

BÀI TẬP XÁC SUẤT THỐNG KÊ CHƯƠNG 9

18.1. Vẽ đồ thị sau chỉ rõ hệ số góc, giao điểm với trục tung

$$Y=2+3X$$

$$Y=2-2X$$

$$Y=-2+4X$$

$$Y=X$$

$$Y=4$$

18.2. Với thông tin trong bảng sau, chỉ ra xem liệu mô hình tuyến tính có khả thi không?

a) X 2 3 5 7 9

Y 6 9 4 7 8

b) X 1 3 5 4 7

Y 5 7 10 9 16

c) X 7 9 2 3 6

Y 4 1 6 10 5

18.3. Vẽ các điểm biểu diễn các quan sát của X,y

X 1 2 3 4 5 6

Y 4 6 7 7 9 11

Vẽ phương trình tuyến tính từ thông tin trên. Chỉ rõ hệ số góc, giao điểm với trục tung.

18.4 Cho 6 điểm sau:

X 1 2 3 4 5 6 7 8 9

Y 5 28 17 14 27 33 39 26 30

a) Vẽ biểu đồ

b) Xác định đường bình phương nhỏ nhất

18.5 Các quan sát về 2 biến sau được ghi lại trong bảng:

X 1 2 3 4 5 6 7 8 9

Y 5 28 17 14 27 33 39 26 30

a) Vẽ biểu đồ phân tán

b) Xây dựng hàm hồi quy tuyến tính

18.6 Có 10 quan sát được tổng hợp theo số liệu sau:

$$\sum x = 31 \quad \sum y = 37 \quad \sum xy = 75 \quad \sum x^2 = 103 \quad \sum y^2 = 445$$

Xây dựng hàm hồi quy tuyến tính

18.7 Có 25 quan sát được tổng hợp như sau:

$$\sum x = 62.5 \quad \sum y = 129 \quad \sum xy = 141 \quad \sum x^2 = 317.5 \quad \sum y^2 = 936.4$$

Tìm hàm hồi quy

18.8 Với bài 18.4, tính tổng $SSE = \sum (\hat{y} - y)^2$

18.9 Với bài 18.5 tính tổng $SSE = \sum (\hat{y} - y)^2$

18.10 Trong một nghiên cứu giữa 2 biến x và y thì thu được kết quả sau:

$$\sum x = 105 \quad \sum y = 4414 \quad \sum xy = 37525 \quad \sum x^2 = 956 \quad \sum y^2 = 1818421$$

$$n = 15$$

Tính SSE

18.11 Cho lợi nhuận bình quân trên mỗi đôla doanh số y tính bằng cent, x là số công nhân trong mỗi cửa hàng?

| | | | | | | | | |
|---|----|----|----|----|----|----|----|----|
| X | 3 | 6 | 4 | 5 | 2 | 5 | 4 | 1 |
| Y | 22 | 30 | 20 | 25 | 18 | 26 | 22 | 19 |

a) Tìm hàm hồi quy biểu hiện mức lợi nhuận trung bình trên mỗi đôla doanh số

b) Vẽ các điểm và đường hồi quy

c) Liệu rằng mô hình tuyến tính có là mô hình tối ưu nhất

d) Nêu ý nghĩa kinh tế của hệ số góc đường hồi quy

18.12 Cho X là số lượng ô tô Y là chi phí trên đầu xe trong một tháng

Sau 12 tháng điều tra thì cho ra kết quả:

$$\sum x = 157 \quad \sum y = 57 \quad \sum xy = 987 \quad \sum x^2 = 4102 \quad \sum y^2 = 413$$

Tìm hàm hồi quy và xác định giá trị \hat{B} ????

18.15 Từ hồi quy được dùng năm 1885 bởi ông Francis Galton trong bài nghiên cứu về mối quan hệ giữa chiều cao trẻ em và cha mẹ. Ông ta đưa ra “quy tắc chung về hồi quy” cụ thể rằng mỗi người đều có điểm chung với họ hàng, nhưng trung bình thì ít hơn. Năm 1903, 2 nhà thống kê, K. Pearson và A. Lee lấy mẫu ngẫu nhiên 1078 cặp cha con để xem xét luật Galton. Dòng hồi quy mẫu là

Chiều cao của con = $33.73 + 0.516 \cdot$ chiều cao của cha

- Giải thích hệ số tương quan
- Dòng hồi quy nói gì về chiều cao của con có bố cao?
- Dòng hồi quy nói gì về chiều cao của con có bố thấp?

