

 TRƯỜNG ĐẠI HỌC BÁCH KHOA- VNUHCM Khoa Khoa học ứng dụng	Thi Cuối Kỳ		Học kỳ/năm học		2	2022-2023
			Ngày thi		02/06/2023	
	Môn học	Xác suất thống kê				
	Mã môn	MT2013	Mã đề	2221		
	Thời gian	100 phút	Ca thi	07:00		
Ghi chú: <ul style="list-style-type: none"> - Sinh viên được sử dụng tài liệu giấy tài liệu giấy ở hình thức in ấn và photo, không được sử dụng tài liệu viết tay. 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. - Với phần trắc nghiệm sinh viên chọn đáp án gần nhất, làm bài trên phiếu trắc nghiệm và nộp lại phiếu trắc nghiệm sau khi bắt đầu làm bài 70 phút. Với phần tự luận, sinh viên làm bài trên đề thi và trình bày đầy đủ các bước tính toán. 						
Họ & tên SV :			CBCT 1:			
MSSV:			CBCT 2:			

Phần I: Trắc nghiệm (7 điểm, 70 phút)

Câu 1 đến câu 4 (L.O.1.1, L.O.2.1, L.O.4) : Giả sử rằng số lỗi trên một chi tiết máy do một máy tiện tự động làm ra là biến ngẫu nhiên có phân phối xác suất như sau:

X	0	1	2	3	4
P	0.25	0.23	0.18	0.21	c

- Tính số lỗi trung bình trên một chi tiết máy do máy tiện này làm ra.
 (A) 1.35 (B) 1.39 (C) 1.74 (D) 1.5 (E) 2.16
- Những chi tiết máy mà có nhiều hơn 2 lỗi thì được phân loại là sản phẩm không đạt. Chọn ngẫu nhiên 13 chi tiết máy trong rất nhiều chi tiết máy do máy tiện này làm ra. Tính xác suất để chọn được tối đa 3 sản phẩm không đạt?
 (A) 0.3043 (B) 0.3583 (C) 0.3263 (D) 0.3103 (E) 0.2893
- Chọn ngẫu nhiên 526 chi tiết máy do máy tiện này làm ra. Gọi X_i là số lỗi trên chi tiết máy thứ i và giả sử rằng số lỗi trên các chi tiết máy là độc lập với nhau. Tính độ lệch chuẩn cho trung bình mẫu này. (A) 0.084 (B) 0.1824 (C) 0.06 (D) 0.2487 (E) 0.0355
- Tính xác suất để số lỗi trung bình trong 526 chi tiết máy được chọn ngẫu nhiên từ máy tiện này là từ 1.7 đến 1.8 lỗi. (A) 0.7154 (B) 0.456 (C) 0.0739 (D) 0.589 (E) 0.1898

Câu 5 đến câu 9 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4) . Một ổ đỡ được sử dụng trong ngành công nghiệp ô tô có đường kính bên trong là đạt tiêu chuẩn khi bằng 3.4 (inch). Đường kính ổ đỡ được biết là có phân phối chuẩn với độ lệch chuẩn 0.52 (inch). Một mẫu ngẫu nhiên gồm 10 ổ đỡ được chọn và đường kính của các ổ đỡ được ghi nhận như bên dưới. 2.11, 2.4, 1.84, 1.85, 1.54, 2.54, 2.47, 2.48, 2.37, 1.61 . Dựa trên mẫu ngẫu nhiên trên, người ta muốn kiểm định xem các ổ đỡ có đạt tiêu chuẩn hay không ở mức ý nghĩa 0.1.

- Chọn cặp giả thuyết không và giả thuyết đối phù hợp cho bài toán trên.
 (A) $H_0: \mu = 3.4$, $H_1: \mu \neq 3.4$. (B) $H_0: \mu \geq 3.4$, $H_1: \mu < 3.4$. (C) $H_0: \mu \neq 3.4$, $H_1: \mu = 3.4$.
 (D) $H_0: \mu \leq 3.4$, $H_1: \mu > 3.4$. (E) $H_0: \mu > 3.4$, $H_1: \mu \leq 3.4$.

