

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2015

Các Phương Pháp Phân Tích Định Lượng

BÀI TẬP 5

KIỂM ĐỊNH GIẢ THUYẾT

Ngày Phát: Thứ ba 03/11/2015

Ngày Nộp: Thứ ba 10/11/2015

Bản in nộp trước 8h20 tại Hộp nộp bài tập trong phòng Lab.

Bản điện tử nộp trước 8h20 tại địa chỉ: <http://www.fetp.edu.vn/vn/tai-nguyen/hoc-vien-hien-tai/>

Câu 1: (50 điểm)

RVT là giống lúa mới chất lượng cao đang được Bộ Nông nghiệp và Phát triển nông thôn (Bộ NN&PTNT) khuyến cáo bà con nông dân sử dụng bởi giống lúa này có khả năng thích nghi rộng, chống chịu các bệnh phổ biến trên cây lúa (như rầy nâu, đạo ôn, khô vằn, bạc lá...) và có chất lượng gạo tốt đảm bảo các tiêu chuẩn xuất khẩu hiện hành. Theo công bố từ nhóm các nhà nghiên cứu đã phát triển ra giống lúa trên, năng suất của giống lúa RVT là khoảng 7,25 tấn/ha. Giống lúa trên đã được nông dân hưởng ứng và trồng thử nghiệm cả 3 miền trong vụ đông xuân vừa qua.

- a. Với góc độ là cơ quan quản lý Nhà nước, Bộ NN&PTNT muốn biết liệu năng suất lúa RVT trên thực tế trong vụ đông xuân vừa qua có phải là 7,25 tấn như thông tin các nhà nghiên cứu đã công bố hay không? Kết quả thu thập khoảng 100 quan sát về sản lượng thực tế ở cả 3 miền cho thấy năng suất trung bình là 7,20 tấn/ha và sai số chuẩn trong mẫu thu thập là 0,35 tấn/ha. Anh chị hãy phát biểu giả thuyết hợp lý và kiểm định giả thuyết trên với mức ý nghĩa 5%.

- Phát biểu giả thuyết:

- $H_0: \mu = \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}} \right)$
- $H_a: \mu \neq \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}} \right)$

- Kiểm định giả thuyết:

- $n = 100 \Rightarrow$ cỡ mẫu lớn $\Rightarrow$ dùng trị thống kê Z
- $\bar{x} = 7,20 \left(\frac{\text{tấn}}{\text{ha}} \right)$
- $s = 0,35 \left(\frac{\text{tấn}}{\text{ha}} \right)$

$$\text{Ta có: } Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{7,20 - 7,25}{0,35 / \sqrt{100}} = -1,43$$

$$\text{Kiểm định 2 phía: } Z_{\alpha/2} = Z_{2,5\%} = 1,96$$

Vì $|Z| = 1,43 < Z_{\alpha/2} = 1,96 \Rightarrow$ chưa đủ cơ sở để bác bỏ H_0 .

Như vậy, chưa đủ cơ sở để nói rằng sản lượng lúa TB vụ đông xuân khác với sản lượng lúa mà các nhà nghiên cứu đã công bố.

- b. Với độ tin cậy 95%, anh chị hãy cho biết khoảng ước lượng về năng suất lúa vụ đông xuân vừa qua tại khu vực miền Bắc? Biết rằng kết quả thu thập ngẫu nhiên 39 mẫu (39 quan sát) tại khu vực miền Bắc cho năng suất trung bình đạt 7,4 tấn/ha và sai số chuẩn trong mẫu thu thập là 0,3 tấn/ha. Anh chị kết luận được gì từ kết quả tính toán khi so sánh với công bố từ các nhà nghiên cứu trong trường hợp này?

- $n = 39 \Rightarrow$ cỡ mẫu lớn $\Rightarrow$ dùng trị thống kê Z
- $\bar{x} = 7,40$ (tấn/ha)
- $s = 0,30$ (tấn/ha)
- $Z_{\alpha/2} = Z_{2,5\%} = 1,96$

Vậy khoảng ước lượng về năng suất lúa vụ đông xuân tại miền Bắc là:

$$\mu \in \left[\bar{x} \pm Z_{\alpha/2} * \frac{s}{\sqrt{n}} \right] \Leftrightarrow \mu \in \left[7,40 \pm 1,96 * \frac{0,30}{\sqrt{39}} \right] \Leftrightarrow \mu \in [7,30 ; 7,49]$$

Với độ tin cậy 95%, khoảng ước lượng năng suất lúa TB tại miền Bắc biến thiên trong khoảng từ 7,30 đến 7,40 (tấn/ha), và như vậy cao hơn năng suất lúa ban đầu do các nhà nghiên cứu công bố là 7,25 (tấn/ha).