18.16. Mục đích của quảng cáo là có càng nhiều người xem nhớ sản phẩm và cuối cùng mua nó. Trong một thí nghiệm để xác định độ dài của quảng cáo tác động đến ký ức con người, 60 mẫu ngẫu nhiên được chọn để xem chương trình tivi dài 1 tiếng. Ở giữa chương trình, một quảng cáo về kem đánh răng xuất hiện. Mỗi người xem quảng cáo có độ dài từ 20 đến 60 giây. Nội dung của quảng cáo là như nhau. Sau chương trình, mỗi người được kiểm tra xem mức độ nhớ của họ về sản phẩm. Thời gian quảng cáo và điểm được lưu ở file XR18-16

- Vẽ đồ thị để xác định liệu mô hình tuyến tính có phù hợp?
- Xác định đường bình phương nhỏ nhất
- Giải thích hệ số tương quan

Sử dụng phần mềm hoặc tính bằng tay $n=60$; $\bar{x}=38.0$; $\bar{y}=13.8$

$SS_{ry} = 3060$

$SS_r = 11440$

$SS_y = 2829.6$

18.17. Sau nhiều kỳ không có nhiều thành công, Pat Statsdud (1 sinh viên thống kê) quyết định cố gắng. Pat cần biết bí mật của thành công cho sinh viên đại học. Sau nhiều giờ tra đổi với sinh viên khác, thành công hơn. Pat đưa ra một lý thuyết cơ bản: học càng nhiều thì điểm càng tốt. Để kiểm tra lý thuyết, Pat lấy mẫu ngẫu nhiên 100 sinh

viên trong 1 môn kinh tế và hỏi mỗi người cho biết thời gian học kinh tế trung bình và điểm nhận được. Dữ liệu được lưu ở cột 1 (thời gian học theo giờ) và 2 (điểm cuối kỳ trên 100) trong file XR18-17

- xác định phương trình hồi quy mẫu
- giải thích mức ý nghĩa
- Dốc của đường hồi quy có hợp lý? Nếu dốc có hướng khác, điều gì có thể nói cho bạn biết?

Dùng phần mềm hoặc tính bằng tay với

$n=60$; $\bar{x}=27.95$; $\bar{y}=74.06$; $SS_{xy}=15\,241$; $SS_x=8118.7$; $SS_y=360360$

18.18 Giả sử một nhà thống kê muốn cập nhật nghiên cứu trong bài tập 18.15. Cô ta thu thập 400 cặp cha con và lưu trữ trong cột 1 (chiều cao của cha theo centimet) và con (chiều cao con theo centimet) ở file XR18-18.

- Xác định đường hồi quy tuyến tính
- Giá trị β_0 cho bạn biết gì?
- Giá trị β_1 cho bạn biết gì

Sử dụng bộ phần mềm hoặc tính bằng tay

$n=400$; $\bar{x}=167.86$; $\bar{y}=171.74$; $SS_{xy}=19626$; $SS_x=40980.6$; $SS_y=35265$

18.22. Kiểm định mỗi giả thiết sau:

a. $H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$

$\beta_1 = 1.87$

$S_{\beta_1} = 0.63$

$n = 25$

$\alpha = 0.05$

b. $H_0: \beta_1 = 0$

$H_A: \beta_1 < 0$

$\beta_1 = -26.32$

$S_{\beta_1} = 14.51$

$n = 100$

$\alpha = 0.01$

c. $H_0: \beta_1 = 0$

$H_A: \beta_1 > 0$

$\beta_1 = 0.056$

$S_{\beta_1} = 0.021$

$n = 10$

$\alpha = 0.05$

18.23. Với độ tin cậy 95%, hãy tính độ dốc của đường hồi quy với dữ liệu sau:

$$\sum x = 55 \quad \sum y = 325 \quad \sum xy = 929$$

$$\sum x^2 = 175 \quad \sum y^2 = 6051 \quad n = 20$$

18.24. 12 quan sát về x và y cho kết quả như sau:

$$\sum x = 65 \quad \sum y = 515 \quad \sum xy = 3085 \quad \sum x^2 = 445 \quad \sum y^2 = 24\,815$$

- Với mức ý nghĩa 1% thì đây có phải là bằng chứng đầy đủ để xác định có một mối quan hệ tuyến tính giữa x và y hay không?
- Xác định ước lượng khoảng của độ dốc với độ tin cậy 99%.
- Tính R^2 . Số liệu này cho biết điều gì về đường hồi quy?

18.25. Liên quan tới bài 18.14.

- Tìm sai số tiêu chuẩn của độ lệch. Giá trị này cho biết điều gì về đường hồi quy mẫu?
- Xác định ước lượng khoảng của độ dốc với độ tin cậy 95%.
- Với mức ý nghĩa 1%, có thể kết luận quảng cáo này đem lại hiệu quả làm tăng doanh thu không?
- Xác định hệ số tương quan và bình luận về giá trị của nó.

