


Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học	1	2023-2024	
			Ngày thi	18/12/2023		
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2311		
	Thời gian	100 phút	Cà thi	13:00		
Ghi chú: <ul style="list-style-type: none"> - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm. 						
Họ & tên SV :			CBCT 1:			
MSSV:			CBCT 2:			

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 15 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 0.67 điểm, với mỗi đáp án sai sinh viên bị trừ 0.13 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.85 và xác suất sinh viên chọn sai là 0.15. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- ☒ A Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
☐ B Chỉ có phát biểu (2) là đúng.
☐ C Chỉ có phát biểu (1) là đúng.
☐ D Chỉ có phát biểu (3) là đúng.
☐ E Chỉ có phát biểu (4) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- ☐ A 0.8375 ☒ B 0.6375 ☐ C Các đáp án còn lại đều sai. ☐ D 0.2375 ☐ E 0.4375

3. Tính xác suất sinh viên bỏ qua 4 câu hỏi.

- ☐ A 0.1252 ☐ B 0.7252 ☒ C 0.2252 ☐ D 0.4252 ☐ E Các đáp án còn lại đều sai.

4. Nếu sinh viên bỏ qua 4 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) Các đáp án còn lại đều sai. (B) 6.05 (C) 7.05 (D) 4.55 (E) 5.05

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là 585 giờ. Một mẫu ngẫu nhiên gồm 10 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.1. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 585.8, 571.1, 582.6, 560.3, 593.3, 589.4, 566.6, 551.2, 586.4, 568.7

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu = 585$, $H_1: \mu \neq 585$.
(B) $H_0: \mu \leq 585$, $H_1: \mu > 585$.
(C) $H_0: \mu > 585$, $H_1: \mu \leq 585$.
(D) $H_0: \mu \neq 585$, $H_1: \mu = 585$.
(E) $H_0: \mu \geq 585$, $H_1: \mu < 585$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối chuẩn với trung bình là 575.54 và độ lệch chuẩn là không biết.
(B) Các đáp án còn lại đều sai.
(C) Phân phối chuẩn với độ lệch chuẩn là 4.4117.
(D) Phân phối Student với 9 bậc tự do.
(E) Phân phối chuẩn với độ lệch chuẩn là không biết.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -29.9223 (B) -46.5888 (C) -2.1443 (D) Các đáp án còn lại đều sai. (E) -7.7003

8. Xác định khoảng tin cậy với độ tin cậy 90% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) Các đáp án còn lại đều sai. (B) [567.4533, 583.6267] (C) [569.4386, 581.6414] (D) [568.3048, 582.7752] (E) [579.353, 590.647]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 13.73 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 90% cho giá trị kỳ vọng μ là không quá 5 (giờ)?

- (A) 56 (B) 16 (C) 21 (D) 31 (E) 36

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 18, \sum_{i=1}^n x_i = 654, \sum_{i=1}^n y_i = 139.646, \sum_{i=1}^n x_i^2 = 27244, \\ \sum_{i=1}^n y_i^2 = 1230.8598 \text{ và } \sum_{i=1}^n x_i y_i = 5781.456.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) Các đáp án còn lại đều sai. (B) 0.5845 (C) 0.9875 (D) 0.9593 (E) 0.8034

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) giảm khoảng 0.2032×10^9 hạt.
- (B) tăng khoảng 0.374×10^9 hạt.
- ☒ (C) tăng khoảng 0.2032×10^9 hạt.
- (D) Các đáp án khác đều sai.
- (E) giảm khoảng 0.374×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) 0.1551 (B) 0.3371 (C) 0.0796 ☒ (D) 0.0081 (E) Các đáp án còn lại đều sai.

13. Tìm khoảng tin cậy với độ tin cậy 90% cho hệ số góc β_1 của mô hình hồi quy.

- ☒ (A) [-0.4498, 0.8563] (B) [0.19, 0.2165] (C) Các đáp án khác đều sai. (D) [0.1924, 0.2141]
- ☒ (E) [0.1891, 0.2174]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) 0.3951 (B) Các đáp án còn lại đều sai. ☒ (C) 0.374 (D) 0.0821 (E) 0.2174

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 2.99$ tại $x = 14$.

- (A) -0.5407 (B) -0.4154 ☒ (C) -0.2293 (D) -0.264 (E) Các đáp án còn lại đều sai.

Từ câu 16 đến câu 20 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 148 sản phẩm từ máy in loại 1 và 414 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 13 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 12 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.01$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) 0.0197 (B) 0.5196 (C) 0.0752 (D) Các đáp án còn lại đều sai. (E) 0.1863

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định z vì so sánh hai kỳ vọng.
- ☒ (B) Kiểm định t vì phương sai tổng thể chưa biết.
- (C) Kiểm định t vì không biết phân phối của tổng thể.
- ☒ (D) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
- ☒ (E) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) 4.8691 (B) 1.5361 ☒ (C) 2.9806 (D) 0.9806 (E) Các đáp án khác đều sai.

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
- ☒ (B) Chọn sử dụng máy in 2 vì giả thuyết H_0 được bác bỏ.
- (C) Chọn sử dụng máy in 1 vì giả thuyết H_0 được bác bỏ.
- (D) Chọn sử dụng máy in 1 vì giả thuyết H_0 KHÔNG được bác bỏ
- (E) Các câu còn lại đều sai.

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 99 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) Các đáp án khác đều sai. (B) [-0.0048, 0.1225] (C) [0.0013, 0.1164] (D) [0.0079, 0.1098]
- ☒ (E) [0.0078, 0.1099]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2312		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm.						
Họ & tên SV :			CBCT 1:			
MSSV:			CBCT 2:			

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 10 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 1 điểm, với mỗi đáp án sai sinh viên bị trừ 0.2 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.75 và xác suất sinh viên chọn sai là 0.25. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- ☐ (A) Chỉ có phát biểu (1) là đúng.
☐ (B) Chỉ có phát biểu (4) là đúng.
☒ (C) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
☐ (D) Chỉ có phát biểu (3) là đúng.
☐ (E) Chỉ có phát biểu (2) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- ☐ (A) 0.1625 ☒ (B) 0.5625 ☐ (C) 0.2625 ☐ (D) 0.8625 ☐ (E) Các đáp án còn lại đều sai.

3. Tính xác suất sinh viên bỏ qua 2 câu hỏi.

- ☐ (A) 0.4816 ☒ (B) 0.2816 ☐ (C) 0.7816 ☐ (D) 0.5816 ☐ (E) Các đáp án còn lại đều sai.

