các trường chuyên nhỏ hơn tỷ lệ tương ứng của thành phố hay không? Sử dụng mức ý nghĩa  $\alpha = 0.01$ .

6. Chọn giả thuyết đối  $H_1$  phù hợp cho bài toán kiểm định trên.  
(A)  $p \neq 0.33$  (B)  $p > 0.33$  (C)  $p \leq 0.33$  (D)  $p < 0.33$  (E)  $p \geq 0.33$
7. Giá trị kiểm định thống kê là:  
(A)  $-1.0403$  (B)  $-2.5403$  (C)  $0.4597$  (D) Các câu còn lại đều sai. (E)  $-0.5403$
8. Miền bác bỏ là:  
(A)  $(-\infty, -2.33) \cup (2.33, +\infty)$  (B)  $(-\infty, -2.58)$  (C)  $(-\infty, -2.33)$  (D)  $(2.58, +\infty)$   
(E)  $(2.33, +\infty)$
9. Chọn kết luận phù hợp với bài toán kiểm định trên  
(A) Giả thuyết  $H_1$  là sai. (B) Chưa đủ cơ sở để bác bỏ giả thuyết  $H_0$ . (C) Đủ cơ sở để bác bỏ giả thuyết  $H_0$ . (D) Đủ cơ sở để bác bỏ giả thuyết  $H_1$ . (E) Các câu còn lại đều sai.
10. Khoảng tin cậy ước lượng với độ tin cậy 99% cho tỷ lệ học sinh là con một trong hệ thống các trường chuyên là:  
(A)  $[0.1912, 0.4199]$  (B) Các câu còn lại đều sai. (C)  $[0.0989, 0.2938]$  (D)  $[0.2023, 0.4088]$   
(E)  $[0.3247, 0.3808]$

**Từ câu 11 đến câu 15, sử dụng dữ kiện sau:**

Trong một thí nghiệm, người ta chọn ngẫu nhiên 14 quả bóng gôn nhãn hiệu thứ nhất và 7 quả bóng gôn nhãn hiệu thứ hai và sử dụng cùng một phương pháp để phát bóng và đo khoảng cách phát bóng của chúng. Dữ liệu thu được cho thấy khoảng cách phát bóng trung bình cho hai nhãn hiệu lần lượt là 269 thước Anh và 260 thước Anh. Người ta cũng tính được độ lệch chuẩn mẫu lần lượt là 12 thước Anh và 4 thước Anh. Với giả thiết các khoảng cách này tuân theo phân phối chuẩn và độ tin cậy  $\alpha = 0.01$ , ta có thể kết luận khoảng cách phát bóng trung bình của loại bóng thứ hai nhỏ hơn của loại bóng thứ nhất không?

11. Chọn giả thuyết đối  $H_1$  :  
(A)  $\mu_1 \leq \mu_2$  (B)  $\mu_1 < \mu_2$  (C)  $\mu_1 \geq \mu_2$  (D)  $\mu_1 \neq \mu_2$  (E)  $\mu_1 > \mu_2$
12. Chọn kiểm định phù hợp  
(A) Các câu còn lại đều sai. (B) F-test (sử dụng phân phối Fisher) với bậc tự do 2 và 19.  
(C) z-test (sử dụng phân phối chuẩn tắc) (D) t-test với bậc tự do 18 (E) t-test (sử dụng phân phối Student) với bậc tự do 16
13. Giá trị kiểm định thống kê  
(A) 2.0383 (B) 1.0383 (C) Các câu còn lại đều sai. (D) 2.5383 (E) 4.0383
14. Xác định miền bác bỏ  
(A)  $(2.5524, +\infty)$  (B)  $(-\infty, 2.5758)$  (C)  $(2.5758, +\infty)$  (D)  $(2.8784, +\infty)$  (E)  $(-\infty, -2.8784) \cup (2.8784, +\infty)$
15. Chọn kết luận phù hợp với bài toán kiểm định trên  
(A) Giả thuyết  $H_1$  là sai. (B) Chưa đủ cơ sở để bác bỏ giả thuyết  $H_0$ . (C) Đủ cơ sở để bác bỏ giả thuyết  $H_0$ . (D) Đủ cơ sở để bác bỏ giả thuyết  $H_1$ . (E) Các câu còn lại đều sai.

**Phần II: Tự luận (4 điểm)**

16. Biên độ ngày của nhiệt độ là chênh lệch giữa nhiệt độ cao nhất và thấp nhất trong một ngày (24 giờ). Bảng số liệu mẫu dưới đây ghi nhận biên độ ngày của nhiệt độ tại 3 thành phố trong 5 ngày lấy mẫu ngẫu nhiên.

Thành phố	Biên độ ngày của nhiệt độ				
Hà Nội	8.1	4.5	7.2	7.3	8.1
Huế	10.4	10.3	12.1	14.3	11.0
TP Hồ Chí Minh	8.8	10.1	11.1	9.1	10.0

Giả sử rằng dữ liệu trên thỏa các giả định của phương pháp Anova.

- Hãy dùng phương pháp Anova để so sánh trung bình biên độ ngày của nhiệt độ tại 3 tỉnh trên và kết luận với mức ý nghĩa 5%.
- Thực hiện so sánh bội bằng phương pháp LSD của Fisher và nhận xét kết quả với mức ý nghĩa 5%.

17. Để nghiên cứu mối quan hệ tuyến tính giữa độ tuổi và huyết áp tâm thu của các bệnh nhi, người ta đã thực hiện khảo sát trên 10 bệnh nhi có độ tuổi nhỏ hơn 14 và kết quả khảo sát giả sử được ghi nhận như bảng sau:

Tuổi ( $X$ )	8	9	14	5	4	3	6	1	13	10
Huyết áp tâm thu ( $Y$ )	98	100	109	106	91	89	94	86	107	102

Giả sử rằng dữ liệu trên thỏa các giả định của mô hình hồi quy tuyến tính đơn với  $Y$  là biến phụ thuộc và  $X$  là biến độc lập.

- Tính hệ số tương quan mẫu và đưa ra nhận xét cho hệ số tương quan mẫu này.
- Viết phương trình hồi quy tuyến tính của  $Y$  theo  $X$ .
- Tìm khoảng tin cậy 95% cho hệ số góc của đường hồi quy tuyến tính của  $Y$  theo  $X$ .

**Đáp án phần trắc nghiệm.**

Question	Key
1	B
2	B
3	A
4	D
5	<del>A</del> <i>C</i>
6	D
7	E
8	C
9	<del>A</del> <i>B</i>
10	A
11	E
12	D
13	D
14	A
15	B
(L.O.1.1, L.O.1.2, L.O.1.3, L.O.2.1, L.O.2.3)	??
(L.O.1.1, L.O.1.2, L.O.1.3, L.O.2.1, L.O.2.3)	??