18.26. Một kế hoạch chia sẻ lợi nhuận mới đã được giới thiệu tại một nhà máy sản xuất các bộ phận của ô tô ở Geelong năm ngoái. Cả người quản lý và đại diện của hiệp hội đều muốn xác định xem số năm kinh nghiệm của một công nhân sẽ ảnh hưởng đến năng suất thu được của nhà máy như thế nào. Sau khi kế hoạch này đã được thực hiện một thời gian thì những số liệu sau đã được thu thập lại:

| Công nhân | Số năm làm việc ở dây chuyền lắp ráp x | Số lượng sản phẩm được sản xuất hàng ngày y |
|-----------|--|---|
| 1 | 15.1 | 110 |
| 2 | 7.0 | 105 |
| 3 | 18.6 | 115 |
| 4 | 23.7 | 127 |

| | | |
|--|------|-----|
| 5 | 11.5 | 98 |
| 6 | 16.4 | 103 |
| 7 | 6.3 | 87 |
| 8 | 15.4 | 108 |
| 9 | 19.9 | 112 |
| $\sum x = 133.9 \quad \sum y = 965 \quad \sum xy = 14\,801.2 \quad \sum x^2 = 2258.73 \quad \sum y^2 = 104\,469$ | | |

- Tìm đường hồi quy bình phương nhỏ nhất?
- Tính sai số tiêu chuẩn của độ lệch. Giá trị này cho biết điều gì về mối quan hệ giữa x và y?
- Với mức ý nghĩa 5%, chúng ta có thể kết luận rằng số năm kinh nghiệm của một công nhân sẽ ảnh hưởng đến năng suất thu được của nhà máy hay không?
- Đo lường xem đường tuyến tính mẫu phù hợp như thế nào?

18. 27. Một chuỗi các siêu thị tiến hành một cuộc khảo sát để giúp xác định vị trí mong muốn cho những cửa hàng mới. Ban quản lý của chuỗi các siêu thị này muốn biết xem mối quan hệ tuyến tính có xuất hiện giữa tiền lương hàng tuần sau khi đã tính thuế và phí hàng tuần dùng cho thực phẩm hay không. Một mẫu ngẫu nhiên gồm 8 hộ gia đình đã chỉ ra những số liệu trong bảng sau:

| Gia đình | Tiền lương hàng tuần sau khi đã tính thuế x | Phí tổn hàng tuần dùng cho thực phẩm y |
|----------|---|--|
| 1 | 600 | 160 |
| 2 | 400 | 110 |
| 3 | 540 | 150 |
| 4 | 360 | 90 |
| 5 | 500 | 130 |
| 6 | 720 | 200 |

| | | |
|---|-----|-----|
| 7 | 450 | 120 |
| 8 | 680 | 180 |

- Tìm đường hồi quy bình phương nhỏ nhất?
- Tính sai số tiêu chuẩn của độ lệch. Giá trị này cho biết điều gì về mối quan hệ giữa x và y?
- Với $\alpha = 0.01$, chúng ta có thể kết luận rằng số năm kinh nghiệm của một công nhân sẽ ảnh hưởng đến năng suất thu được của nhà máy hay không?
- Tính R^2 . Giá trị này cho biết điều gì về độ mạnh của đường hồi quy tuyến tính?

18.28. Để xác định giá lý tưởng cho một sản phẩm mới mà một công ty muốn bán, phòng nghiên cứu của công ty đã chọn 10 địa điểm được cho là có tiềm năng để bán hàng và đề xuất sản phẩm tại mỗi một mức giá khác nhau. Kết quả bán hàng được ghi lại trong bảng sau:

| Vị trí | Giá x | Lượng bán (\$000s) y |
|--------|-------|----------------------|
| 1 | 15.00 | 15 |
| 2 | 15.50 | 14 |
| 3 | 16.00 | 16 |
| 4 | 16.50 | 9 |
| 5 | 17.00 | 12 |
| 6 | 17.50 | 10 |
| 7 | 18.00 | 8 |
| 8 | 18.50 | 9 |
| 9 | 19.00 | 6 |
| 10 | 19.50 | 5 |

- Tìm phương trình của đường hồi quy.
- Với mức ý nghĩa 5% thì đây có phải là bằng chứng đầy đủ để xác định có một mối quan hệ tuyến tính âm giữa giá và lượng bán hay không?

c. Tính hệ số tương quan và R^2 .

18.29. Liên quan tới bài 18.16.

a. Tính sai số chuẩn của ước lượng và cho biết ý nghĩa của giá trị này.

b. Với mức ý nghĩa 5%, chúng ta có thể kết luận rằng độ rộng của khoảng lợi nhuận và trí nhớ của con người liên quan tuyến tính đến nhau hay không?

c. Tính R^2 . Giá trị này cho biết điều gì về đường hồi quy?

