

BÀI TẬP PHÂN TÍCH THỐNG KÊ NHIỀU CHIỀU
TUẦN 4-5: BIỂU ĐỒ TÁN XẠ - PHÂN BỐ CHUẨN NHIỀU CHIỀU -
KIỂM ĐỊNH GTTB MỘT MẪU, HAI MẪU

Câu 1. Sử dụng câu lệnh *set.seed(100)* và sinh mẫu ngẫu nhiên gồm 100 quan sát tuân theo phân phối chuẩn 3 chiều với giá trị trung bình $(0, 0, 1)'$ và ma trận hiệp phương sai

$$\begin{pmatrix} 1 & .8 & .6 \\ .8 & 1 & .8 \\ .6 & .8 & 1 \end{pmatrix}.$$

- a. Đặt tên các biến lần lượt là A , B và C . Thêm một biến vào mẫu trên, đặt tên D gồm 100 quan sát được sinh ngẫu nhiên biết nó tuân theo phân phối chuẩn với hai tham số 0.5 và 0.5.
- b. Tính giá trị trung bình mẫu, ma trận tương quan mẫu và ma trận hiệp phương sai mẫu.
- c. Vẽ biểu đồ phân tán của các biến và biểu đồ nhiệt của ma trận tương quan mẫu.
- d. Kiểm tra tính chuẩn 4–chiều.

Câu 2. Tập dữ liệu **Smartket** trong thư viện **ISLR** gồm 9 biến số được đo trong 1250 thời điểm từ năm 2001 đến 2005 của thị trường chứng khoán S&P. Xét vectơ X gồm các biến định lượng trong tập dữ liệu.

- a. Tính giá trị trung bình mẫu, ma trận tương quan mẫu và ma trận hiệp phương sai mẫu của X .
- b. Vẽ biểu đồ phân tán của các biến và biểu đồ nhiệt của ma trận tương quan mẫu của X .
- c. Kiểm tra tính chuẩn nhiều chiều của X .
- d. Có sự khác biệt giữa lợi nhuận trung bình một ngày trước và ba ngày trước ngày đang xét không?
- e. Lợi nhuận trung bình tại ngày đang xét và các ngày trước đó có thực sự khác 0' không?
- f. Kiểm định xem có sự khác biệt giữa năm 2004 và 2005 về lợi nhuận trung bình tại ngày đang xét và các ngày trước đó không?