4. Nếu sinh viên bỏ qua 2 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) 7.6 (B) 5.6 (C) Các đáp án còn lại đều sai. (D) 3.6 (E) 4.6

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là nhỏ hơn 767.5 giờ. Một mẫu ngẫu nhiên gồm 7 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.1. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 774.5, 785.9, 768.3, 761.3, 752.6, 776.8, 752.1

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu = 767.5, H_1: \mu < 767.5$.
(B) $H_0: \mu \neq 767.5, H_1: \mu = 767.5$.
(C) $H_0: \mu > 767.5, H_1: \mu \leq 767.5$.
(D) $H_0: \mu = 767.5, H_1: \mu \neq 767.5$.
(E) $H_0: \mu \leq 767.5, H_1: \mu > 767.5$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối chuẩn với độ lệch chuẩn là 4.8111.
(B) Phân phối chuẩn với độ lệch chuẩn là không biết.
(C) Phân phối Student với 6 bậc tự do.
(D) Phân phối chuẩn với trung bình là 767.3571 và độ lệch chuẩn là không biết.
(E) Các đáp án còn lại đều sai.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -0.0297 (B) Các đáp án còn lại đều sai. (C) -50.0297 (D) -38.9187 (E) -22.2522

8. Xác định khoảng tin cậy với độ tin cậy 90% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [761.3418, 773.6582] (B) [760.4292, 774.2851] (C) Các đáp án còn lại đều sai. (D) [758.0092, 776.7051] (E) [759.467, 775.2473]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 16.15 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 90% cho giá trị kỳ vọng μ là không quá 4 (giờ)?

- (A) 24 (B) 94 (C) 84 (D) 44 (E) 14

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 18, \sum_{i=1}^n x_i = 624, \sum_{i=1}^n y_i = 106.184, \sum_{i=1}^n x_i^2 = 26302, \\ \sum_{i=1}^n y_i^2 = 758.6983 \text{ và } \sum_{i=1}^n x_i y_i = 4455.832.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.8885 (B) Các đáp án còn lại đều sai. (C) 0.8934 (D) 0.7765 (E) 0.9857

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) giảm khoảng 0.1477×10^9 hạt.
- (B) tăng khoảng 0.1477×10^9 hạt.
- (C) Các đáp án khác đều sai.
- (D) tăng khoảng 0.1659×10^9 hạt.
- (E) giảm khoảng 0.1659×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.
 (A) 0.2921 (B) 0.1193 (C) 0.3984 (D) 0.0071 (E) Các đáp án còn lại đều sai.
13. Tìm khoảng tin cậy với độ tin cậy 95% cho hệ số góc β_1 của mô hình hồi quy.
 (A) Các đáp án khác đều sai. (B) $[-0.1471, 0.4789]$ (C) $[0.152, 0.1798]$ (D) $[0.1509, 0.181]$
 (E) $[0.1535, 0.1783]$
14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.
 (A) 0.2104 (B) 0.0678 (C) 0.1551 (D) Các đáp án còn lại đều sai. (E) 0.1477
15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 9.03$ tại $x = 55$.
 (A) Các đáp án còn lại đều sai. (B) -0.2407 (C) -0.2426 (D) 0.17 (E) 0.0097

Từ câu 16 đến câu 20(L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 206 sản phẩm từ máy in loại 1 và 256 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 14 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 13 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.05$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.
 (A) 0.1886 (B) Các đáp án còn lại đều sai. (C) 0.2997 (D) 0.0775 (E) 0.022
17. Xác định phương pháp kiểm định phù hợp.
 (A) Kiểm định z vì so sánh hai kỳ vọng.
 (B) Kiểm định t vì phương sai tổng thể chưa biết.
 (C) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
 (D) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
 (E) Kiểm định t vì không biết phân phối của tổng thể.
18. Tính giá trị thống kê của bài toán kiểm định.
 (A) 1.0045 (B) 2.671 (C) 1.56 (D) Các đáp án khác đều sai. (E) 0.7825
19. Chọn khẳng định đúng nhất cho bài toán kiểm định.
 (A) Máy in 1 là tốt hơn máy in 2.
 (B) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
 (C) Chọn sử dụng máy in 1 vì $\hat{p}_1 < \hat{p}_2$.
 (D) Các câu còn lại đều sai.
 (E) Máy in 2 là tốt hơn máy in 1.
20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 95 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.
 (A) Các đáp án khác đều sai. (B) $[-0.0265, 0.0608]$ (C) $[-0.026, 0.0603]$ (D) $[-0.0193, 0.0537]$
 (E) $[-0.0259, 0.0602]$

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2313		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm.						
Họ & tên SV :				CBCT 1:		
MSSV:				CBCT 2:		

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 10 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 1 điểm, với mỗi đáp án sai sinh viên bị trừ 0.2 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.8 và xác suất sinh viên chọn sai là 0.2. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
- (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (A) Chỉ có phát biểu (2) là đúng.
 - (B) Chỉ có phát biểu (1) là đúng.
 - (C) Chỉ có phát biểu (3) là đúng.
 - (D) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
 - (E) Chỉ có phát biểu (4) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- (A) 0.9 (B) 0.6 (C) 0.3 (D) Các đáp án còn lại đều sai. (E) 0.8

3. Tính xác suất sinh viên bỏ qua 4 câu hỏi.

- (A) Các đáp án còn lại đều sai. (B) 0.646 (C) 0.146 (D) 0.046 (E) 0.446

4. Nếu sinh viên bỏ qua 4 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) Các đáp án còn lại đều sai. (B) 4.06 (C) 2.56 (D) 4.56 (E) 3.06

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là lớn hơn 958 giờ. Một mẫu ngẫu nhiên gồm 9 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.05. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 905.7, 1072.9, 989.9, 886.4, 959.5, 1071.8, 1088, 944.1, 1050.9

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu \neq 958$, $H_1: \mu = 958$.
(B) $H_0: \mu > 958$, $H_1: \mu \leq 958$.
(C) $H_0: \mu \geq 958$, $H_1: \mu < 958$.
(D) $H_0: \mu = 958$, $H_1: \mu \neq 958$.
☒ (E) $H_0: \mu = 958$, $H_1: \mu > 958$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối chuẩn với độ lệch chuẩn là không biết.
☒ (B) Phân phối Student với 8 bậc tự do.
(C) Các đáp án còn lại đều sai.
(D) Phân phối chuẩn với độ lệch chuẩn là 25.6481.
(E) Phân phối chuẩn với trung bình là 996.5778 và độ lệch chuẩn là không biết.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) 7.0591 (B) 45.9476 ☒ (C) 1.5041 (D) 34.8366 (E) Các đáp án còn lại đều sai.