18.30. Liên quan tới bài 18.17.

a. Tính sai số tiêu chuẩn của độ lệch và cho biết ý nghĩa của giá trị này.

b. Với mức ý nghĩa 5% thì đây có phải là bằng chứng đầy đủ để xác định có một mối quan hệ tuyến tính giữa thời gian học và điểm cuối kì hay không?

c. Tính R^2 . Giá trị này cho biết điều gì về đường hồi quy?

18.50. Cho 10 cặp quan sát của biến phân loại theo số lượng như sau. Tính hệ số tương quan Spearman:

| Quan sát | X | Y |
|----------|----|----|
| 1. | 1 | 2 |
| 2. | 3 | 3 |
| 3. | 4 | 5 |
| 4. | 7 | 6 |
| 5. | 6 | 8 |
| 6. | 5 | 4 |
| 7. | 8 | 9 |
| 8. | 2 | 1 |
| 9. | 10 | 7 |
| 10. | 9 | 10 |

18.51: kiểm định giả thiết sau:

a. $H_0: P_3=0$

$H_1: P_3>0$

$R_3= 0.482$

$N= 20$

$\alpha=0.01$

b. $H_0: P_3=0$

$H_1: P_3\neq 0$

$R_3= 0.413$

$N= 25$

$\alpha=0.05$

c. $H_0: P_3=0$

$H_1: P_3<0$

$R_3= -0.558$

$N= 10$

$\alpha=0.05$

18.52. cho bảng số liệu sau:

| | | | | | | |
|---|-----|-----|-----|-----|-----|-----|
| X | 115 | 220 | 86 | 99 | 50 | 110 |
| Y | 1.0 | 1.3 | 0,6 | 0,8 | 0,5 | 0.7 |
| | | | | | | |

Tính hệ số tương quan Pearson và kiểm tra iệu giữa hai biến có quan hệ tuyến tính hay không? (với $\alpha=0.5$)

b. Tính hệ số tương quan Spearman và kiểm tra iệu giữa hai biến có quan hệ tuyến tính hay không? (với $\alpha=0.5$)

18.53. Quay trở lại bài 18.16

a. xác định hệ số tương quan. Với mức ý nghĩa 5% liệu có tồn tại quan hệ tuyến tính giữa thương mại và điểm bài kiểm tra trí nhớ hay không?

b. giả thiết rằng những điều kiện để thực hiện kiểm tra ở bài tập 18.16 không được đáp ứng đủ. Liệu với những số liệu trên đã đủ để kết luận thương mại càng kéo dài thì điểm số kiểm tra trí nhớ sẽ càng cao hay không?

18.54. Lượng hai loại cổ phiếu bị trả lại hàng tuần được ghi lại trong thwoi gian 13 tuần.sau đây là ghi chép của những dữ liệu này.

a. giả thiết rằng việc trả lại được phân bố như thông thường, liệu có thể kết luận hai cổ phiếu tương quan với nhau ko?(mức ý nghĩa 5%)

b. nếu điều giả sử là ngược lại, thì có thể kết luận như trên không
bảng:

| Tuần | CP1 | CP2 |
|------|-----|-----|
| 1 | -7 | 6 |
| 2 | -4 | 6 |
| 3 | -7 | -4 |
| 4 | -3 | 9 |
| 5 | 2 | 3 |
| 6 | -10 | -3 |
| 7 | -10 | 7 |
| 8 | 5 | -3 |
| 9 | 1 | 4 |
| 10 | -4 | 7 |
| 11 | 2 | 9 |
| 12 | 6 | 5 |
| 13 | -13 | -7 |

18.55. quản lý của một nhà máy muốn biết liệu kinh nghiệm làm việc của các công nhân có ảnh hưởng đến hiệu suất của họ hay không. Bà ta đã chọn ngẫu nhiên 24 công nhân và ghi chép lại kinh nghiệm cũng như thành tích công việc mà họ đạt được (qua giám sát). Cột 1: năm kinh nghiệm, cột 2: năng suất làm việc với các mức độ : 5: xuất sắc, 4: tốt, 3: trung bình, 2: tạm được, 1: kém. với $\alpha=0.10$ liệu có thể khẳng định kinh nghiệm làm việc có ảnh hưởng tới năng suất làm việc hay không?

| Công nhân | Kinh nghiệm | Năng suất |
|-----------|-------------|-----------|
| 1 | 1 | 1 |
| 2 | 17 | 4 |
| 3 | 20 | 4 |
| 4 | 9 | 5 |
| 5 | 2 | 2 |
| 6 | 13 | 4 |
| 7 | 9 | 3 |
| 8 | 23 | 5 |
| 9 | 7 | 2 |
| 10 | 10 | 5 |
| 11 | 12 | 5 |
| 12 | 24 | 2 |
| 13 | 8 | 2 |
| 14 | 20 | 5 |
| 15 | 21 | 3 |
| 16 | 19 | 2 |
| 17 | 1 | 1 |
| 18 | 22 | 3 |
| 19 | 20 | 4 |

| | | |
|----|----|---|
| 20 | 11 | 3 |
| 21 | 18 | 5 |
| 22 | 14 | 4 |
| 23 | 21 | 3 |
| 24 | 21 | 1 |

18.56. lượng đổi trả hàng tháng từ tháng 6/1996 tới t5/2001 của cổ phiếu ý và hồng koong được lưu trữ tổng file XR18-56. Điều này đã đủ khẳng định 2 cổ phiếu trên có quan hệ tuyến tính chưa?