6. Phân phối của trung bình mẫu trong bài toán kiểm định trên là gì?
- (A) Phân phối chuẩn với trung bình là 3.4 và độ lệch chuẩn là 0.52.
 (B) Phân phối chuẩn với trung bình là 2.121 và độ lệch chuẩn là 0.1644.
 (C) Phân phối chuẩn với trung bình là 3.4 và độ lệch chuẩn là 0.0383.
 (D) Phân phối chuẩn với trung bình là 3.4 và độ lệch chuẩn là 0.1644.
 (E) Phân phối chuẩn với trung bình là 3.4 và độ lệch chuẩn là 0.3826.
7. Tính giá trị kiểm định thống kê cho bài toán trên.
- (A) -7.178 (B) -9.278 (C) -6.878 (D) -9.578 (E) -7.778
8. Xác định khoảng tin cậy với độ tin cậy 90% cho trung bình đường kính trong của các ổ đỡ này.
- (A) [3.1303 , 3.6697] (B) [1.8513 , 2.3907] (C) [1.9105 , 2.3315] (D) [2.0509 , 2.1911]
 (E) [1.8936 , 2.3484]
9. Cần phải kiểm tra tối thiểu bao nhiêu ổ đỡ để sai số ước lượng của khoảng tin cậy với độ tin cậy 90% cho giá trị kỳ vọng μ không quá 0.25 (inch). (A) 6 (B) 9 (C) 8 (D) 12 (E) 16

Câu 10 đến câu 15 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Các nhà nghiên cứu y học muốn biết liệu bốn loại thuốc khác nhau có dẫn đến giảm huyết áp trung bình khác nhau ở bệnh nhân hay không. Họ chỉ định ngẫu nhiên 5 bệnh nhân sử dụng mỗi loại thuốc trong một tháng, sau đó đo huyết áp cả trước và sau khi bệnh nhân bắt đầu sử dụng thuốc để tìm mức giảm huyết áp trung bình cho mỗi loại thuốc. Bảng dữ liệu được ghi nhận như bên dưới.

Loại 1	13.3	20.3	18.7	13.2	17
Loại 2	-0.2	-2.4	-1.5	4.2	1.7
Loại 3	7.1	5.9	-0.1	3.1	3.9
Loại 4	6	5.4	6.4	2.8	4.3

Xét mô hình ANOVA một nhân tố với mức ý nghĩa $\alpha = 0.05$.

10. Tính tổng biến thiên do các thành phần sai số ngẫu nhiên gây ra (tổng bình phương) trong dữ liệu này. (A) 839.6495 (B) 100.8879 (C) 107.888 (D) 731.7615 (E) 108.8875
11. Tính giá trị kiểm định thống kê cho bài toán ANOVA.
- (A) 36.1739 (B) 34.1738 (C) 40.1735 (D) 30.174 (E) 38.1736
12. Tìm miền bác bỏ cho bài toán ANOVA.
- (A) $[4.08, +\infty]$ (B) $[-\infty, 2.21]$ (C) $[-\infty, 3.24]$ (D) $[2.21, +\infty]$ (E) $[3.24, +\infty]$
13. Tính giá trị chênh lệch nhỏ nhất (LSD) của phương pháp so sánh bội Fisher.
- (A) 3.4817 (B) 4.4708 (C) 1.8203 (D) 1.7388 (E) 2.0283
14. Xác định khoảng tin cậy với độ tin cậy 95% cho chênh lệch trung bình giữa mức giảm huyết áp của loại thuốc 1 và mức giảm huyết áp của loại thuốc 2.
- (A) [12.9914, 19.9548] (B) [14.1024, 21.0658] (C) [9.6584, 16.6218] (D) [12.6583, 19.6217]
 (E) [11.8804, 18.8438]

15. Xác định khoảng tin cậy với độ tin cậy 95% cho mức giảm huyết áp trung bình của loại thuốc
- (A) [11.5937; 16.5175] (B) [13.0183; 19.9817] (C) [14.9267; 19.8505] (D) [13.2602; 18.184]
 - (E) [14.0381; 18.9619]

Câu 16 đến câu 20 (L.O.1.2, L.O.2.1, L.O.2.2, L.O.4). Để nghiên cứu mối quan hệ tuyến tính giữa thu nhập (y) (triệu đồng/tháng) và độ tuổi của các nhân viên trong một công ty lớn (x), người ta đã thu thập thông tin từ 18 nhân viên từ công ty này và số liệu được tóm tắt như sau.

	Age (x)	Income (y)
Tổng các giá trị quan trắc ($\sum x_i$ or $\sum y_i$)	383	890
Tổng bình phương ($\sum x_i^2$ or $\sum y_i^2$)	8243	44054.34

Và, $\sum x_i y_i = 18994.2$. Áp dụng mô hình hồi quy tuyến tính đơn cho y theo x .