8. Xác định khoảng tin cậy với độ tin cậy 95% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- ☒ (A) [937.4332 , 1055.7224] (B) [915.937 , 1000.063] (C) [946.3074 , 1046.8481] (D) Các đáp án còn đều sai. (E) [948.8722 , 1044.2833]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 96.39 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 95% cho giá trị kỳ vọng μ là không quá 3 (giờ)?

- ☒ (A) 3966 (B) 3986 (C) 3996 (D) 3946 (E) 3936

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 19, \sum_{i=1}^n x_i = 669, \sum_{i=1}^n y_i = 213.852, \sum_{i=1}^n x_i^2 = 26361, \\ \sum_{i=1}^n y_i^2 = 2670.8027 \text{ và } \sum_{i=1}^n x_i y_i = 8384.988.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.9941 (B) 0.8289 (C) 0.9189 (D) 0.896 (E) Các đáp án còn lại đều sai.

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) Các đáp án khác đều sai.
- (B) giảm khoảng 0.3048×10^9 hạt.
- (C) tăng khoảng 0.3048×10^9 hạt.
- (D) giảm khoảng 0.5215×10^9 hạt.
- (E) tăng khoảng 0.5215×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) 0.0838 (B) Các đáp án còn lại đều sai. (C) 0.0081 (D) 0.3037 (E) 0.0191

13. Tìm khoảng tin cậy với độ tin cậy 95% cho hệ số góc β_1 của mô hình hồi quy.

- (A) [0.2908, 0.3189] (B) [-0.7954, 1.4051] (C) Các đáp án khác đều sai. (D) [0.289, 0.3207]
- (E) [0.2878, 0.3219]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) 0.5215 (B) 0.516 (C) 0.6163 (D) 0.7634 (E) Các đáp án còn lại đều sai.

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 6.05$ tại $x = 19$.

- (A) 0.0134 (B) -0.2636 (C) -0.6161 (D) -0.2945 (E) Các đáp án còn lại đều sai.

Từ câu 16 đến câu 20 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 374 sản phẩm từ máy in loại 1 và 202 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 7 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 18 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 1 hay không, với mức ý nghĩa $\alpha = 0.05$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) Các đáp án còn lại đều sai. (B) 0.1844 (C) 0.0733 (D) 0.4066 (E) 0.0178

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
- (B) Kiểm định t vì không biết phân phối của tổng thể.
- (C) Kiểm định z vì so sánh hai kỳ vọng.
- (D) Kiểm định t vì phương sai tổng thể chưa biết.
- (E) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) -3.9564 (B) -2.0679 (C) -4.8454 (D) Các đáp án khác đều sai. (E) -5.9564

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Các câu còn lại đều sai.
- (B) Chọn sử dụng máy in 1 vì giả thuyết H_0 được bác bỏ.
- (C) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
- (D) Chọn sử dụng máy in 2 vì giả thuyết H_0 được bác bỏ.
- (E) Chọn sử dụng máy in 1 vì $\hat{p}_1 < \hat{p}_2$

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 95 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) Các đáp án khác đều sai. (B) [-0.1053 , -0.0354] (C) [-0.1053 , -0.0355] (D) [-0.112 , -0.0288]
- (E) [-0.1052 , -0.0356]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2314		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: <ul style="list-style-type: none"> - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm. 						
Họ & tên SV :				CBCT 1:		
MSSV:				CBCT 2:		

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 20 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 0.5 điểm, với mỗi đáp án sai sinh viên bị trừ 0.1 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.9 và xác suất sinh viên chọn sai là 0.1. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
- (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (A) Chỉ có phát biểu (3) là đúng.
 - (B) Chỉ có phát biểu (1) là đúng.
 - (C) Chỉ có phát biểu (2) là đúng.
 - (D) Chỉ có phát biểu (4) là đúng.
 - (E) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- (A) 0.775 (B) 0.975 (C) Các đáp án còn lại đều sai. (D) 0.375 (E) 0.675

3. Tính xác suất sinh viên bỏ qua 5 câu hỏi.

- (A) 0.1023 (B) 0.0023 (C) 0.2023 (D) 0.4023 (E) Các đáp án còn lại đều sai.

4. Nếu sinh viên bỏ qua 5 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) 5.1 (B) 4.6 (C) Các đáp án còn lại đều sai. (D) 6.6 (E) 5.6

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là nhỏ hơn 854 giờ. Một mẫu ngẫu nhiên gồm 9 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.01. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 816.4, 841.3, 859.6, 879.5, 878.5, 814.7, 857.5, 845, 835.6

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu = 854$, $H_1: \mu \neq 854$.
(B) $H_0: \mu \leq 854$, $H_1: \mu > 854$.
(C) $H_0: \mu > 854$, $H_1: \mu \leq 854$.
(D) $H_0: \mu \neq 854$, $H_1: \mu = 854$.
(E) $H_0: \mu = 854$, $H_1: \mu < 854$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối chuẩn với độ lệch chuẩn là 7.8711.
(B) Phân phối chuẩn với trung bình là 847.5667 và độ lệch chuẩn là không biết.
(C) Phân phối chuẩn với độ lệch chuẩn là không biết.
(D) Các đáp án còn lại đều sai.
(E) Phân phối Student với 8 bậc tự do.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -28.5953 (B) -23.0398 (C) -39.7063 (D) -0.8173 (E) Các đáp án còn lại đều sai.

8. Xác định khoảng tin cậy với độ tin cậy 99% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [821.1591 , 873.9743] (B) [827.2592 , 867.8741] (C) [835.6603 , 872.3397] (D) Các đáp án còn đều sai. (E) [824.764 , 870.3693]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 31.22 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 99% cho giá trị kỳ vọng μ là không quá 8 (giờ)?

- (A) 52 (B) 82 (C) 107 (D) 122 (E) 102

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 20, \sum_{i=1}^n x_i = 716, \sum_{i=1}^n y_i = 155.512, \sum_{i=1}^n x_i^2 = 29766, \\ \sum_{i=1}^n y_i^2 = 1390.1835 \text{ và } \sum_{i=1}^n x_i y_i = 6430.862.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.7073 (B) 0.9036 (C) 0.9984 (D) Các đáp án còn lại đều sai. (E) 0.9149

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) Các đáp án khác đều sai.
- (B) giảm khoảng 0.2089×10^9 hạt.
- (C) tăng khoảng 0.2089×10^9 hạt.
- (D) tăng khoảng 0.2961×10^9 hạt.
- (E) giảm khoảng 0.2961×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) 0.1352 (B) Các đáp án còn lại đều sai. (C) 0.0722 (D) 0.0028 (E) 0.1139

13. Tìm khoảng tin cậy với độ tin cậy 99% cho hệ số góc β_1 của mô hình hồi quy.

- (A) [0.201,0.2169] (B) Các đáp án khác đều sai. (C) [0.2019,0.216] (D) [0.2018,0.2161]
- (E) [-0.6432,1.0611]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) 0.3548 (B) 0.2961 (C) 0.2156 (D) Các đáp án còn lại đều sai. (E) 0.4636

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 7.12$ tại $x = 32$.

