非加法性에 대한 Tukey의 統計量에 관하여

白 雲 鵬*

1. 緖 論

A, B 두 要因의 영향을 받고 있다고 생각되는 rc 개 測定値가 있고 이것이 다음과 같이 $r \times c$ 二元分類表로 整理되었다고 하자.

〔表 1.1〕 $r \times c$ 二元分類表

要因 A	要因 B				計	平均
	B_0	B_1	·····	B_{c-1}		
A_0	y_{00}	y_{01}	·····	$y_{0,c-1}$	y_0	$\bar{y}_0$
A_1	y_{10}	y_{11}	·····	$y_{1,c-1}$	y_1	$\bar{y}_1$
⋮	⋮	⋮		⋮	⋮	⋮
A_{r-1}	$y_{r-1,0}$	$y_{r-1,1}$	·····	$y_{r-1,c-1}$	y_{r-1}	$\bar{y}_{r-1}$
計	$y_{\cdot 0}$	$y_{\cdot 1}$	·····	$y_{\cdot,c-1}$	$y_{\cdot\cdot}$	
平均	$\bar{y}_{\cdot 0}$	$\bar{y}_{\cdot 1}$	·····	$\bar{y}_{\cdot,c-1}$		$\bar{y}_{\cdot\cdot}$

여기에서

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

와 같은 加法模型을 생각한다. 그리고 ϵ_{ij} 는 殘餘項으로써 平均이 0, 分散이 σ^2 인 正規分布를 한다고 假定하는 것이 보통이다. 또 이것은 母數模型인 경우 $E(y_{ij}) = \mu + \alpha_i + \beta_j$, $v(y_{ij}) = \sigma^2$ 임을 意味하는 것으로 된다. 그러나 資料에 따라서는 위에서와 같은 加法的 模型을 適用한다는 것이 適當하지 못한 경우가 있다.

加法模型이 適合하지 않다고 생각 될 경우, 이것을 解決하는 方法으로 測定値의 變數變換

*高麗大學校 敎授

을 생각할 수있다. 투키(J.W. Tukey)[6],[7]는 (i) 變數變換이 必要하다면 이와같은 決定을 돕기 위하여, (ii) 適切한 變數變換을 구하기 위하여, (iii) 加法性을 갖도록 變數變換이 되었나 의여부를 알아보는 方法을 생각하였다.

스네데코(G.W. Snedecor)와 코크란(W.G. Cochran)[5] (1967)은 이것을 다음과 같이 說明하고있다. 變數變換으로는 $x = y^p$와 같은 形式을 생각한다. $p = \frac{1}{2}$일 때는 平方根變換을, $p = -1$일 때는 逆數變換을 나타내며 $p=0$일 때는 對數變換이 必要하다는 것을 나타낸다. $y_{ij} = x_{ij}^{\frac{1}{p}}$로 놓고 이것을 테일러(Taylor) 級數로 展開한 다음에 $\bar{x}.. \doteqdot \bar{y}^p$와 같은 關係를 사용하여 y 尺度에서의 非加法性을 나타내는 첫 項을 다음과 같이 近似的으로 구한다.

$$\frac{1-p}{\bar{y}..}(\bar{y}_{i\cdot}-\bar{y}..)(\bar{y}._j-\bar{y}..) \tag{1.1}$$

여기에서 결국 다음과 같은 結論을 얻는다. 첫째로 y 尺度에 의한 資料에서 위와 같은 非加法性이 나타날 경우 加法的 模型을 適合시켜 얻은 값 $\bar{y}_{ij}$에 대한 偏差 $y_{ij}-\hat{y}_{ij}$는 變量 $(\bar{y}_{0\cdot}-\bar{y}..)(\bar{y}._j-\bar{y}..)$에 대한 一次回歸로써 나타낼 수 있다는 것이며, 둘째로 이때의 回歸係數를 B로 表示하면 이것이 $(1-p)/\bar{y}..$의 推定値로 되는 것이며 따라서 p의 推定値는 $(1_0-B\bar{y}..)$에 의하여 구해 진다는 것이다. 그러므로 투키의 非加法性의 檢定은 B의 값이 0이라는 것을 檢定하는 것으로 된다. 이때 투키의 非加法性을 나타내는 統計量 N은 回歸係數 B 값에서 다음과 같이 얻어진다.