18.57. làm tương tự bài tập 18.56. giả sử rằng 2 biến số trên không tuân theo phân bố thường. hãy kiểm định giả thiết liệu hai đại lượng trên có quan hệ tuyến tính không?

18.58. cho 6 điểm sau đây:

| | | | | | | |
|---|----|----|---|---|---|---|
| X | -5 | -2 | 0 | 3 | 4 | 7 |
| Y | 15 | 9 | 7 | 6 | 4 | 1 |

- Xác định phương trình hồi quy
- Dùng phương trình đó tìm giá trị dự đoán của y
- Từ giá trị dự đoán và giá trị thật của y hãy tính thặng dư.
- Tính giá trị thặng dư tiêu chuẩn
- Xác định các giá trị outliers (???) có thể

18.59. quan sát của 2 biến được ghi lại trong bảng sau:

| | | | | | | | | | |
|---|---|----|----|----|----|----|----|----|----|
| X | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| y | 5 | 28 | 17 | 14 | 27 | 33 | 39 | 26 | 30 |

- Tính phương trình hồi quy
- Dùng phương trình đó tìm giá trị dự đoán của y
- Từ giá trị dự đoán và giá trị thật của y hãy tính thặng dư. Tính giá trị thặng dư tiêu chuẩn

- d. Xác định các giá trị outliers (???) có thể
- e. Vẽ đồ thị giá trị thặng dư với giá trị dự đoán của y . biến ở đây có trở thành một hằng số được không? Giải thích?

18.60. mỗi cặp giá trị dưới đây là giá trị thực và giá trị dự đoán của y (dựa trên mẫu hồi quy đơn giản). vẽ biểu đồ giá trị c của y (trên trục hoành) và giá trị thặng dư (trục tung).

| | | | | | | | | | | | | | | | | | | |
|---|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|
| a | y | 155 | 112 | 163 | 130 | 143 | 182 | 160 | 104 | 125 | 161 | 189 | 102 | 142 | 149 | 180 | | |
| | \hat{y} | 143 | 108 | 180 | 133 | 146 | 193 | 140 | 101 | 126 | 176 | 200 | 97 | 145 | 151 | 158 | | |
| b | y | 10 | 22 | 29 | 15 | 24 | 13 | 17 | 23 | 11 | 27 | 19 | 26 | 20 | 14 | | | |
| | \hat{y} | 7 | 21 | 29 | 13 | 25 | 16 | 19 | 22 | 14 | 27 | 17 | 27 | 22 | 11 | | | |
| c | y | 46 | 40 | 53 | 60 | 56 | 62 | 44 | 49 | 52 | 59 | 45 | 55 | 47 | 61 | 42 | 57 | 50 |
| | \hat{y} | 48 | 43 | 54 | 63 | 54 | 65 | 46 | 47 | 49 | 56 | 41 | 53 | 44 | 57 | 45 | 62 | 51 |

18.61. quay trở lại bài 18.4

- a. Tính phương trình hồi quy
- b. Dùng phương trình đó tìm giá trị dự đoán của y
- c. Từ giá trị dự đoán và giá trị thật của y hãy tính thặng dư. Tính giá trị thặng dư tiêu chuẩn
- d. Xác định các giá trị outliers (???) có thể
- e. Vẽ đồ thị giá trị thặng dư với giá trị dự đoán của y . biến ở đây có trở thành một hằng số được không? Giải thích?

18.62. quay trở lại bài 18.5

- a. Tính phương trình hồi quy
- b. Dùng phương trình đó tìm giá trị dự đoán của y
- c. Từ giá trị dự đoán và giá trị thật của y hãy tính thặng dư. Tính giá trị thặng dư tiêu chuẩn
- d. Xác định các giá trị outliers (???) có thể

e. Vẽ đồ thị giá trị thặng dư với giá trị dự đoán của y . biến ở đây có trở thành một hằng số được không? Giải thích?