- (A) Các đáp án còn lại đều sai. (B) 0.4136 (C) 0.1383 (D) -0.2243 (E) 0.504

Từ câu 16 đến câu 20(L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 148 sản phẩm từ máy in loại 1 và 244 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 10 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 22 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 1 hay không, với mức ý nghĩa $\alpha = 0.05$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) Các đáp án còn lại đều sai. (B) 0.5284 (C) 0.1951 (D) 0.3062 (E) 0.0285

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định t vì phương sai tổng thể chưa biết.
- (B) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
- (C) Kiểm định z vì so sánh hai kỳ vọng.
- (D) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
- (E) Kiểm định t vì không biết phân phối của tổng thể.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) -2.7921 (B) -0.5701 (C) Các đáp án khác đều sai. (D) -2.2366 (E) -0.7921

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
- (B) Máy in 1 là tốt hơn máy in 2.
- (C) Máy in 2 là tốt hơn máy in 1.
- (D) Các câu còn lại đều sai.
- (E) Chọn sử dụng máy in 1 vì $\hat{p}_1 < \hat{p}_2$.

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 95 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) [-0.0787 , 0.0335] (B) Các đáp án khác đều sai. (C) [-0.0679 , 0.0227] (D) [-0.0767 , 0.0315]
- (E) [-0.0785 , 0.0333]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TP HCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TP HCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2315		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm.						
Họ & tên SV :			CBCT 1:			
MSSV:			CBCT 2:			

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 20 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 0.5 điểm, với mỗi đáp án sai sinh viên bị trừ 0.1 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.75 và xác suất sinh viên chọn sai là 0.25. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (A) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
 (B) Chỉ có phát biểu (3) là đúng.
 (C) Chỉ có phát biểu (4) là đúng.
 (D) Chỉ có phát biểu (1) là đúng.
 (E) Chỉ có phát biểu (2) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- (A) 0.3625 (B) 0.2625 (C) 0.5625 (D) Các đáp án còn lại đều sai. (E) 0.1625

3. Tính xác suất sinh viên bỏ qua 5 câu hỏi.

- (A) 0.6023 (B) 0.7023 (C) Các đáp án còn lại đều sai. (D) 0.4023 (E) 0.2023

4. Nếu sinh viên bỏ qua 5 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) Các đáp án còn lại đều sai. (B) 3.25 (C) 3.75 (D) 5.25 (E) 4.25

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là lớn hơn 847.6 giờ. Một mẫu ngẫu nhiên gồm 7 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.01. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 784.5, 847.6, 899.7, 895.2, 828.3, 863.2, 904

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu = 847.6$, $H_1: \mu \neq 847.6$.
(B) $H_0: \mu > 847.6$, $H_1: \mu \leq 847.6$.
(C) $H_0: \mu = 847.6$, $H_1: \mu > 847.6$.
(D) $H_0: \mu \neq 847.6$, $H_1: \mu = 847.6$.
(E) $H_0: \mu \geq 847.6$, $H_1: \mu < 847.6$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối chuẩn với độ lệch chuẩn là không biết.
(B) Phân phối chuẩn với trung bình là 860.3571 và độ lệch chuẩn là không biết.
(C) Phân phối chuẩn với độ lệch chuẩn là 16.6329.
(D) Phân phối Student với 6 bậc tự do.
(E) Các đáp án còn lại đều sai.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) 0.767 (B) Các đáp án còn lại đều sai. (C) 6.322 (D) 39.655 (E) 28.544

8. Xác định khoảng tin cậy với độ tin cậy 99% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [808.0801 , 912.6342] (B) Các đáp án còn lại đều sai. (C) [798.6992 , 922.0151] (D) [808.8454 , 886.3546] (E) [817.4444 , 903.2699]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 50.32 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 99% cho giá trị kỳ vọng μ là không quá 4 (giờ)?

- (A) 1054 (B) 1059 (C) 1034 (D) 1079 (E) 1099

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 20, \sum_{i=1}^n x_i = 763, \sum_{i=1}^n y_i = 114.65, \sum_{i=1}^n x_i^2 = 32445, \\ \sum_{i=1}^n y_i^2 = 752.5325 \text{ và } \sum_{i=1}^n x_i y_i = 4922.55.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.973 (B) 0.7855 (C) Các đáp án còn lại đều sai. (D) 0.5886 (E) 0.6187

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) Các đáp án khác đều sai.
- (B) giảm khoảng 0.1644×10^9 hạt.
- (C) tăng khoảng 0.1644×10^9 hạt.
- (D) giảm khoảng 0.5408×10^9 hạt.
- (E) tăng khoảng 0.5408×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) Các đáp án còn lại đều sai. (B) 0.0561 (C) 0.0092 (D) 0.4154 (E) 0.1839

13. Tìm khoảng tin cậy với độ tin cậy 95% cho hệ số góc β_1 của mô hình hồi quy.

- (A) Các đáp án khác đều sai. (B) [1.3005,-0.9717] (C) [0.1464,0.1825] (D) [0.1451,0.1838]
- (E) [0.1485,0.1804]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) -0.5016 (B) -0.5901 (C) Các đáp án còn lại đều sai. (D) -0.4875 (E) -0.5408

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 9.9$ tại $x = 60$.

- (A) 0.7542 (B) 0.5746 (C) Các đáp án còn lại đều sai. (D) 0.2623 (E) 0.9557

Từ câu 16 đến câu 20(L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 390 sản phẩm từ máy in loại 1 và 324 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 39 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 3 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.01$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) 0.5176 (B) 0.2954 (C) Các đáp án còn lại đều sai. (D) 0.0732 (E) 0.0177

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định t vì không biết phân phối của tổng thể.
- (B) Kiểm định z vì so sánh hai kỳ vọng.
- (C) Kiểm định t vì phương sai tổng thể chưa biết.
- (D) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
- (E) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) 6.4633 (B) 4.2413 (C) 5.1303 (D) 5.3523 (E) Các đáp án khác đều sai.

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Chọn sử dụng máy in 1 vì giả thuyết H_0 được bác bỏ.
- (B) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
- (C) Chọn sử dụng máy in 1 vì giả thuyết H_0 KHÔNG được bác bỏ
- (D) Các câu còn lại đều sai.
- (E) Chọn sử dụng máy in 2 vì giả thuyết H_0 được bác bỏ.