$$N = \frac{[\sum\sum(\bar{y}_{i\cdot}-\bar{y}..)(\bar{y}._j-\bar{y}..)y_{ij}]^2}{\sum(\bar{y}_{i\cdot}-\bar{y}..)^2 \sum(\bar{y}._j-\bar{y}..)^2} \tag{1.2}$$

그러므로 加法模型에서 自由度$(r-1)(c-1)$의 殘餘平方合을 R로 表示할 때 $E=R-N$은 N과는 獨立이며 自由度 $(r-1)(c-1)-1$ 인 平方合으로 된다. 그러므로 N/Ve (단 $Ve=\frac{E}{(r-1)(c-1)-1}$)에 의한 F-檢定을 할 수 있는 것이다.

라오(C.R. Rao)[3]는 數學的 模型으로써

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + \lambda\alpha_i\beta_j$$

를 생각하여 λ의 推定値로써 위의 回歸係數 B와 같은 값을 얻고 있다. 그러나 그레이빌(F.A.Graybill)[1]의 模型設定과 說明方法에는 納得하기 어려운 點이 있는 것 같다. 그레이빌은 交互作用項이 包含된 模型으로 $E(y_{ij})=\mu+\alpha_i+\beta_j+\alpha\beta_{ij}$를 設定하고 투키의 非加法性에 관한 統計量을 모든 i, j 에 관해서 $\alpha\beta_{ij}=0$이라는 歸無假說을 檢定하는데 利用하고 있다. 그러나 위에서 모든 i, j에 관해서 $\alpha\beta_{ij}=0$으로 되는 것이 N/Ve가 自由度 $(1,(v-1)(c-1)-1)$인 F-分布를 하는데 必要한 條件은 아닌 것이다.

本論文에서는 加法模型에 의하여 求해지는 自由度 $(r-1)(c-1)$ 인 殘餘平方合에서 自由

度1인 非加法性에 관한 統計量을 分離해 내는 생각을 一般化하여 殘餘平方合을 自由度 1인 $(r-1)(c-1)$ 개의 平方合으로 分割한다. 그리고 투키의 非加法性에 관한 統計量은 二元分類表의 周邊平均値를 利用한 自由度 1인 한 交互作用(A의 一次成分 $\times$ B의 二次成分)을 나타내는 平方合임을 分明히 하는데 本論文의 目的이 있다.

2. 二元分類表의 周邊平均値를 利用한 殘餘平方合의 直交分割法

〔表 1.1〕과 같은 二元分類表에 대해서 交互作用이 存在하는 模型

$$y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ij} \tag{2.1}$$

를 생각할 경우 $\alpha\beta_{ij}$와 ϵ_{ij}를 서로 分類해서 推定할 수는 없다. 그러나 여기에서 생각할 수 있는 것은 交互作用과 誤差가 混同되어 있는 自由度 $(r-1)(c-1)$의 殘餘平方合을 自由度 1인 平方合의 合으로 分割하고 그 크기를 檢討하는 方法이다.

模型(2.1)에서 母數 α_i, β_j, 그리고 $\alpha\beta_{ij}$를 다음과 같이 變換할 수가 있다.

$$\begin{aligned} \alpha_i &= \sum_{m=0}^{r-1} a_m p_m(i) \\ \beta_j &= \sum_{n=0}^{j-1} b_n q_n(j) \\ \alpha\beta_{ij} &= \sum_{m=0}^{r-1}\sum_{n=0}^{c-1} c_{mn} p_m(i) q_n(j) \end{aligned} \tag{2.2}$$

여기에서 a_m, b_n 그리고 c_{mn}은 變換後에 우리가 推定하여야 할 母數이고 $p_m(i)$는 次數가 m 以下의 $(i-\bar{i})$에 관한 多項式이며 $q_n(j)$는 次數 n 以下인 $(j-\bar{j})$에 관한 多項式으로서 다음과 같은 條件을 滿足하여야 한다.