16. Tính hệ số gốc β_1 và hệ số chặn β_0 .
- (A) $\beta_1 = -0.2913$; $\beta_0 = 35.9934$ (B) $\beta_1 = 0.6087$; $\beta_0 = 36.4934$ (C) $\beta_1 = 1.4087$; $\beta_0 = 36.9934$
 - (D) $\beta_1 = 1.0087$; $\beta_0 = 36.3934$ (E) $\beta_1 = 0.1087$; $\beta_0 = 35.6934$
17. Một nhân viên có tuổi là 38, tìm thu nhập trung bình của nhân viên này (triệu đồng/tháng).
- (A) 66.5 (B) 68 (C) 65 (D) 60 (E) 67.5
18. Ước lượng phương sai của các thành phần sai số ngẫu nhiên.
- (A) 0.6915 (B) 0.8815 (C) 0.3835 (D) 0.5705 (E) 0.7475
19. Xây dựng khoảng tin cậy hai phía 95% cho hệ số gốc:
- (A) [0.4929; 0.7344] (B) [0.1329; 1.1544] (C) [0.4029; 0.8144] (D) [-0.0471; 0.9044]
 - (E) [-0.0571; 0.7944]
20. Xác định hệ số xác định R^2 . (A) 0.7709 (B) 0.5409 (C) 0.7109 (D) 0.8209 (E) 0.4409

Phần II: Tự luận (3 điểm, 30 phút)

21. Điểm thi đánh giá năng lực của Đại học Quốc gia TP.HCM năm 2023 là biến ngẫu nhiên tuân theo quy luật phân phối chuẩn với điểm trung bình là 639.2 và độ lệch chuẩn là 118.6.
- Xác định mức điểm sao cho có 70% thí sinh tham gia kỳ thi có điểm lớn hơn mức điểm này.
 - Tìm xác suất để trong 200 thí sinh được chọn ngẫu nhiên, có từ 100 đến 120 thí sinh có điểm lớn hơn 650.

22. Một công ty sản xuất hàng thủ công mỹ nghệ có 2 cơ sở sản xuất cùng một mặt hàng. Khi theo dõi về quá trình hoàn thiện mỗi sản phẩm ở từng cơ sở, công ty thu được các số liệu mẫu như dưới đây. Ở cơ sở 1:

Thời gian hoàn thiện mỗi sản phẩm (giờ)	14-16	16-18	18-20	20-22	22-24	24-26
Số sản phẩm tương ứng	3	7	16	22	14	8

Ở cơ sở 2, người ta theo dõi thời gian hoàn thiện của 9 sản phẩm thì thấy tổng thời gian hoàn thiện của chúng 195 giờ và phương sai mẫu là 7 giờ. Giả thiết rằng thời gian hoàn thiện mỗi sản phẩm ở hai cơ sở là các biến ngẫu nhiên có phân phối chuẩn.

- (a) Tìm khoảng tin cậy 95% cho tỷ lệ sản phẩm được hoàn thiện dưới 20 giờ ở cơ sở 1.
 (b) Chúng ta có thể xem như cơ sở 1 hoàn thiện sản phẩm nhanh hơn so với cơ sở 2 hay không, hãy kiểm định với mức ý nghĩa 1%?.

Ghi chú:

Dùng điểm chính giữa mỗi khoảng khi tính trung bình mẫu ở cơ sở 1. Phần hỗ trợ bảng tra số:

Bậc của hàm student	Một số giá trị phân vị phải	
	$t_{0.005}$	$t_{0.01}$
79	2.6395	2.3745
78	2.6403	2.3751
77	2.6412	2.3758
10	3.1693	2.3758 2.764
9	3.2498	2.8214

