

Chương trình Giảng dạy Kinh tế Fulbright

Học kỳ Thu năm 2014

Các Phương Pháp Phân Tích Định Lượng

BÀI TẬP 4

LẤY MẪU VÀ ƯỚC LƯỢNG THỐNG KÊ

Ngày Phát: Thứ Ba 21/10/2014

Ngày Nộp: Thứ Ba 28/10/2014

Bản in nộp lúc **8h20 sáng**, tại Hộp nộp bài tập trong phòng Lab

Bản điện tử gửi lên <http://intranet.fetp.edu.vn:81>

Bài 1: (20 điểm)

Một cửa hàng thương mại ghi nhận doanh thu trong 5 tháng như sau: 300; 350; 250; 300; 400 (triệu đồng). Giả sử doanh thu trong 5 tháng trên đây được xem như một tổng thể. Tiến hành chọn mẫu ngẫu nhiên 2 trong 5 tháng và tìm hiểu về doanh thu trung bình của cửa hàng trong 2 tháng.

- Tính tham số thống kê: trung bình, độ lệch chuẩn của tổng thể.
- Tìm phân phối xác suất của trung bình mẫu \bar{X} .
- Tính và nêu ý nghĩa các trị thống kê: trung bình, độ lệch chuẩn trực tiếp từ phân phối xác suất của \bar{X} .
- Chỉ ra và giải thích các mối quan hệ giữa trị thống kê và tham số thống kê nói trên.

Bài 2: (30 điểm)

Mục GIADAT trong tập tin dữ liệu *MPP07-521-P04V-Data.xls* trình bày số liệu khảo sát giá đất (triệu đồng/m²) trên 635 hộ gia đình ở thành phố Hồ Chí Minh. Coi đây là một tổng thể.

- Tính toán các giá trị thống kê mô tả của tập dữ liệu GIADAT.
- Chọn ngẫu nhiên một mẫu có cỡ bất kỳ từ tổng thể nêu trên, hãy dự đoán một cách định tính:
 - Giá trị kỳ vọng của mẫu này.
 - Độ phân tán của giá trị trung bình trong mẫu này xung quanh giá trị trung bình của tổng thể. Nêu rõ các giả định cần thiết để dự đoán của bạn có độ tin cậy cao.
- Hãy chọn ngẫu nhiên một mẫu có cỡ $n = 75$ từ tập dữ liệu nói trên, chỉ ra các giá trị:
 - Kỳ vọng của mẫu.
 - Khoảng cách từ kỳ vọng của mẫu tới kỳ vọng của tổng thể.
 - So sánh với dự đoán ở phần (b) và giải thích các quan sát thu được.
- Dùng Excel để chọn ra 25 mẫu ngẫu nhiên bất kỳ từ tập dữ liệu trên với cỡ mẫu tăng dần (theo cấp số cộng, công sai là 5) bắt đầu từ $n = 150$.
 - Với mỗi mẫu hãy lặp lại các yêu cầu trong phần (c).
 - Vẽ đồ thị của các giá trị vừa tìm được theo sự tăng dần của n .

- Mô tả xu thế thay đổi của các giá trị này theo n . Bạn rút ra được điều gì từ câu hỏi này?
- e. Giả sử rằng, số quan sát của biến GIADAT rất lớn (lớn hơn rất nhiều so với 635 quan sát) các giá trị trung bình và phương sai được cho là không đổi so với các giá trị trong phần (a). Với một mong muốn có độ lệch chuẩn của mẫu chọn từ tập dữ liệu này là 45 nghìn đồng/m², thì cỡ mẫu cần thiết là bao nhiêu?

Bài 3: (15 điểm)

Từ một mẫu tổng quát $W = (X_1, X_2)$, giả sử phương sai của X_1, X_2 là như nhau, chúng ta xem xét ước lượng trung bình của tổng thể μ như sau:

$$\bar{X} = aX_1 + bX_2, \text{ với } (a + b) = 1, a > 0, b > 0$$

- Chứng minh rằng \bar{X} là ước lượng không chệch của μ .
- Với điều kiện nào của a, b thì ước lượng \bar{X} hiệu quả nhất?

Bài 4: (20 điểm)

Chọn ngẫu nhiên 240 người dân trong khu phố để khảo sát thì thấy 160 người có tham gia bảo hiểm y tế (BHYT). Hãy trả lời các câu hỏi sau, nêu rõ cách tính toán và các lập luận:

- Ước lượng sai số chuẩn của phân phối mẫu.
- Với độ tin cậy 95%, tỉ lệ người dân trong khu phố tham gia BHYT là bao nhiêu?
- Giả định tỉ lệ người dân trong khu phố tham gia BHYT trên thực tế là 70%. Xác suất để tỉ lệ người dân tham gia BHYT trong mẫu thấp hơn 60% là bao nhiêu?
- Theo thông tin câu (c), để độ chính xác là 5% và độ tin cậy là 95% cần khảo sát bao nhiêu người dân trong khu phố?

Bài 5: (15 điểm)

Nghiên cứu tìm hiểu nguyên nhân vắng học của học viên với lí do phải chăm sóc con nhỏ, nhà trường đã chọn ra một mẫu ngẫu nhiên gồm 69 học viên có con nhỏ dưới 10 tuổi và thấy rằng thời gian vắng học trung bình là 2,25 giờ mỗi tháng với độ lệch chuẩn 2,09 giờ mỗi tháng. Song song đó nhà trường cũng chọn ra một mẫu ngẫu nhiên độc lập khác gồm 206 học viên không có con nhỏ dưới 10 tuổi và thấy thời gian vắng học trung bình là 1,69 giờ mỗi tháng, độ lệch chuẩn của mẫu là 1,91 giờ mỗi tháng.

- Xác định khoảng tin cậy 90% và khoảng tin cậy 99% cho sự khác biệt về thời gian vắng học trung bình giữa nhóm học viên có con nhỏ dưới 10 tuổi và nhóm học viên không có con nhỏ dưới 10 tuổi.
- Giải thích kết quả tính toán.
- Kết quả này có phải là bằng chứng xác thực giúp chúng ta kết luận rằng học viên có con nhỏ dưới 10 tuổi vắng học nhiều hơn học viên không có con nhỏ dưới 10 tuổi hay không?

Yêu cầu: Nộp kèm file tính toán thực hiện bằng excel.

--- Hết ---