

CHƯƠNG 9

PHÂN TÍCH PHƯƠNG SAI

9.1 PHÂN TÍCH PHƯƠNG SAI MỘT YẾU TỐ:

PHÂN TÍCH PHƯƠNG SAI MỘT YẾU TỐ LÀ PHÂN TÍCH ẢNH HƯỞNG CỦA MỘT YẾU TỐ NGUYÊN NHÂN (ĐỊNH TÍNH) ẢNH HƯỞNG ĐẾN MỘT YẾU TỐ KẾT QUẢ (ĐỊNH LƯỢNG) ĐANG NGHIÊN CỨU.

9.1.1 TRƯỜNG HỢP k TỔNG THỂ CÓ PHÂN PHỐI CHUẨN VÀ PHƯƠNG SAI BẰNG NHAU

GỌI n_1, n_2, \dots, n_k LÀ SỐ QUAN SÁT TỪ k TỔNG THỂ KHÁC NHAU CÓ PHÂN PHỐI CHUẨN.

$\mu_1, \mu_2, \dots, \mu_k$ LÀ TRUNG BÌNH CỦA CÁC TỔNG THỂ. MÔ HÌNH PHÂN TÍCH PHƯƠNG SAI MỘT YẾU TỐ MÔ TẢ DƯỚI DẠNG KIỂM ĐỊNH GIẢ THUYẾT NHƯ SAU:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

TA THỰC HIỆN CÁC BƯỚC SAU:

BƯỚC 1: TÍNH CÁC TRUNG BÌNH MẪU

BẢNG SỐ LIỆU TỔNG QUÁT THỰC HIỆN PHÂN TÍCH PHƯƠNG SAI:

TỔNG THỂ			
1	2	...	k
x_{11}	x_{21}	...	x_{k1}
x_{12}	x_{22}	...	x_{k1}
...
x_{1n_1}	x_{2n_2}	...	x_{kn_k}

TÍNH TRUNG BÌNH MẪU $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$:

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \quad (i = 1, 2, \dots, k)$$

VÀ TRUNG BÌNH CHUNG CỦA k MẪU:

$$\bar{x} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{\sum_{i=1}^k n_i}$$

BƯỚC 2: TÍNH CÁC TỔNG CÁC ĐỘ LỆCH BÌNH PHƯƠNG

- **TỔNG CÁC ĐỘ LỆCH BÌNH PHƯƠNG TRONG NỘI BỘ**

NHÓM SSW:

TỔNG CÁC ĐỘ LỆCH BÌNH PHƯƠNG CỦA TỪNG NHÓM ĐƯỢC TÍNH THEO CÔNG THỨC:

$$\text{NHÓM 1: } SS_1 = \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2$$

$$\text{NHÓM 2: } SS_2 = \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2$$

.....

$$SSW = SS_1 + SS_2 + \dots + SS_k$$

$$\text{Hay : } SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

- TỔNG CÁC ĐỘ LỆCH BÌNH PHƯƠNG GIỮA CÁC NHÓM SSG :

$$SSG = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

- TỔNG CÁC ĐỘ LỆCH BÌNH PHƯƠNG TOÀN BỘ SST:

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

CÓ THỂ DỄ DÀNG CHỨNG MINH:

$$SST = SSW + SSG$$

BƯỚC 3: TÍNH CÁC PHƯƠNG SAI (TRUNG BÌNH CỦA CÁC ĐỘ LỆCH BÌNH PHƯƠNG)

- TÍNH PHƯƠNG SAI TRONG NỘI BỘ NHÓM

MSW:
$$MSW = \frac{SSW}{n - k}$$

- TÍNH PHƯƠNG SAI GIỮA CÁC NHÓM MSG:

$$MSG = \frac{SSG}{k - 1}$$

BƯỚC 4: KIỂM ĐỊNH GIẢ THUYẾT

TÍNH:
$$F = \frac{MSG}{MSW}$$

NẾU : $F > F_{k-1, n-k, \alpha}$ TA BÁC BỎ H_0

9.1.2 PHÂN TÍCH SÂU ANOVA

TRƯỜNG HỢP BÁC BỎ GIẢ THUYẾT H_0 , NGHĨA LÀ TRUNG BÌNH CỦA CÁC TỔNG THỂ KHÔNG BẰNG NHAU. VÌ VẬY, VẤN ĐỀ TIẾP THEO LÀ PHÂN TÍCH SÂU HƠN ĐỂ XÁC ĐỊNH NHÓM (TỔNG THỂ) NÀO KHÁC NHÓM NÀO.