TÀI LIỆU SƯU TẬP
 BỞI HCMUT-CNCP

F_MAMH	MADE	CAU	F_KEY
MT2013	2221	1	C
MT2013	2221	2	A
MT2013	2221	3	C
MT2013	2221	4	D
MT2013	2221	5	A
MT2013	2221	6	D
MT2013	2221	7	E
MT2013	2221	8	B
MT2013	2221	9	D
MT2013	2221	10	C
MT2013	2221	11	A
MT2013	2221	12	E
MT2013	2221	13	A
MT2013	2221	14	D
MT2013	2221	15	E
MT2013	2221	16	B
MT2013	2221	17	D
MT2013	2221	18	B
MT2013	2221	19	C
MT2013	2221	20	C
MT2013	2222	1	C
MT2013	2222	2	A
MT2013	2222	3	D
MT2013	2222	4	D
MT2013	2222	5	B
MT2013	2222	6	B
MT2013	2222	7	C
MT2013	2222	8	A
MT2013	2222	9	A
MT2013	2222	10	D
MT2013	2222	11	D
MT2013	2222	12	C
MT2013	2222	13	E
MT2013	2222	14	D
MT2013	2222	15	E
MT2013	2222	16	C
MT2013	2222	17	E

MT2013	2222	18	E
MT2013	2222	19	E
MT2013	2222	20	D
MT2013	2223	1	B
MT2013	2223	2	D
MT2013	2223	3	C
MT2013	2223	4	A
MT2013	2223	5	E
MT2013	2223	6	C
MT2013	2223	7	E
MT2013	2223	8	C
MT2013	2223	9	C
MT2013	2223	10	B
MT2013	2223	11	C
MT2013	2223	12	A
MT2013	2223	13	C
MT2013	2223	14	C
MT2013	2223	15	B
MT2013	2223	16	D
MT2013	2223	17	E
MT2013	2223	18	A
MT2013	2223	19	D
MT2013	2223	20	D
MT2013	2224	1	E
MT2013	2224	2	D
MT2013	2224	3	D
MT2013	2224	4	C
MT2013	2224	5	E
MT2013	2224	6	C
MT2013	2224	7	B
MT2013	2224	8	A
MT2013	2224	9	A
MT2013	2224	10	B
MT2013	2224	11	D
MT2013	2224	12	A
MT2013	2224	13	C
MT2013	2224	14	E
MT2013	2224	15	D

MT2013	2224	16	A
MT2013	2224	17	B
MT2013	2224	18	E
MT2013	2224	19	D
MT2013	2224	20	C



II. PHẦN TỰ LUẬN

Bài 21. Điểm thi đánh giá năng lực của Đại học Quốc gia TP HCM năm 2023 là biến ngẫu nhiên tuân theo quy luật phân phối chuẩn với điểm trung bình là 639.2 và độ lệch chuẩn là 118.6.

Câu 1. (0.5 điểm) Xác định mức điểm sao cho có 70% thí sinh tham gia kỳ thi có điểm lớn hơn mức điểm này.

Lời giải

Gọi X là điểm thi của thí sinh tham gia kỳ thi đánh giá năng lực năm 2023. X là BNN liên tục có phân phối chuẩn với điểm trung bình là $\mu_X = 639.2$ và độ lệch chuẩn là $\sigma_X = 118.6$. Gọi x_0 mức điểm sao cho 70 phần trăm thí sinh tham gia kỳ thi có điểm lớn hơn mức điểm x_0 này. Khi đó ta có

$$\begin{aligned}P(X > x_0) &= 70\% \Rightarrow P(X \leq x_0) = 100\% - 70\% \\&\Rightarrow P\left(\frac{X - \mu_X}{\sigma_X} \leq \frac{x_0 - \mu_X}{\sigma_X}\right) = 30\% = 0.3 \\&\Rightarrow \frac{x_0 - \mu_X}{\sigma_X} \approx -0.53 \Rightarrow x_0 \approx -0.53 \times \sigma_X + \mu_X \approx -0.53 \times 118.6 + 639.2 \approx 576.342\end{aligned}$$

□

Câu 2. (0.5 điểm) Tìm xác suất để trong 200 thí sinh được chọn, có từ 100 đến 120 thí sinh có điểm lớn hơn 650.

Lời giải

Gọi X là điểm thi của thí sinh tham gia kỳ thi đánh giá năng lực năm 2023. X là BNN liên tục có phân phối chuẩn với điểm trung bình là $\mu_X = 639.2$ và độ lệch chuẩn là $\sigma_X = 118.6$ nên xác suất để thí sinh đó có điểm lớn hơn 650 là

$$P(X > 650) = P\left(\frac{X - \mu_X}{\sigma_X} > \frac{650 - \mu_X}{\sigma_X}\right) \approx 0.4637$$

