

THỐNG KÊ MÁY TÍNH & ỨNG DỤNG – 18HCB

BÀI TẬP 4

---oOo---

Suy diễn thống kê (statistical inference) là quá trình rút ra kết luận cho toàn bộ tổng thể từ các quan sát trên mẫu (kích thước nhỏ, đại diện cho tổng thể). Suy diễn thống kê gồm các bước phân tích có định hướng hơn và sâu hơn trên dữ liệu, sau khi đã thực hiện *phân tích dữ liệu thăm dò* (exploratory data analysis). Hai bài toán quan trọng của suy diễn thống kê là xây dựng *khoảng tin cậy* (confidence interval) và *kiểm định giả thuyết* (hypotheses testing).

Hướng dẫn

Sinh viên thực hành theo tài liệu: “Khoảng tin cậy” và “Kiểm định thống kê” được đính kèm trên trang môn học.

Sinh viên đọc thêm lý thuyết trong tài liệu: Robin H. Lock et al. *Statistics: Unlocking the power of data*, Wiley, 2013.

Bài tập nộp

Sinh viên trình bày bài trong file Word (bao gồm code R, kết quả phân tích, ...) và nộp bài vào link nộp bài trên trang môn học.

Sinh viên tiếp tục việc phân tích dữ liệu trên dataset **SpeedDating** trong package **Lock5withR** ở Bài tập 3 bằng việc xây dựng khoảng tin cậy 95% cho các tham số và kiểm định (với mức ý nghĩa 95%) cho các nghi vấn trên các tham số quan tâm của các biến và cặp biến đã chọn ở Bài tập 3. Sinh viên cũng đưa ra các nhận xét, đánh giá, nhận định trên kết quả phân tích được.