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 99 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) [0.0451 , 0.1364] (B) [0.0532 , 0.1282] (C) [0.0492 , 0.1323] (D) Các đáp án khác đều sai.
- (E) [0.0451 , 0.1364]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2316		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: <ul style="list-style-type: none"> - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm. 						
Họ & tên SV :				CBCT 1:		
MSSV:				CBCT 2:		

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 15 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 0.67 điểm, với mỗi đáp án sai sinh viên bị trừ 0.13 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.7 và xác suất sinh viên chọn sai là 0.3. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- ☐ (A) Chỉ có phát biểu (1) là đúng.
☐ (B) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
☐ (C) Chỉ có phát biểu (4) là đúng.
☐ (D) Chỉ có phát biểu (2) là đúng.
☐ (E) Chỉ có phát biểu (3) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- ☐ (A) 0.625 ☐ (B) 0.325 ☐ (C) 0.525 ☐ (D) Các đáp án còn lại đều sai. ☐ (E) 0.425

3. Tính xác suất sinh viên bỏ qua 3 câu hỏi.

- ☐ (A) 0.2252 ☐ (B) 0.1252 ☐ (C) Các đáp án còn lại đều sai. ☐ (D) 0.0252 ☐ (E) 0.3252

4. Nếu sinh viên bỏ qua 3 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) 5.16 (B) 6.16 (C) Các đáp án còn lại đều sai. (D) 3.66 (E) 5.66

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là 524.9 giờ. Một mẫu ngẫu nhiên gồm 8 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.05. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 613.7, 576.4, 501.7, 441.6, 625.9, 478, 509.4, 359.5

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu \neq 524.9$, $H_1: \mu = 524.9$.
(B) $H_0: \mu > 524.9$, $H_1: \mu \leq 524.9$.
(C) $H_0: \mu \leq 524.9$, $H_1: \mu > 524.9$.
(D) $H_0: \mu \geq 524.9$, $H_1: \mu < 524.9$.
(E) $H_0: \mu = 524.9$, $H_1: \mu \neq 524.9$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Các đáp án còn lại đều sai.
(B) Phân phối chuẩn với độ lệch chuẩn là không biết.
(C) Phân phối Student với 7 bậc tự do.
(D) Phân phối chuẩn với trung bình là 513.275 và độ lệch chuẩn là không biết.
(E) Phân phối chuẩn với độ lệch chuẩn là 31.8848.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -5.9206 (B) -0.3646 (C) -44.8091 (D) -11.4761 (E) Các đáp án còn lại đều sai.

8. Xác định khoảng tin cậy với độ tin cậy 95% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [437.8675 , 588.6825] (B) [450.7809 , 575.7691] (C) [452.8534 , 573.6966] (D) [472.609 , 577.191] (E) Các đáp án còn lại đều sai.

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 88.27 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 95% cho giá trị kỳ vọng μ là không quá 8 (giờ)?

- (A) 483 (B) 518 (C) 438 (D) 468 (E) 428

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 20, \sum_{i=1}^n x_i = 695, \sum_{i=1}^n y_i = 166.63, \sum_{i=1}^n x_i^2 = 28351, \\ \sum_{i=1}^n y_i^2 = 1616.8726 \text{ và } \sum_{i=1}^n x_i y_i = 6768.034.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.9978 (B) 0.7483 (C) 0.7275 (D) 0.6942 (E) Các đáp án còn lại đều sai.

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) giảm khoảng 0.2422×10^9 hạt.
- (B) tăng khoảng 0.2422×10^9 hạt.
- (C) giảm khoảng 0.2328×10^9 hạt.
- (D) Các đáp án khác đều sai.
- (E) tăng khoảng 0.2328×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) 0.0037 (B) 0.3005 (C) 0.0432 (D) Các đáp án còn lại đều sai. (E) 0.4439

13. Tìm khoảng tin cậy với độ tin cậy 99% cho hệ số góc β_1 của mô hình hồi quy.

- (A) [0.2222, 0.2433] (B) Các đáp án khác đều sai. (C) [0.2234, 0.2421] (D) [-0.4644, 0.9299]
- (E) [0.2233, 0.2422]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) 0.2602 (B) Các đáp án còn lại đều sai. (C) -0.0669 (D) 0.6715 (E) 0.2422

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 10.6$ tại $x = 44$.

- (A) -0.1504 (B) 0.2931 (C) 0.1152 (D) Các đáp án còn lại đều sai. (E) -0.3032

Từ câu 16 đến câu 20(L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 154 sản phẩm từ máy in loại 1 và 222 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 3 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 2 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.05$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) 0.5119 (B) Các đáp án còn lại đều sai. (C) 0.0675 (D) 0.1786 (E) 0.012

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
- (B) Kiểm định t vì không biết phân phối của tổng thể.
- (C) Kiểm định t vì phương sai tổng thể chưa biết.
- (D) Kiểm định z vì so sánh hai kỳ vọng.
- (E) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) 2.7602 (B) 2.2047 (C) 0.8717 (D) Các đáp án khác đều sai. (E) 1.6492

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Chọn sử dụng máy in 1 vì $\hat{p}_1 < \hat{p}_2$.
- (B) Các câu còn lại đều sai.
- (C) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
- (D) Máy in 1 là tốt hơn máy in 2.
- (E) Máy in 2 là tốt hơn máy in 1.

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 95 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) Các đáp án khác đều sai. (B) [-0.0131 , 0.0341] (C) [-0.0146 , 0.0356] (D) [-0.0105 , 0.0315] (E) [-0.0131 , 0.034]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2317		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: <ul style="list-style-type: none"> - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm. 						
Họ & tên SV :				CBCT 1:		
MSSV:				CBCT 2:		

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 10 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 1 điểm, với mỗi đáp án sai sinh viên bị trừ 0.2 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.9 và xác suất sinh viên chọn sai là 0.1. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- ☐ (A) Chỉ có phát biểu (2) là đúng.
☐ (B) Chỉ có phát biểu (3) là đúng.
☐ (C) Chỉ có phát biểu (1) là đúng.
☐ (D) Chỉ có phát biểu (4) là đúng.
☐ (E) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- ☐ (A) 0.675 ☐ (B) Các đáp án còn lại đều sai. ☐ (C) 0.875 ☐ (D) 0.275 ☐ (E) 0.975

3. Tính xác suất sinh viên bỏ qua 2 câu hỏi.

- ☐ (A) Các đáp án còn lại đều sai. ☐ (B) 0.1816 ☐ (C) 0.6816 ☐ (D) 0.2816 ☐ (E) 0.4816

4. Nếu sinh viên bỏ qua 2 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) 8.04 (B) 6.04 (C) Các đáp án còn lại đều sai. (D) 7.04 (E) 6.54