Gọi BNN Y là số thí sinh có điểm lớn hơn 650 trong 200 thí sinh được chọn. Xác suất để thí sinh có điểm lớn hơn 650 là $p \approx 0.4637$. Xác suất để trong 200 thí sinh được chọn, có từ 100 đến 120 thí sinh có điểm lớn hơn 650 là

$$P(100 \leq Y \leq 120) = \sum_{k=100}^{120} C_{200}^k p^k (1-p)^{200-k}$$

Xác suất này có thể được xấp xỉ bởi BNN có phân phối chuẩn chuẩn tắc với hiệu chỉnh liên tục ± 0.5 như sau:

$$P(100 \leq Y \leq 120) \approx P\left(\frac{100 - 0.5 - 200 \times p}{\sqrt{200 \times p(1-p)}} \leq \frac{Y - 200 \times p}{\sqrt{200 \times p(1-p)}} \leq \frac{120 + 0.5 - 200 \times p}{\sqrt{200 \times p(1-p)}}\right) \approx 0.1689.$$

□

Bài 22. Một công ty sản xuất hàng thủ công mỹ nghệ có 2 cơ sở sản xuất cùng một mặt hàng. Khi theo dõi về quá trình hoàn thiện mỗi sản phẩm ở từng cơ sở, công ty thu được các số liệu mẫu như dưới đây. Ở cơ sở 1:

Thời gian hoàn thiện mỗi sản phẩm (giờ)	14-16	16-18	18-20	20-22	22-24	24-26
Số sản phẩm tương ứng	3	7	16	22	14	8

Ở cơ sở 2, người ta theo dõi thời gian hoàn thiện của 9 sản phẩm thì thấy tổng thời gian hoàn thiện của chúng 195 giờ và phương sai mẫu là 7 giờ. Giả thiết rằng thời gian hoàn thiện mỗi sản phẩm ở hai cơ sở là các biến ngẫu nhiên có phân phối chuẩn.

Ghi chú: Dùng điểm chính giữa mỗi khoảng khi tính trung bình mẫu ở cơ sở 1. Phần hỗ trợ bảng tra số:

	Một số giá trị phân vị phải	
Bậc của hàm student	$t_{0.005}$	$t_{0.01}$
79	2.6395	2.3745
78	2.6403	2.3751
77	2.6412	2.3758
10	3.1693	2.3758
9	3.2498	2.8214

Câu 3. (0.5 điểm) Tìm khoảng tin cậy 95% cho tỷ lệ sản phẩm được hoàn thiện dưới 20 giờ ở cơ sở 1.

Lời giải

Ta có $n = 70$; $f = 26/70 \approx 0.3714$

Ngưỡng sai số:

$$\varepsilon = \frac{z_{\alpha/2} \times \sqrt{f(1-f)}}{\sqrt{n}} = \frac{1.96 \times \sqrt{\frac{26}{70} \left(1 - \frac{26}{70}\right)}}{\sqrt{70}} \approx 0.1132$$

Khoảng tin cậy cần tìm: (0.2582; 0.4846)

□

Câu 4. (1.5 điểm) Chúng ta có thể xem như cơ sở 1 hoàn thiện sản phẩm nhanh hơn so với cơ sở 2 hay không, hãy kiểm định với mức ý nghĩa 1%?

Lời giải

Ta có

$$n_1 = 70; \quad \bar{x}_1 = 20.7429; \quad s_1^2 = 6.7155; \quad s_1 = 2.5914$$

$$n_2 = 9; \quad \bar{x}_2 = 21.6667; \quad s_2^2 = 7; \quad s_2 = 2.6458$$

Ký hiệu $\mu_1; \mu_2$ lần lượt là thời gian trung bình để hoàn thiện một sản phẩm ở cơ sở 1 và cơ sở 2.

Giả thuyết $H_0: \mu_1 = \mu_2$

Đối thuyết $H_1: \mu_1 < \mu_2$

Do $s_1/s_2 \in [1/2; 2]$ nên xem như $\sigma_1^2 = \sigma_2^2$

Phương sai gộp: $s^2 = 6.7451$ (sinh viên cần ghi công thức tính)

Miền bác bỏ $RR = (-\infty; -t_{0.01}^{(77)}) = (-\infty; -2.3758)$

Tính giá trị tktđ: $t_{qs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} = -1.0045$

Do $t_{qs} \notin RR$ nên chưa bác bỏ được H_0 . Chưa thể nói cơ sở 1 hoàn thiện sản phẩm nhanh hơn so với cơ sở 2.

□

