

# Đáp án phần tự luận cuối kỳ 212

Ngày 25 tháng 5 năm 2022

## Ca 1

**Câu 16** Để so sánh chất lượng sinh viên giữa 4 ngành học tại một trường đại học kinh tế, người ta đã thực hiện khảo sát điểm trung bình học kỳ (thang điểm 10) cho 20 sinh viên được chọn ngẫu nhiên từ mỗi ngành và kết quả khảo sát được ghi nhận như bên dưới.

Ngành học	Điểm trung bình				
Kế toán	7.1	6.7	6.6	6.6	7.36
Tài chính	6.7	6.8	7.0	5.6	6.2
Nhân sự	7.5	7.5	8.4	7.6	9.8
Tiếp thị	4.5	6.0	5.1	4.5	5.8

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

- Hãy dùng phương pháp Anova để so sánh điểm trung bình học kỳ giữa 4 ngành học 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%.

Giải

- (1.25 điểm) mỗi ý là 0.25 điểm.

- Phát biểu giả thuyết thống kê:

$H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ : Không có sự khác biệt về điểm trung bình học kỳ giữa 4 ngành học.

$H_1: \exists i \in \overline{1, 4}$  sao cho  $\tau_i \neq 0$ : Có sự khác biệt về điểm trung bình học kỳ giữa 4 ngành học.

Tính các tổng và trung bình nhóm:

Ngành học	Điểm trung bình					tổng	trung bình
Kế toán	7.1	6.7	6.6	6.6	7.36	34.36	6.872
Tài chính	6.7	6.8	7.0	5.6	6.2	32.3	6.46
Nhân sự	7.5	7.5	8.4	7.6	9.8	40.8	8.16
Tiếp thị	4.5	6.0	5.1	4.5	5.8	25.9	5.18
						133.36	6.668

Tổng bình phương của các giá trị quan trắc:  $\sum_{ij} y_{ij} = 919.5296$

- Tính các tổng bình phương:

$$\begin{aligned} SST &= 919.5296 - \frac{133.36^2}{20} = 30.28512 \\ SSB &= \frac{(34.36^2 + 32.3^2 + 40.8^2 + 25.9^2)}{5} - \frac{133.36^2}{20} = 22.62544 \\ SSE &= SST - SSB = 7.65968 \end{aligned}$$

- Tính các trung bình phương:

$$\begin{aligned} MSB &= \frac{SSB}{k-1} = \frac{22.62544}{3} = 7.541813 \\ MSE &= \frac{SSE}{k(n-1)} = \frac{7.65968}{16} = 0.47873 \end{aligned}$$

- Tính giá trị kiểm định thống kê:

$$f = \frac{MSB}{MSE} = \frac{7.541813}{0.47873} = 15.75379$$

1

- Xác định miền bác bỏ:

$$RR = (3.24, +\infty)$$

Kết luận: vì  $f \in RR$  nên đủ cơ sở để bác bỏ  $H_0$ , tức là dữ liệu cung cấp đủ chứng cứ để kết luận rằng có sự khác biệt trong điểm trung bình học kỳ giữa 4 ngành học.

(b) **(0.75 điểm)** Mỗi ý là **0.25 điểm**.

- Tính giá trị LSD:

$$LSD = t_{\alpha/2}^{k(n-1)} \sqrt{\frac{2MSE}{n}} = 2.12 \times 0.437598 = 0.9277077$$

- Tính chênh lệch cho từng cặp trung bình mẫu:

$$|y_1 - y_2| = 0.412 < LSD$$

$$|y_1 - y_3| = 1.288 > LSD$$

$$|y_1 - y_4| = 1.692 > LSD$$

$$|y_2 - y_3| = 1.7 > LSD$$

$$|y_2 - y_4| = 1.28 > LSD$$

$$|y_3 - y_4| = 2.98 > LSD$$

- Kết luận: Không đủ chứng cứ để kết luận rằng có sự khác biệt về điểm trung bình giữa hai ngành học kế toán và tài chính, các cặp còn lại đều có sự khác biệt về điểm trung bình.

**Câu 17** Khi đo chỉ số chất lượng không khí AQI trong 5 ngày ở Hà nội, người ta ghi nhận được số liệu dưới đây. Ký hiệu X là chỉ số AQI của NO2 và Y là chỉ số AQI của O3.

X	27	29	35	21	13
Y	36	31	41	30	26

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 thẳng hồi quy tuyến tính của Y theo X.

Giải

- **(0.5 điểm)**

Tính các tổng:

$$\begin{aligned} \sum x_i &= 125, & \sum y_i &= 164 \\ \sum x_i y_i &= 4274, & \sum x_i^2 &= 3405 \\ \sum y_i^2 &= 5514 \end{aligned}$$

Tính các tổng bình phương:

$$\begin{aligned} S_{xx} &= 3405 - \frac{125^2}{5} = 280 \\ S_{xy} &= 4274 - \frac{125 \times 164}{5} = 174 \\ S_{yy} &= 5514 - \frac{164^2}{5} = 134.8 \end{aligned}$$

(a) **(0.5 điểm)**

Tính hệ số tương quan mẫu:

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}} = \frac{174}{\sqrt{280 \times 134.8}} = 0.8956231$$

Nhận xét: Giá trị  $r_{xy}$  dự đoán mối quan hệ tuyến tính đồng biến mạnh giữa x và y.