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là nhỏ hơn 502.4 giờ. Một mẫu ngẫu nhiên gồm 7 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.1. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 487.2, 544.2, 571.7, 429, 480.8, 419.8, 490.6

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu = 502.4$, $H_1: \mu < 502.4$.
(B) $H_0: \mu > 502.4$, $H_1: \mu \leq 502.4$.
(C) $H_0: \mu \leq 502.4$, $H_1: \mu > 502.4$.
(D) $H_0: \mu \neq 502.4$, $H_1: \mu = 502.4$.
(E) $H_0: \mu = 502.4$, $H_1: \mu \neq 502.4$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Phân phối Student với 6 bậc tự do.
(B) Phân phối chuẩn với độ lệch chuẩn là không biết.
(C) Các đáp án còn lại đều sai.
(D) Phân phối chuẩn với trung bình là 489.0429 và độ lệch chuẩn là không biết.
(E) Phân phối chuẩn với độ lệch chuẩn là 20.9011.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -6.1951 (B) -0.6391 (C) -11.7506 (D) Các đáp án còn lại đều sai. (E) -33.9726

8. Xác định khoảng tin cậy với độ tin cậy 90% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [448.432 , 529.6537] (B) Các đáp án còn đều sai. (C) [454.765 , 523.3207] (D) [475.6466 , 529.1534] (E) [458.9452 , 519.1405]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 50.28 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 90% cho giá trị kỳ vọng μ là không quá 5 (giờ)?

- (A) 322 (B) 272 (C) 247 (D) 282 (E) 267

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 16, \sum_{i=1}^n x_i = 508, \sum_{i=1}^n y_i = 241.78, \sum_{i=1}^n x_i^2 = 18402, \\ \sum_{i=1}^n y_i^2 = 4139.6372 \text{ và } \sum_{i=1}^n x_i y_i = 8725.12.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.886 (B) Các đáp án còn lại đều sai. (C) 0.7593 (D) 0.9569 (E) 0.9976

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) giảm khoảng 0.4613×10^9 hạt.
 (B) Các đáp án khác đều sai.
 (C) giảm khoảng 0.464×10^9 hạt.
 (D) tăng khoảng 0.4613×10^9 hạt.
 (E) tăng khoảng 0.464×10^9 hạt.
12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.
 (A) 0.2539 (B) 0.0085 (C) 0.2599 (D) 0.3747 (E) Các đáp án còn lại đều sai.
13. Tìm khoảng tin cậy với độ tin cậy 99% cho hệ số góc β_1 của mô hình hồi quy.
 (A) [0.4391, 0.4836] (B) [-0.9199, 1.8426] (C) [0.4361, 0.4866] (D) [0.4395, 0.4832] (E) Các đáp án khác đều sai.
14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.
 (A) Các đáp án còn lại đều sai. (B) 0.2226 (C) 0.7115 (D) 0.464 (E) 0.6116
15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 17.42$ tại $x = 37$. (A) -0.1132 (B) 0.0831 (C) 0.2528 (D) Các đáp án còn lại đều sai. (E) 0.0048
- Từ câu 16 đến câu 20** (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 304 sản phẩm từ máy in loại 1 và 470 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 24 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 9 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.05$?
16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.
 (A) 0.0149 (B) 0.2926 (C) 0.1815 (D) 0.4037 (E) Các đáp án còn lại đều sai.
17. Xác định phương pháp kiểm định phù hợp.
 (A) Kiểm định t vì không biết phân phối của tổng thể.
 (B) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
 (C) Kiểm định t vì phương sai tổng thể chưa biết.
 (D) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
 (E) Kiểm định z vì so sánh hai kỳ vọng.
18. Tính giá trị thống kê của bài toán kiểm định.
 (A) 4.7989 (B) 3.1324 (C) 4.0214 (D) 4.2434 (E) Các đáp án khác đều sai.
19. Chọn khẳng định đúng nhất cho bài toán kiểm định.
 (A) Chọn sử dụng máy in 2 vì giả thuyết H_0 được bác bỏ.
 (B) Các câu còn lại đều sai.
 (C) Chọn sử dụng máy in 1 vì giả thuyết H_0 được bác bỏ.
 (D) Chọn sử dụng máy in 1 vì giả thuyết H_0 KHÔNG được bác bỏ
 (E) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$
20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 95 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.
 (A) [0.0306 , 0.089] (B) Các đáp án khác đều sai. (C) [0.0324 , 0.0872] (D) [0.0307 , 0.0889]
 (E) [0.0271 , 0.0925]

Phần II: Tự luận (3 điểm)


21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Lecturer:	<i>Date</i>	Approved by:	<i>Date</i>
Dr. Phan Thi Huong	28/11/2023		28/11/2023

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		1	2023-2024
			Ngày thi		18/12/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2318		
	Thời gian	100 phút	Ca thi	13:00		
Ghi chú: - Sinh viên được sử dụng các tài liệu giấy. Riêng tài liệu viết tay trực tiếp cần phải có đầy đủ họ tên, MSSV trên mỗi trang tài liệu. Sinh viên được sử dụng máy tính bỏ túi không có chức năng lập trình. - Đề thi gồm 20 câu hỏi trắc nghiệm và 2 câu hỏi tự luận trên 4 trang giấy A4. - Không làm tròn kết quả trung gian. Kết quả cuối cùng được làm tròn đến 4 chữ số thập phân. Sinh viên chọn đáp án gần đúng nhất cho mỗi câu hỏi trắc nghiệm.						
Họ & tên SV :				CBCT 1:		
MSSV:				CBCT 2:		

Phần I: Trắc nghiệm (7 điểm)

Từ câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4). Một sinh viên làm một bài kiểm tra gồm 20 câu hỏi trắc nghiệm. Trong mỗi câu hỏi chỉ có duy nhất một lựa chọn là đúng và sinh viên chỉ chọn một đáp án cho một câu hỏi. Biết rằng với mỗi đáp án đúng sinh viên ghi được 0.5 điểm, với mỗi đáp án sai sinh viên bị trừ 0.1 điểm, và sinh viên không bị trừ điểm nếu bỏ qua câu hỏi (không chọn đáp án nào). Giả sử rằng với mỗi câu hỏi thì xác suất để sinh viên bỏ qua câu hỏi này là 0.25, với những câu hỏi sinh viên có chọn đáp án thì xác suất sinh viên chọn đúng là 0.75 và xác suất sinh viên chọn sai là 0.25. Kết quả của các câu hỏi được giả sử là độc lập với nhau.