모든 i, j에 대하여

$$p_0(i) = \frac{1}{\sqrt{r}}$$

$$q_0(j) = \frac{1}{\sqrt{c}}$$

$$\sum_{i=0}^{r-1} p_m(i)\, p_{m'}(i) = \begin{cases} 1 & m = m' \text{ 일 때} \\ 0 & m \neq m' \text{ 일 때} \end{cases}$$

$$\sum_{j=0}^{j-1} q_n(i)\, q_{n'}(j) = \begin{cases} 1 & n = n' \text{ 일 때} \\ 0 & n \neq n' \text{ 일 때} \end{cases}$$

이와같은 方法에 의하여 非直交多元 分類表의 分析을 論議한 것이 筆者와 훼더러(W.T. Federer)〔2〕이다. 그러나 이 方法은 質的인 要因에 대해서는 그 水準設定이 任意的인 것으로 되므로 다음과 같은 다른 方法을 생각하지 않으면 안된다. 그런데 이와같은 경우 水準數 i, j 를 使用한 直交變換 대신에 二元分類表의 周邊平均値 $\bar{y}_{i\cdot}$, $\bar{y}_{\cdot j}$ 를 利用하여 다음과 같은 變換을 생각할 수 있을 것이다.

$$\alpha_i = \sum_{m=0}^{r-1} a_m\, p_m(\bar{y}_{i\cdot})$$

$$\beta_j = \sum_{n=0}^{c-1} b_n\, q_n(\bar{y}_{\cdot j}) \qquad (2.3)$$

$$\alpha\beta_{ij} = \sum_{m=0}^{r-1}\sum_{n=0}^{c-1} c_{mn}\, p_m(\bar{y}_{i\cdot}) q_n(\bar{y}_{\cdot j})$$

여기에서 a_m, b_n 그리고 c_{mn} 은 먼저와 같이 우리가 推定하여야 할 母數이고 $p_m(\bar{y}_{i\cdot})$는 $(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})$에 관해서 m 次以下의 多項式이며 $q_n(\bar{y}_{\cdot j})$ 는 $(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})$에 관한 n 次 以下의 多項式이다. 그리고 두 多項式은 다음과 같은 條件을 滿足하여야 한다.

모든 i, j에 대하여

$$p_0(\bar{y}_{i\cdot}) = \frac{1}{\sqrt{r}}$$

$$q_0(\bar{y}_{\cdot j}) = \frac{1}{\sqrt{c}}$$

이고

$$\sum_{i=0}^{r-1} p_m(\bar{y}_{i\cdot})\, p_{m'}(\bar{y}_{i\cdot}) = \begin{cases} 1 & m = m' \text{ 일 때} \\ 0 & m \neq m' \text{ 일 때} \end{cases}$$

$$\sum_{j=0}^{c-1} q_n(\bar{y}_{\cdot j})\, q_{n'}(\bar{y}_{\cdot j}) = \begin{cases} 1 & n = n' \text{ 일 때} \\ 0 & n \neq n' \text{ 일 때} \end{cases}$$

그런데 이러한 $p_m(\bar{y}_{i\cdot})$, $q_n(\bar{y}_{\cdot j})$는 롭슨(D.S. Robson)〔4〕에 따라서 구할 수가 있다.

