

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2010

Các Phương Pháp Phân Tích Định Lượng

Đáp án Bài tập 4

LẤY MẪU VÀ ƯỚC LƯỢNG

Bài 1: (25 điểm)

Một mẫu có cỡ $n = 25$ được chọn ngẫu nhiên từ một tổng thể. Giả sử tổng thể có trị trung bình $\mu = 200$ và độ lệch chuẩn $\sigma = 40$.

- a. Hãy cho biết giá trị kỳ vọng và phương sai của trung bình mẫu \bar{x}

Giá trị kỳ vọng của trung bình mẫu $E(\bar{X}) = \mu = 200$

Phương sai của trung bình mẫu $\sigma_{\bar{X}}^2$:

$$\sigma_{\bar{X}}^2 = \frac{\sigma_x^2}{n} = \frac{40^2}{25} = 64$$

- b. Bây giờ, giả sử cỡ mẫu là 49. Hãy tính lại các yêu cầu trong câu a và cho nhận xét về sự giống nhau và khác nhau trong các kết quả của hai cỡ mẫu khác nhau. Giải thích.

Bây giờ $n = 49$

Giá trị kỳ vọng của trung bình mẫu $E(\bar{X}) = \mu = 200$

Phương sai của trung bình mẫu $\sigma_{\bar{X}}^2$:

$$\sigma_{\bar{X}}^2 = \frac{\sigma_x^2}{n} = \frac{40^2}{49} = 32.65$$

Nhận xét:

Đối với cỡ mẫu 25 và 49, giá trị kỳ vọng của trung bình mẫu là như nhau và bằng giá trị trung bình của tổng thể.

Với cỡ mẫu 49, phương sai của trung bình mẫu sẽ nhỏ hơn phương sai của trung bình mẫu đối với cỡ mẫu 25. Khi kích thước của mẫu ngẫu nhiên càng lớn, các trung bình mẫu có tính ổn định hơn nhờ vào sự triệt tiêu lẫn nhau của các quan sát riêng biệt. Nói cách khác, khi kích thước của mẫu ngẫu nhiên nhỏ, các giá trị rất lớn hoặc rất nhỏ trong mẫu sẽ ảnh hưởng mạnh lên các trung bình mẫu. Vì vậy mà mẫu ngẫu nhiên có kích thước mẫu càng lớn thì độ lệch chuẩn của phân phối mẫu càng nhỏ.

- c. Nếu Anh/Chị mong muốn độ lệch chuẩn của trung bình mẫu tối đa là 10. Hãy tính cỡ mẫu tối thiểu mà Anh/Chị cần phải lấy.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \Rightarrow \text{cỡ mẫu tối thiểu } n = \frac{\sigma_x^2}{\sigma_{\bar{x}}^2} = \frac{1600}{100} = 16$$

Bài 2: (25 điểm)

Vòng đời của nhãn hiệu bóng đèn Bình Minh có giá trị trung bình 1500 giờ và độ lệch chuẩn là 500 giờ. Cho phân phối xác suất của tổng thể là phân phối chuẩn. Giả sử bạn mua 16 bóng đèn của nhãn hiệu này và xem đây như là một mẫu ngẫu nhiên sản phẩm của nhà sản xuất này.

- a. Hãy tính giá trị trung bình của trung bình mẫu

$$\text{Giá trị kỳ vọng của trung bình mẫu } E(\bar{X}) = \mu = 1500 \text{ (giờ)}$$

- b. Hãy tính độ lệch chuẩn của trung bình mẫu.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{500}{4} = 125$$

$$\Rightarrow \sigma_{\bar{x}} = 125 \text{ (giờ)}$$

- c. Hãy tính xác suất, trung bình mà nói, 16 bóng đèn này có vòng đời thấp hơn 1200 giờ.

$$P(\bar{x} < 1200) = P(z < (1200 - 1500)/125)$$

$$= P(z < -2.4) = 0.5 - P(0 < z < 2.4) = 0.5 - 0.4918 = 0.0082$$

Bài 3: (35 điểm)

Để khuyến khích người dân mua hàng ở khu vực trung tâm thành phố, lãnh đạo thành phố dự định xây dựng một khu giữ xe công cộng nằm gần khu trung tâm mua sắm của thành phố. Thành phố dự định sẽ sử dụng phí giữ xe để bù đắp chi phí xây dựng. Khảo sát trong vòng 44 ngày người ta thấy chi phí trung bình thu được trong một ngày là \$126 với độ lệch chuẩn là \$15.

- a) Những giả thiết nào cần được đảm bảo để có thể sử dụng những số liệu thống kê cho việc suy diễn?

Các điều kiện cần phải được đảm bảo là lấy mẫu độc lập và các quan sát được chọn ngẫu nhiên.

- b) Hãy ước lượng khoảng tin cậy 90% cho doanh thu trung bình mà khu giữ xe này có thể tạo ra trong một ngày.