(b) **(0.5 điểm)**

Tính các hệ số hồi quy:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{174}{280} = 0.6214286$$
$$\hat{\beta}_0 = \frac{\sum x_i}{n} - \hat{\beta}_1 \times \frac{\sum y_i}{n} = 17.26429$$

Viết đường thẳng hồi quy:

$$\hat{y} = 17.26429 + 0.6214286x$$

(c) **(0.5 điểm)**

Tính

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{S_{yy} - \hat{\beta}_1 \times S_{xy}}{n} = 8.890476$$

Tìm

$$t_{\alpha/2}^{n-2} = 3.182$$

Kết luận: 95%KTC cho  $\beta_1$  là:  $[0.05442777, 1.188429]$ .

## Ca 2

**Câu 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				
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.

- (a) 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%.
- (b) Thực hiện so sánh bội bằng phương pháp LSD của Fisher và nhận xét kết quả với mức ý nghĩa 5%.

Giải

(a) **(1.25 điểm) mỗi ý là 0.25 điểm.**

- Phát biểu giả thuyết thống kê:

$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$ : Không có sự khác biệt về trung bình biên độ ngày của nhiệt giữa 3 thành phố trên.

$H_1 : \exists i \in \overline{1,3}$  sao cho  $\tau_i \neq 0$ : Có sự khác biệt về trung bình biên độ ngày của nhiệt giữa 3 thành phố trên.

Tính các tổng và trung bình nhóm:

Thành phố	Biên độ ngày của nhiệt					tổng	trung bình
Hà nội	8.1	4.5	7.2	7.3	8.1	35.2	7.04
Huế	10.4	10.3	12.1	14.3	11.0	58.1	11.62
TP Hồ Chí Minh	8.8	10.1	11.1	9.1	10.0	49.1	9.82
						142.4	9.493333

Tổng bình phương của các giá trị quan trắc:  $\sum_{ij} y_{ij} = 1428.22$

- Tính các tổng bình phương:

$$SST = 1428.22 - \frac{142.4^2}{15} = 76.36933$$
$$SSB = \frac{(35.2^2 + 58.1^2 + 49.1^2)}{5} - \frac{142.4^2}{15} = 53.24133$$
$$SSE = SST - SSB = 23.128$$

- Tính các trung bình bình phương:

$$MSB = \frac{SSB}{k-1} = \frac{53.24133}{2} = 26.62067$$
$$MSE = \frac{SSE}{k(n-1)} = \frac{23.128}{12} = 1.927333$$

- Tính giá trị kiểm định thống kê:

$$f = \frac{MSB}{MSE} = \frac{26.62067}{1.927333} = 13.81218$$

- Xác định miền bác bỏ:

$$RR = (3.89, +\infty)$$

Kết luận: vì  $f \in RR$  nên đủ cơ sở để bác bỏ  $H_0$ , tức là dữ liệu cung cấp đủ chứng cứ để kết luận rằng khác biệt về trung bình biên độ ngày của nhiệt giữa 3 thành phố trên.

(b) **(0.75 điểm)** Môi ý là **0.25 điểm**.

- Tính giá trị LSD:

$$LSD = t_{\alpha/2}^{k(n-1)} \sqrt{\frac{2MSE}{n}} = 2.179 \times 0.8780281 = 1.913223$$

- Tính chênh lệch cho từng cặp trung bình mẫu:

$$|y_1 - y_2| = 4.58 > LSD$$

$$|y_1 - y_3| = 2.78 > LSD$$

$$|y_2 - y_3| = 1.8 < LSD$$

- Kết luận: Không đủ chứng cứ để kết luận rằng có sự khác biệt về biên độ ngày của nhiệt giữa hai thành phố Huế và TPHCM, các cặp còn lại đều có sự khác biệt về biên độ ngày của nhiệt

**Câu 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 thẳng hồi quy tuyến tính của  $Y$  theo  $X$ .

Giải

- **(0.5 điểm)**

Tính các tổng:

$$\begin{aligned} \sum x_i &= 73, & \sum y_i &= 982 \\ \sum x_i y_i &= 7432, & \sum x_i^2 &= 697 \\ & & \sum y_i^2 &= 97008 \end{aligned}$$

Tính các tổng bình phương:

$$\begin{aligned} S_{xx} &= 697 - \frac{73^2}{10} = 164.1 \\ S_{xy} &= 7432 - \frac{73 \times 982}{10} = 263.4 \\ S_{yy} &= 97008 - \frac{982^2}{10} = 575.6 \end{aligned}$$

(a) **(0.5 điểm)**

Tính hệ số tương quan mẫu:

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}} = \frac{263.4}{\sqrt{164.1 \times 575.6}} = 0.85704$$

Nhận xét: Giá trị  $r_{xy}$  dự đoán mối quan hệ tuyến tính đồng biến mạnh giữa  $x$  và  $y$ .

(b) **(0.5 điểm)**

Tính các hệ số hồi quy:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{263.4}{164.1} = 1.605119$$
$$\hat{\beta}_0 = \frac{\sum x_i}{n} - \hat{\beta}_1 \times \frac{\sum y_i}{n} = 86.48263$$

Viết đường thẳng hồi quy:

$$\hat{y} = 86.48263 + 1.605119x$$

(c) **(0.5 điểm)**

Tính

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{S_{yy} - \hat{\beta}_1 \times S_{xy}}{n} = 19.10146$$

Tìm

$$t_{\alpha/2}^{n-2} = 2.306$$

Kết luận: 95%KTC cho  $\beta_1$  là:  $[0.818366, 2.391872]$ .