- c. Theo thông tin phản hồi từ một số nông dân ở khu vực miền Trung cho rằng, sản lượng thực tế khi trồng giống RVT là thấp hơn so với năng suất đã công bố của nhóm nghiên cứu khi dựa vào kết quả từ 25 mẫu (25 quan sát) thu thập ngẫu nhiên tại miền Trung cho thấy năng suất chỉ đạt 6,8 tấn/ha và sai số chuẩn trong mẫu là 0,3 tấn/ha. Anh chị hãy phát biểu giả thuyết hợp lý và kiểm định giả thuyết trên với độ tin cậy 95%. Kết luận của anh/chị đối với phản hồi của nông dân ở khu vực này là gì?

- Phát biểu giả thuyết:

- $H_0: \mu = \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}} \right)$
- $H_a: \mu < \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}} \right)$

- Kiểm định giả thuyết:

- $n = 25 \Rightarrow$ cỡ mẫu nhỏ $\Rightarrow$ dùng trị kiểm định t
- $\bar{x} = 6,80$ (tấn/ha)
- $s = 0,30$ (tấn/ha)

Ta có: $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{6,80 - 7,25}{\frac{0,30}{\sqrt{25}}} = -7,5$

Kiểm định 1 phía: $t_{\alpha; df=24} = t_{5\%; df=24} = 1,71$

Vì $|t| = 7,5 < t_{\alpha; df=24} = 1,71 \Rightarrow$ có đủ cơ sở để bác bỏ H_0 .

Vậy với độ tin cậy 95%, chúng ta hoàn toàn có cơ sở để nói rằng năng suất lúa tại miền Trung là thấp hơn 7,25 (tấn/ha). Điều này có thể do điều kiện đất đai tại miền Trung kém màu mỡ hơn các vùng khác.

- d. Trong khi đó, phản hồi từ các hộ dân ở Miền Nam lại cho rằng, năng suất của họ cao hơn năng suất nhóm nghiên cứu đã công bố. Kết quả thu thập mẫu của 36 mẫu (36 quan sát) ở khu vực này cho thấy, năng suất lúa trung bình là 7,4 tấn/ha và sai số chuẩn trong mẫu thu thập là 0,2 tấn/ha. Anh chị hãy đưa ra giả thuyết hợp lý và kiểm định giả thuyết trên với độ tin cậy 95%. Kết luận của anh/chị đối với phản hồi của nông dân ở khu vực này là gì?

- Phát biểu giả thuyết:

- $H_0: \mu = \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}}\right)$
- $H_a: \mu > \mu_0 = 7,25 \left(\frac{\text{tấn}}{\text{ha}}\right)$

- Kiểm định giả thuyết:

- $n = 36 \Rightarrow$ cỡ mẫu lớn $\Rightarrow$ dùng trị thống kê Z
- $\bar{x} = 7,40$ (tấn/ha)
- $s = 0,20$ (tấn/ha)

$$\text{Ta có: } Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{7,40 - 7,25}{0,20/\sqrt{36}} = 4,50$$

$$\text{Kiểm định 1 phía: } Z_\alpha = Z_{5\%} = 1,645$$

$$\text{Vì } |Z| = 4,50 > Z_\alpha = 1,645 \Rightarrow \text{có đủ cơ sở để bác bỏ } H_0.$$

Như vậy, với độ tin cậy 95% chúng ta hoàn toàn có cơ sở để nói rằng sản lượng lúa TB vụ đông xuân tại miền là cao hơn so với sản lượng lúa mà các nhà nghiên cứu đã công bố.