PHƯƠNG PHÁP TUKEY:

NẾU CÓ k NHÓM THÌ SỐ LƯỢNG CẶP LÀ:

$$C_k^2 = \frac{k!}{2!(k-2)!} = \frac{k(k-1)}{2}$$

VÍ DỤ: TA CÓ $K = 3$, THÌ SỐ CẶP SO SÁNH TRONG KIỂM ĐỊNH LÀ 3.

$$C_3^2 = \frac{3!}{2!(3-2)!} = 3$$

CÁC GIẢ THUYẾT CẦN KIỂM ĐỊNH LÀ:

1. $H_0: \mu_1 = \mu_2$	2. $H_0: \mu_2 = \mu_3$	3. $H_0: \mu_1 = \mu_3$
$H_1: \mu_1 \neq \mu_2$	$H_1: \mu_2 \neq \mu_3$	$H_1: \mu_1 \neq \mu_3$

GIÁ TRỊ GIỚI HẠN TUKEY ĐƯỢC TÍNH THEO

CÔNG THỨC:
$$T = q_{\alpha, k, n-k} \sqrt{\frac{MSW}{n_i}}$$

TRONG ĐÓ:

- $q_{\alpha, k, n-k}$ LÀ GIÁ TRỊ TRA BẢNG PHÂN PHỐI KIỂM ĐỊNH TUKEY (STUDENTIZED RANGE DISTRIBUTION)
- n LÀ TỔNG SỐ QUAN SÁT MẪU ($n = \sum n_i$)
- MSW LÀ PHƯƠNG SAI TRONG NỘI BỘ NHÓM
- n_i LÀ SỐ QUAN SÁT TRONG 1 NHÓM (TỔNG THỂ), TRONG TRƯỜNG HỢP MỖI NHÓM CÓ SỐ QUAN SÁT n_i KHÁC NHAU, SỬ DỤNG GIÁ TRỊ n_i NHỎ NHẤT

TIÊU CHUẨN QUYẾT ĐỊNH LÀ BÁC BỎ
GIẢ THUYẾT H_0 KHI ĐỘ LỆCH TUYỆT
ĐỐI GIỮA CÁC CẶP TRUNG BÌNH MẪU
LỚN HƠN HAY BẰNG T.

9.1.3 TRƯỜNG HỢP CÁC TỔNG THỂ ĐƯỢC GIẢ ĐỊNH CÓ PHÂN PHỐI BẤT KỲ (PHƯƠNG PHÁP PHI THAM SỐ)

TRONG TRƯỜNG HỢP NÀY TA CÓ THỂ CHUYỂN ĐỔI DỮ LIỆU YẾU TỐ KẾT QUẢ TỪ DẠNG ĐỊNH LƯỢNG VỀ DẠNG ĐỊNH TÍNH (DỮ LIỆU THỨ BẬC) VÀ ÁP DỤNG MỘT KIỂM ĐỊNH PHI THAM SỐ PHÙ HỢP LÀ **KRUSKAL - WALLIS**.

GIẢ SỬ RẰNG CHÚNG TA CÓ CÁC MẪU NGẪU
NHIÊN ĐỘC LẬP GỒM n_1, n_2, \dots, n_k QUAN SÁT
TỪ k TỔNG THỂ CÓ PHÂN PHỐI BẤT KỲ. TA SỬ
DỤNG KIỂM ĐỊNH KRUSKAL – WALLIS BẰNG
CÁCH XẾP HẠNG CÁC QUAN SÁT MẪU. MẶC DÙ
SỐ QUAN SÁT CỦA n_k MẪU LÀ KHÁC NHAU
NHƯNG KHI XẾP HẠNG THÌ ĐƯỢC SẮP XẾP MỘT
CÁCH LIÊN TỤC TỪ NHỎ ĐẾN LỚN, NẾU GIÁ TRỊ
QUAN SÁT TRÙNG NHAU THÌ HẠNG GIỐNG NHAU
BẰNG CÁCH DÙNG SỐ TRUNG BÌNH CỘNG CÁC
HẠNG CỦA CHÚNG ĐỂ CHIA ĐỀU.