1. Chọn phát biểu đúng trong các phát biểu sau:

- (1) Số câu hỏi mà sinh viên bỏ qua là một biến ngẫu nhiên có phân phối nhị thức.
 - (2) Số câu hỏi mà sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (3) Số câu trả lời đúng trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
 - (4) Số câu trả lời sai trong các câu sinh viên có làm là một biến ngẫu nhiên có phân phối nhị thức.
- (A) Chỉ có phát biểu (1) là đúng.
 (B) Chỉ có phát biểu (4) là đúng.
 (C) Chỉ có phát biểu (3) là đúng.
 (D) Tất cả các phát biểu (1), (2), (3) và (4) đều đúng.
 (E) Chỉ có phát biểu (2) là đúng.

2. Chọn ngẫu nhiên một câu hỏi trong bài kiểm tra, tính xác suất để câu hỏi này là một câu mà sinh viên ghi được điểm.

- (A) 0.6625 (B) Các đáp án còn lại đều sai. (C) 0.5625 (D) 0.3625 (E) 0.1625

3. Tính xác suất sinh viên bỏ qua 5 câu hỏi.

- (A) 0.3023 (B) 0.2023 (C) 0.1023 (D) Các đáp án còn lại đều sai. (E) 0.4023

4. Nếu sinh viên bỏ qua 5 câu hỏi, hãy tính số điểm trung bình mà sinh viên đạt được.

- (A) 5.25 (B) Các đáp án còn lại đều sai. (C) 4.25 (D) 5.75 (E) 3.25

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một hãng điện tử tuyên bố rằng các bóng đèn do họ sản xuất có tuổi thọ trung bình là nhỏ hơn 551.2 giờ. Một mẫu ngẫu nhiên gồm 9 bóng đèn được chọn và tuổi thọ được ghi nhận như bên dưới. Dựa trên mẫu ngẫu nhiên này, người ta muốn kiểm định xem tuyên bố của hãng có được chấp nhận hay không ở mức ý nghĩa 0.05. Biết rằng tuổi thọ của bóng đèn có phân phối chuẩn. Dữ liệu: 446.4, 544.8, 594.9, 591.4, 491.7, 561.3, 516.4, 631.3, 494.9

5. Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.

- (A) $H_0: \mu \leq 551.2$, $H_1: \mu > 551.2$.
(B) $H_0: \mu \neq 551.2$, $H_1: \mu = 551.2$.
(C) $H_0: \mu > 551.2$, $H_1: \mu \leq 551.2$.
(D) $H_0: \mu = 551.2$, $H_1: \mu \neq 551.2$.
(E) $H_0: \mu = 551.2$, $H_1: \mu < 551.2$.

6. Phân phối của thống kê kiểm định trong bài toán kiểm định trên là gì?

- (A) Các đáp án còn lại đều sai.
(B) Phân phối chuẩn với trung bình là 541.4556 và độ lệch chuẩn là không biết.
(C) Phân phối Student với 8 bậc tự do.
(D) Phân phối chuẩn với độ lệch chuẩn là không biết.
(E) Phân phối chuẩn với độ lệch chuẩn là 19.7674.

7. Tính giá trị kiểm định thống kê cho bài toán trên.

- (A) -28.271 (B) Các đáp án còn lại đều sai. (C) -11.6045 (D) -0.493 (E) -50.493

8. Xác định khoảng tin cậy với độ tin cậy 95% cho tuổi thọ trung bình của các bóng đèn sản xuất bởi công ty này.

- (A) [518.7814 , 583.6186] (B) [495.8718 , 587.0393] (C) Các đáp án còn đều sai. (D) [504.6881 , 578.223] (E) [502.7114 , 580.1998]

9. Biết thêm rằng độ lệch chuẩn của thọ bóng đèn là 57.41 giờ. Cần phải kiểm tra tối thiểu bao nhiêu bóng đèn để sai số ước lượng của khoảng tin cậy với độ tin cậy 95% cho giá trị kỳ vọng μ là không quá 8 (giờ)?

- (A) 198 (B) 203 (C) 223 (D) 188 (E) 173

Từ câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Người ta làm lạnh các nguyên tử Cesium bằng việc chiếu các tia sáng laser. Các nguyên tử Cesium sau khi làm lạnh có thể được ứng dụng trong chế tạo các đồng hồ nguyên tử với giá thành thấp. Mô hình hồi quy tuyến tính đơn được áp dụng để nghiên cứu mối quan hệ giữa số lượng hạt Cesium và cường độ của nguồn chiếu sáng. Sau đây là một số kết quả tính toán từ thực nghiệm, trong đó y là số hạt Cesium (đơn vị: 10^9 hạt) và x là cường độ của nguồn chiếu sáng (đơn vị: mW):

$$n = 18, \sum_{i=1}^n x_i = 645, \sum_{i=1}^n y_i = 234.645, \sum_{i=1}^n x_i^2 = 25733, \\ \sum_{i=1}^n y_i^2 = 3390.1683 \text{ và } \sum_{i=1}^n x_i y_i = 9338.613.$$

10. Tính hệ số tương quan mẫu cho dữ liệu này.

- (A) 0.7215 (B) 0.6868 (C) Các đáp án còn lại đều sai. (D) 0.9985 (E) 0.8182

11. Nếu cường độ nguồn sáng tăng 1 mW thì số lượng hạt Cesium được kỳ vọng sẽ

- (A) giảm khoảng 0.312×10^9 hạt.
- (B) giảm khoảng 0.3551×10^9 hạt.
- (C) tăng khoảng 0.3551×10^9 hạt.
- (D) Các đáp án khác đều sai.
- (E) tăng khoảng 0.312×10^9 hạt.

12. Đưa ra ước lượng cho độ lệch chuẩn (sai số chuẩn) của $\hat{\beta}_1$.

- (A) Các đáp án còn lại đều sai. (B) 0.2025 (C) 0.1028 (D) 0.0048 (E) 0.4269

13. Tìm khoảng tin cậy với độ tin cậy 99% cho hệ số góc β_1 của mô hình hồi quy.

- (A) [0.341, 0.3691] (B) [0.3427, 0.3675] (C) [-0.5561, 1.2662] (D) Các đáp án khác đều sai.
- (E) [0.3427, 0.3675]

14. Tìm ước lượng bình phương bé nhất cho hệ số chặn của đường thẳng hồi quy.

- (A) 0.4808 (B) Các đáp án còn lại đều sai. (C) 0.0475 (D) 0.312 (E) 0.6249

15. Từ đường thẳng hồi quy, hãy ước lượng thành phần sai số cho giá trị quan trắc $y = 14.54$ tại $x = 40$. (A) -0.1678 (B) 0.2361 (C) 0.3481 (D) Các đáp án còn lại đều sai. (E) 0.0246

Từ câu 16 đến câu 20(L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Hai loại máy in 3D được khảo sát chất lượng. Một loại máy in được đánh giá là tốt hơn nếu nó sản xuất ra ít sản phẩm hỏng hơn và sẽ được lựa chọn để sử dụng cho công ty. Người ta chọn ngẫu nhiên 330 sản phẩm từ máy in loại 1 và 474 sản phẩm từ máy in loại 2. Trong các sản phẩm được chọn từ máy loại 1 có 30 sản phẩm hỏng và trong các sản phẩm được chọn từ máy loại 2 có 43 sản phẩm hỏng được ghi nhận. Liệu dữ liệu trên có đủ chứng cứ để người ta lựa chọn sử dụng máy in loại 2 hay không, với mức ý nghĩa $\alpha = 0.01$?