- e. Trước phản hồi của nông dân ở 02 khu vực trên, các nhà nghiên cứu đã đặt ra câu hỏi, liệu năng suất lúa ở khu vực miền Nam có cao hơn khu vực miền Trung hay không bởi xét về điều kiện tự nhiên thì đất đai ở khu vực miền Nam được đánh giá là màu mỡ hơn? Anh chị hãy đưa ra giả thuyết phù hợp để giúp các nhà nghiên cứu trả lời câu hỏi trên. Tiến hành kiểm định và đưa ra kết luận của anh/chị dựa vào các thông tin về mẫu ở 02 khu vực trên trong câu c và câu d?
- f. Hãy tính toán giá trị p-value trong mỗi trường hợp ở câu a, c, d, e. Kết luận của anh/chị từ kết quả tính toán và ý nghĩa của các giá trị p-value tính toán được?
- g. Viết báo cáo tổng hợp về năng suất lúa thực tế trong vụ đông xuân vừa qua một cách ngắn gọn để báo cáo lên Bộ trưởng (tối đa 300 từ)?

Câu 2: (30 điểm)

Trung tâm đào tạo lái xe A công bố tỷ lệ thí sinh đăng ký và được đào tạo tại trung tâm vượt qua kỳ thi và được cấp bằng là 85%. Kết quả kiểm tra ngẫu nhiên từ thanh tra giao thông trên 90 người đã tham gia học tại trung tâm A cho thấy có 70 người là vượt qua kỳ thi sát hạch để cấp bằng. Câu hỏi mà tổ thanh tra đặt ra là tỷ lệ thí sinh học tại trung tâm đào tạo lái xe A vượt qua kỳ thi sát hạch và được cấp bằng có thấp hơn con số mà trung tâm A đã công bố?

- Hãy phát biểu giả thuyết hợp lý để trả lời câu hỏi trên của tổ thanh tra?
- Với mức ý nghĩa 5%, kết luận của anh chị cho câu hỏi của tổ thanh tra?

- c. Xác suất anh/chị mắc sai lầm loại 1 nếu anh/chị bác bỏ thông tin mà trung tâm đào tạo A đã công bố và chấp nhận nghi ngờ của tổ thanh tra là đúng là bao nhiêu %?
- d. Trong khi đó, kết quả kiểm tra ngẫu nhiên 80 người đã tham gia đăng ký và được đào tạo tại trung tâm đào tạo B cho thấy 65 người có thể vượt qua kỳ thi sát hạch để cấp bằng. Câu hỏi là có sự khác biệt về tỷ lệ học viên vượt qua kỳ thi sát hạch và được cấp bằng ở 2 trung tâm đào tạo trên? Kết luận của anh chị từ kết quả tính toán?

Câu 3: (20 điểm)

Định giá đất để phục vụ cho công tác giải phóng mặt bằng là một vấn đề rất nóng trong thời gian gần đây sau khi Luật đất đai 2013 ra đời và có hiệu lực. Theo đó, mức giá đất đền bù phải được căn cứ theo giá trị thị trường. Giá trị thị trường thường được thẩm định độc lập bởi các đơn vị thẩm định giá tư nhân, sau đó sẽ được Sở tài chính (tại địa phương nơi tài sản định giá hiện hữu) thẩm tra và ra quyết định. Một số ý kiến cho rằng, mức giá thẩm tra của Sở tài chính thường thấp hơn mức giá do các công ty thẩm định giá tư nhân đề xuất. Kết quả một mẫu khảo sát gồm 8 khu đất đã được thu thập thông tin bao gồm cả giá đất do đơn vị thẩm định giá tư nhân đề xuất (gọi tắt là P_{Private}) và mức giá do Sở tài chính thẩm tra (gọi tắt là P_{Public}) như sau:

Khu đất	P_{Private} (tỷ đồng)	P_{Public} (tỷ đồng)
1	36.3	35.1
2	48.4	46.8
3	40.2	37.3
4	54.7	50.6
5	28.7	29.1
6	42.8	41.0
7	36.1	35.3
8	39.0	39.1

- a. Anh chị hãy phát biểu giả thuyết và kiểm định sự khác nhau giữa giá đất trong 2 nhóm trên (01 nhóm là bao gồm các giá đất do các công ty thẩm định giá tư nhân thực hiện và nhóm còn lại bao gồm các giá đất do Sở tài chính thẩm tra). Kết luận của anh/chị từ kết quả tính toán?
- b. Từ thông tin đề bài cung cấp, liệu anh/chị có đủ bằng chứng để nói rằng giá đất do Sở tài chính định giá thường THẤP HƠN giá đất do các công ty thẩm định giá tư nhân đề xuất? Nếu có, hãy chứng minh.

---HẾT---