위의 條件에 따라서 $m \geq 1$, $n \geq 1$일 때 $\sum_i p_m(\bar{y}_{i\cdot}) = 0$, $\sum_j q_n(\bar{y}_{\cdot j}) = 0$ 와 같이 되므로 c_{mn} 의 推定値는 다음과 같이 구해진다.

$$C_{mn} = \boldsymbol{w}_{mn}'\boldsymbol{y}$$

단, $\boldsymbol{w}_{mn} = [p_m(\bar{y}_{0\cdot})\, q_n(\bar{y}_{\cdot 0}),\ p_m(\bar{y}_{0\cdot})\, q_n(\bar{y}_{\cdot 1}), \cdots, p_m(\bar{y}_{0\cdot}) q_n(\bar{y}_{\cdot,c-1}), \cdots, p_m(\bar{y}_{1\cdot}) q_n(\bar{y}_{\cdot 0}),$
$\cdots\cdots, p_m(\bar{y}_{r-1,\cdot}) q_n(\bar{y}_{\cdot,c-1})]'$

$$\boldsymbol{y}=[y_{00},\ y_{01},\cdots\cdots,y_{0,c-1},\ y_{10},\cdots\cdots,y_{r-1,c-1}]'$$

여기에서 $m\neq 0,\ n\neq 0$일 때는 $\mathbf{1}'\boldsymbol{w}_{mn}=0$ (단, $\mathbf{1}$은 모든 元素가 1인 $rc\times 1$ 벡터)이고 $\boldsymbol{w}_{mn}'\boldsymbol{w}_{mn}=1$와 같은 特性을 가지고 있다는 것을 쉽게 理解할 수가 있다. 따라서 $\hat{c}_{mn}$에 관한 다음과 같은 自由度 1인 平方合이 구해진다.

$$SSc_{mn}=[\sum\sum p_m(\bar{y}_{i\cdot})q_n(\bar{y}_{\cdot j})y_{ij}]^2$$

이것은 二元分類表[1.1]에서 주어진 周邊平均値를 基準으로 한 A, B 두 要因間의 自由度 1인 交互作用(A의 m 次成分 × B의 n 次成分)에 관한 平方合이다.

그런데 롭슨[4]에 의하면 $m=1$, $n=1$ 일 때

$$p_1(\bar{y}_{i\cdot})=\frac{\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}}{\sqrt{\sum(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2}}$$

$$q_1(\bar{y}_{\cdot j})=\frac{\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot}}{\sqrt{\sum(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2}}$$

와 같이 되므로

$$SSc_{11}=\frac{[\sum^i\sum^j(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{ij}]^2}{\sum^i(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2\sum^j(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2}$$

를 얻는다.

이것이 다름아닌 투키의 非加法에 관한 統計量(1.2)과 똑같은 것이다. 이와 같이 우리는 모든 m과 n 값에 대한 서로 獨立인 SSc_{mn}을 구할 수 있을 것이며 이들을 서로 比較 檢討할 수 있을 것이다. 例를 들면 SSc_{12}, SSc_{21}을 구한 다음에 $E=R-SSc_{11}-SSc_{12}-SSc_{21}$와 같이 殘餘平方合을 구하여 ($A$의 一次成分) × ($B$의 二次成分), ($A$의 二次成分) × ($B$의 一成 次分)에 대한 F-檢定을 할 수 있을 것이며 이러한 檢定은 模型의 非加法性에 대한 좀더 詳細한 情報를 얻는데 도움이 될 것이다.

參考 直交多項式을 구하는 롭슨의 方法

最小自乘法에 의한 回歸方程式

$$z_i=a_0{}^*+a_1{}^*y_i+\cdots\cdots+a_r{}^*y_i,\quad i=1,2,\cdots,n>r$$

는 다음과 같이 表現될 수 있다.

$$z_i=a_0p_0(y_i)+a_1p_1+\cdots+a_rp_r(y_i),\ i=1,2,\cdots,\ n>r \tag{1}$$

단 $p_m(y_i)$는 y_i 에 관한 m次 多項式이고 다음과 같은 條件을 滿足한다.

$$\sum_i p_m(y_i)p_{m'}(y_i)=\begin{cases}1 & m=m'\text{ 일 때}\\ 0 & m\neq m'\text{ 일 때}\end{cases} \tag{2}$$

式(1)에 있어서 最小自乘法에 의한 a_m은 條件(2)가 滿足될 경우 測定値 $z_1, z_2, \cdots, z_n$의 線型式으로서 다음과 같이 구해진다.