18.70: Một sinh viên nông nghiệp của một trường đại học ở New England đã lấy được sổ ghi chép của bố anh ta về một vài dữ liệu trong đó có đề cập đến sản lượng thu hoạch và lượng thuốc trừ sâu đã sử dụng, lượng mưa trung bình, số giờ chiếu sáng trung bình và nhiệt độ trung bình hàng ngày. Theo như con số xấp xỉ đầu tiên, anh ta muốn xem lại sản lượng thu hoạch dựa trên lượng thuốc trừ sâu đã sử dụng, dựa vào các dữ liệu được cung cấp như sau:

| Lượng thuốc trừ sâu (kg/hecta) X | Sản lượng (nghìn kg/ hecta) y |
|-------------------------------------|---|
| 220 | 36 |
| 450 | 72 |
| 250 | 48 |
| 320 | 51 |
| 500 | 80 |
| 250 | 40 |
| 330 | 55 |
| 430 | 72 |
| 240 | 39 |
| 280 | 45 |
| 370 | 62 |
| 400 | 71 |
| 410 | 79 |
| 450 | 75 |
| $\sum x = 4900$ | $\sum y = 825$ |
| $\sum x^2 = 1\,825\,600$ | $\sum y^2 = 51\,891$ $\sum xy = 307\,190$ |

- a. Minh họa bằng đồ thị và nhận xét về sự thích hợp của việc sử dụng mô hình đường suy thoái
- b. Tìm những góc vuông có đường suy thoái nhỏ nhất cho những dữ liệu trên
- c. Kiểm tra để quyết định xem mối quan hệ đường suy thoái có tồn tại giữa 2 biến số hay không?
- c. Tính toán hệ số xác định và giải thích giá trị đó
- d. Dự đoán sản lượng thu hoạch với độ tin cậy 90% dựa vào việc đã sử dụng 500kg thuốc sâu. Sản lượng vừa dự đoán so sánh với sản lượng thực sự khi có sử dụng 500kg thuốc sâu như thế nào?

18.71: Để phản hồi lại với tất cả các lời than phiền của phụ huynh và học sinh về việc mức phí tài liệu học tập của học sinh 9-12 là quá cao, hội đồng nhà trường đang nỗ lực để giữ nguyên mức phí như trước đây. Nhà trường đã lựa chọn ngẫu nhiên 3 học sinh từ năm 9-12 với kết quả như sau:

| Chi phí hàng năm | | | |
|------------------|-----|-----|-----|
| Học sinh | | | |
| Năm | 1 | 2 | 3 |
| 9 | 215 | 200 | 210 |
| 10 | 210 | 205 | 220 |
| 11 | 210 | 220 | 235 |
| 12 | 225 | 225 | 215 |

- a. Tìm phương trình của đường suy thoái
- b. Liệt kê những dữ liệu trên cung cấp đủ dẫn chứng để chỉ ra rằng những năm lớn hơn thì xảy ra tình trạng chi phí cao hơn (với $\alpha = 0.10$).
- c. Với độ tin cậy = 95%, hãy dự đoán mức chi phí hàng năm của việc học hết năm thứ 11

d. Với độ tin cậy 98%, hãy ước lượng giá trị trung bình của chi phí hàng năm cho việc học hết năm 9.

18.72: Quản lý của một hãng thiết bị gia dụng đang xem lại số tiền chi tiêu cho quảng cáo hàng tuần. Trong 6 tháng vừa qua tất cả các quảng cáo cho cửa hàng đã xuất hiện trên báo địa phương. Số lượng các quảng cáo trên tuần nằm trong khoảng 1-7. Nhân viên bán hàng của cửa hàng đang theo dõi lượng khách đến cửa hàng mỗi tuần. Số lượng các quảng cáo và lượng khách hàng hàng tuần trong 26 tuần qua đã được ghi chép và lưu trong file XR18-72.

a. Xác định đường suy thoái mẫu

b. giải thích các hệ số

c. với $\alpha = 5\%$, liệu nhà quản lý có thể suy ra rằng càng nhiều quảng cáo thì càng có nhiều khách hàng.

d. Tìm và giải thích hệ số xác định

e. Theo quan điểm của bạn liệu đây là 1 bài tập đáng giá để sử dụng phương trình đường suy thoái cho việc dự đoán lượng khách hàng sẽ đến cửa hàng, với thông tin rằng hãng này có ý định quảng cáo 5 lần trên báo địa phương. Nếu bạn nghĩ vậy thì tìm khoảng dự đoán với độ tin cậy 95%. Nếu không, hãy giải thích tại sao.

($n = 26$, giá trị trung bình (\bar{x}) của các quảng cáo = 4.115, giá trị trung bình (\bar{y}) của lượng khách hàng = 384.808. $SS_{xy} = 1850.6$; $SS_x = 86.654$; $SS_y = 463802$).