16. Với giả thuyết H_0 là đúng, hãy ước lượng cho độ lệch chuẩn (sai số chuẩn) của chênh lệch tỷ lệ sản phẩm kém chất lượng ở hai mẫu trong bài toán kiểm định trên.

- (A) 0.4094 (B) 0.0761 (C) 0.1872 (D) 0.0206 (E) Các đáp án còn lại đều sai.

17. Xác định phương pháp kiểm định phù hợp.

- (A) Kiểm định z vì tỷ lệ mẫu có phân phối chuẩn.
- (B) Kiểm định t vì phương sai tổng thể chưa biết.
- (C) Kiểm định z vì tỷ lệ mẫu xấp xỉ phân phối chuẩn.
- (D) Kiểm định z vì so sánh hai kỳ vọng.
- (E) Kiểm định t vì không biết phân phối của tổng thể.

18. Tính giá trị thống kê của bài toán kiểm định.

- (A) 0.7868 (B) Các đáp án khác đều sai. (C) 0.0093 (D) 1.3423 (E) 0.2313

19. Chọn khẳng định đúng nhất cho bài toán kiểm định.

- (A) Các câu còn lại đều sai.
- (B) Máy in 2 là tốt hơn máy in 1.
- (C) Máy in 1 là tốt hơn máy in 2.
- (D) Chọn sử dụng máy in 1 vì $\hat{p}_1 < \hat{p}_2$.
- (E) Chọn sử dụng máy in 2 vì $\hat{p}_2 < \hat{p}_1$

20. Xây dựng khoảng tin cậy (hai phía) với độ tin cậy 99 % cho chênh lệch tỷ lệ sản phẩm kém chất lượng sản xuất bởi máy in loại 1 và máy in loại 2.

- (A) [-0.053 , 0.0533] (B) [-0.0478 , 0.0482] (C) [-0.053 , 0.0533] (D) [-0.053 , 0.0534]
- (E) Các đáp án khác đều sai.

Phần II: Tự luận (3 điểm)

21. (L.O.1.1, L.O.2.1, L.O.4) Một công nhân có nhiệm vụ sửa chữa ngay các lỗi xuất hiện ngẫu nhiên trên một dây chuyền sản xuất. Giả sử thời gian (đơn vị: phút) để sửa xong một lỗi là biến ngẫu nhiên có phân phối đều trên đoạn $[1; 5]$ và các lỗi xuất hiện là độc lập với nhau.
- (a) Tìm xác suất trong 5 lỗi ngẫu nhiên thì có 2 lỗi mà khi sửa mỗi lỗi này, người công nhân chỉ cần dưới 2 phút.
 - (b) Giả thiết một tháng người công nhân sửa 50 lỗi. Tìm xác suất tổng thời gian sửa lỗi trong tháng đó của người công nhân lớn hơn 180 phút.
22. (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Dưới đây là bảng số liệu về chỉ số bụi mịn PM2.5 trong không khí ở TPHCM được thu thập tại 6 thời điểm khác nhau trong 3 tháng đầu của mỗi năm, từ 2020 đến 2022. Giả thiết rằng các số liệu thỏa mãn mô hình Anova.

Năm	Chỉ số bụi mịn PM2.5					
2022	74	96	63	74	121	84
2021	78	92	68	71	99	86
2020	73	56	64	74	62	43

- (a) Hãy sử dụng phương pháp Anova để so sánh chỉ số bụi mịn PM2.5 ở TPHCM trong 3 tháng đầu của 3 năm, với mức ý nghĩa 5%.
- (b) Tìm khoảng tin cậy 95% cho chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022.
- (c) Tìm khoảng tin cậy 95% cho chênh lệch chỉ số bụi mịn PM2.5 trung bình trong 3 tháng đầu năm của năm 2022 với năm 2020. Chúng ta có thể so sánh được 2 giá trị trung bình này hay không?

–HẾT–

Answers Sheet

Question sheet code 2311:

1 A. 2 B. 3 C. 4 B. 5 A. 6 D. 7 C. 8 B. 9 C. 10 C. 11 C. 12 D. 13 E. 14 C. 15 C. 16 A.
17 D. 18 C. 19 B. 20 B.

Question sheet code 2312:

1 C. 2 B. 3 B. 4 B. 5 A. 6 C. 7 A. 8 D. 9 D. 10 E. 11 D. 12 D. 13 D. 14 E. 15 C. 16 E.
17 C. 18 E. 19 D. 20 B.

Question sheet code 2313:

1 D. 2 B. 3 C. 4 D. 5 E. 6 B. 7 C. 8 A. 9 A. 10 A. 11 C. 12 C. 13 E. 14 A. 15 B. 16 E.
17 A. 18 A. 19 B. 20 D.

Question sheet code 2314:

1 E. 2 E. 3 C. 4 D. 5 E. 6 E. 7 D. 8 A. 9 E. 10 C. 11 C. 12 D. 13 A. 14 B. 15 C. 16 E. 17 B.
18 E. 19 D. 20 D.

Question sheet code 2315:

1 A. 2 C. 3 E. 4 D. 5 C. 6 D. 7 A. 8 C. 9 A. 10 A. 11 C. 12 C. 13 D. 14 E. 15 B. 16 E.
17 E. 18 C. 19 E. 20 C.

Question sheet code 2316:

1 B. 2 C. 3 A. 4 A. 5 E. 6 C. 7 B. 8 A. 9 D. 10 A. 11 E. 12 A. 13 A. 14 E. 15 C. 16 E.
17 E. 18 C. 19 B. 20 C.

Question sheet code 2317:

1 E. 2 A. 3 D. 4 D. 5 A. 6 A. 7 B. 8 A. 9 B. 10 E. 11 D. 12 B. 13 C. 14 D. 15 A. 16 A.
17 D. 18 C. 19 A. 20 E.

Question sheet code 2318:

1 D. 2 C. 3 B. 4 A. 5 E. 6 C. 7 D. 8 B. 9 A. 10 D. 11 C. 12 D. 13 A. 14 D. 15 E. 16 D.
17 C. 18 C. 19 A. 20 A.