$$a_m=\sum_{i=1}^{n} z_i p_m(y_i) \tag{3}$$

한편 만일 $i=1,2,\cdots,n$에 대하여 $z_i=y_i^m$ 일 경우는 $0\leq m\leq n-1$ 을 滿足하는 모든 m에 대하여 最小自乘法에 의한 m次 多項式은 誤差없이 完全히 適合된다. 따라서 이 때 $0\leq m\leq n-1$ 에 대하여 다음과 같은 關係式이 成立한다.

$$y_i^m=[\sum_{j=1}^{n} y_j^m p_0(y_j)]p_0(y_i)+\cdots\cdots+[\sum_{j=1}^{n} y_j^m p_m(y_j)]p_m(y_i) \tag{4}$$

이 式에서

$$[\sum_{j=1}^{n} y_j^m p_m(y_j)]p_m(y_i)=y_i^m-\sum_{s=0}^{n-1} p_s(y_i)\sum_{j=1}^{n} y_j^m p_s(y_j) \tag{5}$$

(5)의 兩邊을 제곱하고 $i=1,2,\cdots,n$에 걸쳐서 合하면 $\sum_{i=1}^{n}[p_m(y_i)]^2=1$ 이므로) 다음과 같이 된다.

$$[\sum_{j=1} y_j^m p_m(y_j)]^2=\sum_{i=1}^{n}[y_i^m-\sum_{s=0}^{r-1} p_s(y_i)\sum_{j=1}^{n} y_j^m p_s(y_i)]^2 \tag{6}$$

이 값을 k_m^2으로 놓기로 한다. 그러면 (5), (6) 式에서 다음과 같은 關係式을 얻게 된다.

$$p_m(y_i)=\frac{1}{k_m}[y_i^m-\sum_{s}^{r-1} p_s(y_i)\sum_{j=1}^{n} y_i^m\ p_s(y_i)]$$

여기에서 몇 개의 直交多項式 $p_m(y_i)$을 구하면 다음과 같이 된다.

$m=0$ 일 때:

$$k_0p_0(y_i)=1 \text{ 따라서 } p_0(y_i)=\frac{1}{\sqrt{n}}$$

$m=1$일 때 : p_0 값을 利用하여

$$k_1p_1(y_i)=y_i-p_0(y_i)\sum_j y_jp_0(y_j)$$

$$=y_i-n^{-\frac{1}{2}}\sum_j y_jn^{-\frac{1}{2}}=y_i-\bar{y}$$

따라서

$$p_1(y_i)=\frac{y_i-\bar{y}}{\sqrt{\sum_j (y_j-\bar{y})^2}}$$

$m = 2$ 일 때 : p_0, p_1 값을 利用하여

$$k_2 p_2(y_i) = y_i^2 - p_0(y_i)\sum_j y_j^2 p_0(y_j) - p_1(y_i)\sum_j y_j^2 p_1(y_j)$$

$$= y_i^2 - \frac{1}{n}\sum_j y_j^2 - (y_i - \bar{y})\frac{\sum y_j^2(y_j - \bar{y})}{\sum (y_j - \bar{y})^2}$$

이와 같이 모든 m 값에 대한 $p_m(y_i)$를 구할 수가 있다.

參 考 文 獻

[1] Graybill, F. A., *An Introduction to Linear Statistical Models*, Vol. 1, McGraw-Hill, 1961, 324~332.

[2] Paik, U. B. and Federer, W. T., "Analysis of Nonorthogonal n-way Classifications," *Annals of Statistics* 2, 1974, 1000~1021.

[3] Rao, C. R., *Linear Statistical Inference and Its Applications*, John Wiley, 1965, 207~209.

[4] Robson, D. S., "A Simple Method for Constructing Orthogonal Polynomials When the Independent Variable Is Unequally Spaced," *Biometrics* 15 (1959), 187~191.

[5] Snedecor, G. W. and Cochran, W. G., *Statistical Methods*, The Iowa State University Press, 1967, 331~334.