18.73: Giám đốc sản xuất của một công ty sản xuất ghế ngồi trong xe ô tô quan tâm đến số lượng và chi phí các máy bị vỡ. Vấn đề là ở chỗ các máy móc quá cũ và trở nên kém tin cậy. Tuy nhiên chi phí để thay mới rất tốn kém, và nhà quản lý thì không cho rằng chi phí là vấn đề đơn giản trong thời buổi kinh tế kém như hiện nay. Để quyết định về việc thay máy mới, ông tập hợp dữ liệu về chi phí sửa chữa tháng trước và(trong các tháng) của kế hoạch ...máy hàn. Những dữ liệu này được lưu trong file XR18-73.

a. Tìm đường suy thoái mẫu

b. giải thích các hệ số

c. xác định sai số chuẩn cho ước lượng và thảo luận về vấn đề con số ước lượng này nói lên điều gì

d. Đánh giá thí nghiệm trên với bất kỳ mức ý nghĩa nào bạn cho là hợp lý để quyết định xem tuổi thọ của máy móc và chi phí sửa chữa hàng tháng là thuộc cùng 1 đường.

e. Tìm và giải thích hệ số xác định

f. Sự vừa vặn của đường suy thoái mẫu có đủ tốt để cho phép nhà quản lý dự đoán chi phí sửa chữa hàng tháng các máy hàn (đã 120 năm sử dụng) hay không? Nếu bạn nghĩ vậy thì tìm khoảng dự đoán với độ tin cậy 95%. Nếu không, hãy giải thích tại sao.

(với $n = 20$, tuổi thọ trung bình của máy (x) = 118.35, chi phí sửa chữa trung bình (y) = 395.21. $SS_{xy} = 17\,800$; $SS_x = 7196.6$; $SS_y = 77\,801$).

18.74: Cách đây rất nhiều năm, Coca cola đã cố gắng để thay đổi công thức đã tồn tại 100 năm qua. Một trong những lý do tại sao ban lãnh đạo công ty cảm thấy thay đổi này là cần thiết đó chính và vì sự cạnh tranh với pepsi cola. Kết quả của cuộc khảo sát những người uống pepsi đã chỉ ra rằng họ thích pepsi bởi vì nó ngọt hơn coca. Theo như 1 phần của bản phân tích về phát triển chậm của coca, lãnh đạo của coca cola đã đưa ra những nghiên cứu sâu hơn về các đối tượng đã thưởng thức nhiều loại coca mới. Giả định rằng 1 mẫu ngẫu nhiên gồm 200 người uống coca được đưa cho các loại coca với lượng đường khác nhau. Sau khi thử sản phẩm, mỗi người uống sẽ được hỏi để đánh giá chất lượng. các phản hồi có thể được quy ước như sau:

5= rất ngon 4= ngon 3= bình thường 2= khá ngon 1= không ngon

Những phản hồi và lượng đường (phần trăm khối lượng) của loại sản phẩm đã được uống thử được lưu lại trong cột 1 và 2 của file XR18-74. Với mức ý nghĩa = 5% liệu nhà quản lý có thể suy ra rằng lượng đường ảnh hưởng đến đánh giá của người uống?

(với $n = 200$, lượng đường trung bình (x) = 12.305, trung bình các đánh giá (y) = 2.91. $SS_{xy} = 1306.5$; $SS_x = 7364.4$; $SS_y = 370.4$).

18.75: Một niềm tin chung của các nhà giám sát kinh doanh thế giới đó là những nam giới có chiều cao sẽ kiếm tiền nhiều hơn những người thấp hơn. Trong một báo cáo

nghiên cứu trên well street journal, 30 cử nhân kinh doanh ở độ tuổi 30, đã được hỏi về mức thu nhập hàng năm và chiều cao của họ. Những số liệu này được lưu trong file XR18-75.

- xác định đường suy thoái mẫu và giải thích các hệ số.
- Tìm sai số chuẩn cho ước lượng và giải thích giá trị đó.
- Với mức ý nghĩa 5%, những dữ liệu này liệu có cung cấp đủ chứng cứ thống kê để suy ra rằng những cử nhân kinh doanh có chiều cao sẽ kiếm được nhiều tiền hơn những người thấp hơn không?
- Cung cấp sự đo đạc về độ mạnh của mối quan hệ đường suy thoái giữa thu nhập và chiều cao.
- Bạn có nghĩ rằng mô hình này là đủ cơ sở để sử dụng cho việc ước lượng và dự đoán thu nhập dựa trên thông số về chiều cao? Nếu bạn cho là như vậy, hãy.

- Hãy ước lượng thu nhập trung bình của những doanh nhân cao 1m83 với độ tin cậy 95%

- Dự đoán thu nhập của 1 cử nhân kinh doanh cao 1m75 với độ tin cậy 95%.

(với $n = 30$, chiều cao trung bình (x) = 174.0, thu nhập trung bình (y) = 36 319.69. $SS_{xy} = 207\,715$; $SS_x = 1607.5$; $SS_y = 236\,061\,504$).