Với độ tin cậy 90%, $Z_{\alpha/2} = 1.645$

$$\bar{X} = 126 \pm 1.645 * \frac{15}{\sqrt{44}}$$

Khoảng tin cậy 90% là [\$122,28; \$129,72]

Dựa trên những thông tin từ mẫu, với độ tin cậy 90%, chúng ta đoan chắc rằng thu nhập trung bình của khu giữ xe công cộng sẽ nằm trong khoảng [\$122,28; \$129,72].

c) *Hãy giải thích 90% độ tin cậy nghĩa là gì trong trường hợp này?*

90% các khoảng tin cậy của tất cả các mẫu ngẫu nhiên sẽ chứa giá trị thật của thu nhập trung bình ngày của khu giữ xe.

d) *Những nhà tư vấn cho dự án này dự đoán doanh thu giữ xe trung bình là vào khoảng \$130 một ngày. Dựa trên khoảng tin cậy bạn hãy cho biết các nhà tư vấn có đúng không? Tại sao?*

Các nhà tư vấn đã dự báo chưa chính xác về thu nhập trung bình ngày của khu giữ xe vì giá trị dự báo của họ nằm ngoài khoảng tin cậy 90% mà chúng ta đã tính toán dựa trên thông tin của mẫu.

e) *Một người cho rằng thành phố nên sử dụng thông tin đã có để ước lượng khoảng tin cậy 95%. Các anh chị hãy cho biết những ưu nhược điểm của đề xuất này?*

$$\bar{X} = 126 \pm 1.96 * \frac{15}{\sqrt{44}}$$

Khoảng tin cậy 95% là [\$121,57; \$130,43].

Chúng ta thấy rằng khoảng tin cậy 95% chứa con số dự báo của các nhà tư vấn bởi vì nó rộng hơn khoảng tin cậy 90%. Điều này cho thấy khoảng tin cậy rộng hơn làm cho các con số dự báo ít bị sai hơn và do đó dự án dễ được chấp thuận hơn. Tuy nhiên, đề xuất này có nhược điểm là khoảng tin cậy rộng do đó việc xác định giá trị trung bình của tổng thể sẽ kém chính xác hơn.

f) *Tại sao khoảng tin cậy 95% là một trở ngại cho công tác hoạch định? Làm thế nào để chúng ta có thể xây dựng được một ước lượng khoảng phục vụ tốt hơn cho công tác hoạch định?*

Một khoảng tin cậy rộng làm cho việc xác định các thông tin cần thiết cho công tác hoạch định trở nên kém chính xác, làm cho chất lượng công tác hoạch định sẽ không cao.

Từ công thức tính sai số biên chúng ta thấy rằng với độ tin cậy không thay đổi sai số biên sẽ nhỏ hơn khi số quan sát trong mẫu tăng lên. Do đó, việc gia tăng cỡ mẫu sẽ làm giảm sai số ước lượng, phục vụ tốt hơn cho công tác hoạch định.

- g) Chúng ta cần thu thập thông tin trong bao nhiêu ngày để có được khoảng tin cậy 95% trong việc ước lượng giá trị trung bình đúng trong vòng \$3?

$$z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} = B$$
$$n = \frac{z_{\alpha/2}^2 * \sigma^2}{B^2} = \frac{1,96^2 * 15^2}{3^2} = 96$$

n = 96 ngày

Bài 4: (15 điểm)

Vào năm 2000, Bộ Giáo dục công bố một kết quả điều tra trên diện rộng về việc tốt nghiệp phổ thông trung học. Các nhà nghiên cứu đã liên hệ với hơn 25.000 người ở độ tuổi 24 để xem họ đã tốt nghiệp phổ thông trung học chưa. 84,9% trong 12.460 đàn ông và 88,1% trong 12.678 phụ nữ được hỏi trả lời rằng họ đã tốt nghiệp.

- a. Hãy xác định khoảng tin cậy 95% cho sự khác biệt về tỷ lệ tốt nghiệp giữa nam giới và nữ giới.

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} = (0,881 - 0,849) \pm 1,96 \sqrt{\frac{0,881 * 0,119}{12.678} + \frac{0,849 * 0,151}{12.460}}$$

Khoảng tin cậy 95% là [0,023;0,04] hay [2,3%;4%]

- b. Giải thích kết quả tính toán của bạn.

Sự khác biệt về tỷ lệ tốt nghiệp giữa phụ nữ và đàn ông nằm trong khoảng 0,023 đến 0,04 với độ tin cậy là 95%.

- c. Kết quả này có phải là bằng chứng xác thực giúp chúng ta kết luận rằng nữ giới có khả năng tốt nghiệp phổ thông trung học nhiều hơn nam giới không?

Kết quả tính toán về khoảng tin cậy 95% cho phép chúng ta kết luận rằng nữ giới có khả năng tốt nghiệp phổ thông trung học nhiều hơn nam giới vì khoảng tin cậy này không chứa giá trị 0.