ĐẶT $\mathbf{n} = \mathbf{n}_1 + \mathbf{n}_2 + \dots + \mathbf{n}_k$ LÀ TỔNG CÁC QUAN SÁT THUỘC CÁC MẪU, VÀ $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_k$ LÀ TỔNG CỦA CÁC HẠNG Ở TỪNG MẪU ĐƯỢC XẾP THEO THỨ TỰ CỦA K MẪU. KIỂM ĐỊNH GIẢ THUYẾT Ở MỨC Ý NGHĨA α CHO TRƯỜNG HỢP NÀY LÀ:

$\mathbf{H}_0 : \mu_1 = \mu_2 = \dots = \mu_k$: TRUNG BÌNH CỦA K TỔNG THỂ ĐỀU BẰNG NHAU. Ở ĐÂY TA SỬ DỤNG BIẾN W THAY CHO TỈ SỐ F TRONG PHẦN TÍNH TOÁN GIÁ TRỊ KIỂM ĐỊNH.

$$\mathbf{W} = \frac{12}{\mathbf{n}(\mathbf{n} + 1)} \sum_{i=1}^k \frac{\mathbf{R}_i^2}{\mathbf{n}_i} - 3(\mathbf{n} + 1)$$

GIẢ THUYẾT \mathbf{H}_0 BỊ BÁC BỎ KHI:

$$\mathbf{W} > \chi_{k-1, \alpha}^2$$

KHI GIẢ THUYẾT VỀ TRUNG BÌNH CỦA K TỔNG THỂ GIỐNG NHAU BỊ BÁC BỎ. TA DÙNG PHƯƠNG PHÁP SO SÁNH TƯƠNG TỰ NHƯ PHƯƠNG PHÁP TUKEY TRONG PHẦN TRƯỚC. SAU ĐÂY LÀ TÓM TẮT CÁC BƯỚC THỰC HIỆN:

BƯỚC 1: TRƯỚC HẾT CHÚNG TA TÍNH HẠNG TRUNG BÌNH CHO TỪNG NHÓM MUỐN SO SÁNH THEO CÔNG THỨC TỔNG QUÁT SAU :

$$\bar{\mathbf{R}}_i = \frac{\mathbf{R}_i}{n_i}$$

BƯỚC 2: TIẾP THEO CHÚNG TA TÍNH CHÊNH LỆCH VỀ HẠNG TRUNG BÌNH GIỮA 2 NHÓM CẦN SO SÁNH

$$\mathbf{D}_{ij} = \left| \bar{\mathbf{R}}_i - \bar{\mathbf{R}}_j \right|$$

D ĐƯỢC COI NHƯ GIÁ TRỊ ĐỂ KIỂM ĐỊNH GIẢ THUYẾT VỀ SỰ BẰNG NHAU CỦA TRUNG BÌNH HAI TỔNG THỂ **i** VÀ **j**

BƯỚC 3 : TÍNH GIÁ TRỊ GIỚI HẠN C_K THEO CÔNG THỨC:

$$C_K = \sqrt{(\chi_{k-1, \alpha}^2) \left(\frac{n(n+1)}{12} \right) \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

TRONG ĐÓ $\chi_{k-1, \alpha}^2$ LÀ GIÁ TRỊ ĐÃ SỬ DỤNG KHI THỰC HIỆN KIỂM ĐỊNH KRUSKAL – WALLIS TRONG PHẦN TRƯỚC.

BƯỚC 4 : NGUYÊN TẮC QUYẾT ĐỊNH : BÁC BỎ GIẢ THUYẾT H_0 VỀ SỰ BẰNG NHAU CỦA HAI TRUNG BÌNH TỔNG THỂ KHI $D > C_K$