18.76: Một nhà nông nghiệp học muốn điều tra những yếu tố quyết định đến sản lượng thu hoạch. Theo đó bà đã tiến hành 1 thí nghiệm trong đó 1 nông trại được chia thành 30 nửa hecta mảnh đất nhỏ. Lượng thuốc trừ sâu sử dụng cho mỗi mảnh đất là khác nhau. Ngô được gieo sau đó và lượng ngô thu được ở cuối mùa đã được ghi chép lại. Những dữ liệu này được lưu trong file XR18-76.

- Xác định đường suy thoái mẫu và giải thích các hệ số.
- Liệu nhà nông nghiệp học đó có thể kết luận rằng có mối quan hệ suy thoái giữa lượng thuốc trừ sâu và sản lượng bắp thu hoạch được?
- Tìm hệ số xác định và giải thích giá trị đó

d. Liệu mô hình đường suy thoái mẫu này là công cụ hữu hiệu cho việc dự đoán sản lượng thu hoạch từ lượng thuốc trừ sâu đã sử dụng hay không. Nếu có, với độ tin cậy 95% hãy tìm khoảng dự đoán sản lượng khi 300kg thuốc trừ sâu được sử dụng. Nếu không, hãy giải thích tại sao?

18.77: Các nhà sản xuất xe hơi được yêu cầu kiểm tra các phương tiện về sự đa dạng của các chất gây ô nhiễm khi thoát khí. Số lượng các chất gây ô nhiễm xảy ra thậm chí là ở cả những loại phương tiện xác định, vì thế rất nhiều phương tiện phải được kiểm tra. Người kỹ sư chịu trách nhiệm kiểm tra đã thu thập dữ liệu (gram/km xe đã đi) về lượng của 2 chất gây ô nhiễm cacbon monoxide và nitrous oxide cho 50 phương tiện. Những dữ liệu này được lưu trong file XR18-77. Kỹ sư đó tin rằng công ty có thể tiết kiệm tiền bằng cách chỉ kiểm tra thông số của 1 loại chất gây ô nhiễm vì 2 chất đó thực ra gần như nhau. Điều đó có nghĩa là, nếu 1 chiếc xe hơi thải với lượng lớn cacbon monoxide, nó cũng sẽ thải ra 1 lượng lớn nitrous oxide. Liệu những dữ liệu đó có ủng hộ cho niềm tin của kỹ sư đó không?

18.78: Trong thập kỷ trước, xã hội nhìn chung và hệ thống xét xử nói riêng đã thay đổi quan điểm của họ về tầm nguy hại của việc lái xe trong tình trạng say xỉn. Trong hầu hết các thẩm quyền, lái xe với nồng độ rượu trong máu vượt quá .05 là hành vi phạm tội. Bởi vì có nhiều yếu tố, rất khó để cung cấp sự hướng dẫn về việc khi nào là an toàn để cho 1 người đã uống rượu có thể lái xe. Trong 1 thí nghiệm để kiểm tra mối quan hệ giữa nồng độ cồn trong máu và cân nặng của người uống, 50 nam giới với cân nặng khác nhau đã được đưa cho 3 chai bia để uống, và sau đó 1 giờ thì họ được đo nồng độ cồn trong máu. Những dữ liệu này được lưu trong file XR18-78.

- a. Nếu chúng ta giả định rằng 2 thông số trên có phân bố chuẩn, liệu chúng ta có thể kết luận rằng nồng độ cồn trong máu và cân nặng có mối quan hệ với nhau?
- b. Sau khi kiểm tra dữ liệu, người thực hành thống kê chịu trách nhiệm về thí nghiệm đã kết luận rằng sự phân tích suy giảm là không hợp lý (bởi anh ta xác định rằng giới hạn

sai số là không bình thường). Bạn có thể rút ra kết luận gì từ dữ liệu trên về mối quan hệ giữa nồng độ cồn trong máu và cân nặng?

18.79: Một vài nhà bình phẩm ti vi than phiền rằng số lượng các chương trình mang tính bạo lực trên ti vi góp phần vào nạn bạo lực trong xã hội. Một vài người khác lại chỉ ra rằng các chương trình truyền hình cũng đóng góp vào sự ngoan ngoãn, vâng lời của đứa trẻ. Chúng ta có thể phải bỏ sung cả vấn đề tài chính vào trong danh sách này. Một nhà tâm lý học đưa ra 1 lý thuyết rằng những người mà xem tivi thường xuyên sẽ được trải nghiệm với rất nhiều vấn đề về thương mại, khiến họ mua sắm nhiều hơn và do đó món nợ của họ cũng tăng lên. Để kiểm tra niềm tin này, 1 mẫu ngẫu nhiên gồm 430 gia đình đã được khảo sát. Mỗi gia đình sẽ được hỏi về tổng số nợ và số giờ mà ti vi được bật lên mỗi tuần. Những dữ liệu này được lưu trong file XR18-79. Hãy biểu diễn thủ tục thống kê để giúp kiểm tra lý thuyết này.