[6] Tukey, J. W, "One Degree of Freedom for Non-Addtivity," *Biometrics* 5 (1949), 232~242.

[7] Tukey, J. W., *Queries in Biometrics* 11 (1955), 111.

SUMMARY

On Tukey's Statistics due to Nonadditivity

U. B. Paik*

In the two-way classification model

$$y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ij},$$

we may use the following orthogonal transformations

$$\alpha_i = \sum_{m=0}^{r-1} a_m p_m(\bar{y}_{i\cdot})$$

$$\beta_j = \sum_{n=0}^{c-1} b_n q_n(\bar{y}_{\cdot j})$$

and

$$\alpha\beta_{ij} = \sum_{m=0}^{r-1}\sum_{n=0}^{c-1} c_{mn}\, p_m(\bar{y}_{i\cdot}) q_n(\bar{y}_{\cdot j}),$$

where a_m, b_n, and c_{mn} are unknown parameters and $p_m(\bar{y}_{i\cdot})$ and $q_n(\bar{y}_{\cdot j})$ are functions of the powers of $\bar{y}_{\cdot j}$ chosen so that each $p_m(\bar{y}_{i\cdot})$ and $q_n(\bar{y}_{\cdot j})$ are the functions of all powers of $(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$ equal to or less than m and of all powers of $(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})$ equal to or less than n respectively, and such that

$$p_0(\bar{y}_{i\cdot}) = \frac{1}{\sqrt{r}} \quad \text{for all } i$$

$$q_0(\bar{y}_{\cdot j}) = \frac{1}{\sqrt{c}} \quad \text{for all } j$$

$$\sum_{i=0}^{r-1} p_m(\bar{y}_{i\cdot}) p_{m'}(\bar{y}_{i\cdot}) = \begin{cases} 1 & \text{if } m = m' \\ 0 & \text{if } m \neq m' \end{cases}$$

* Professor of Statistics, Korea University

$$\sum_{j=0}^{c-1} q_n(\bar{y}_{\cdot j}) q_{n'}(\bar{y}_{\cdot j}) = \begin{cases} 1 & \text{if } n = n' \\ 0 & \text{if } n \neq n' \end{cases}$$

Since $\sum_c p_m(\bar{y}_{i\cdot}) = 0$ for $m \geqslant 1$ and $\sum_j q_n(\bar{y}_{\cdot j}) = 0$ for $n \geqslant 1$, we obtain an estimate of c_{mn} as follows:

$$\hat{c}_{mn} = \boldsymbol{w}_{mn}' \boldsymbol{y}$$

where $\boldsymbol{w}_{mn} = [p_m(\bar{y}_{0\cdot}) q_n(\bar{y}_{\cdot 0}),\ p_m(\bar{y}_{0\cdot}) q_n(\bar{y}_{\cdot 1}), \cdots,\ p_m(\bar{y}_{r-1,\cdot}) q_n \bar{y}_{\cdot,c-1})]'$

and $\boldsymbol{y} = [y_{00},\ y_{01},\ \cdots,\ y_{r-1,c-1}]'$.

Note that $\mathbf{1}\boldsymbol{w}_{mn} = 0$ and $\boldsymbol{w}_{mn}'\boldsymbol{w}_{mn} = 1$ for $m \neq 0,\ n \neq 0$.

Hence

$$SSc_{mn} = [\sum_i \sum_j p_m(\bar{y}_{i\cdot}) q_n(\bar{y}_{\cdot j}) y_{ij}]^2.$$

Particulary, in the case of $m=1$, $n=1$ (see Robson [4]),

$$p_1(\bar{y}_{i\cdot}) = \frac{\bar{y}_{i\cdot} - \bar{y}}{\sqrt{\sum (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}}, \qquad p_1(\bar{y}_{\cdot j}) = \frac{\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}}{\sqrt{\sum (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2}}$$

so

$$SSc_{11} = \frac{[\sum_i \sum_j (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}) y_{ij}]^2}{\sum_j (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \sum_j (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2}$$

This is the Tukey's sum of squares due to non-additivity.