

vi. Contrastes orthogonaux: Cas spécial, utile pour les comparaisons pré-planifiées.

Deux contrastes à coefficients c_i et d_i sont orthogonaux si $\sum_{i=1}^a c_i d_i = 0$.

Et pour un plan non balancé, on a $\sum_{i=1}^a n_i c_i d_i = 0$

5. Méthode de Scheffé pour la comparaison de tous les contrastes:

Dans beaucoup de situations, on ne peut pas connaître à l'avance quels contrastes on doit comparer; ou bien on peut s'intéresser à comparer plus que $(a - 1)$ comparaisons possibles. Et, dans les expériences exploratoires, les comparaisons d'intérêt sont définies et révélées uniquement après examen préliminaire des données.

La Méthode de Scheffé (1953) définit une comparaison de tous types possibles de contrastes entre traitements moyens. On suppose un ensemble de m contrastes dans les traitements moyens d'intérêts déterminés

$$\Gamma_u = c_{1u} \mu_1 + \dots + c_{au} \mu_a \\ \text{avec } u = 1, \dots, m$$

Le contraste correspondant dans les moyennes traitements, \bar{y}_i , est

$$C_u = c_{1u} \bar{y}_1 + \dots + c_{au} \bar{y}_a. \\ \text{avec } u = 1, \dots, m$$

Et l'erreur standard du contraste est définie par

$$S_{C_u} = \sqrt{CM_{SSE} \sum_{i=1}^a \frac{c_{iu}^2}{n_i}}$$

où n_i : nombre d'observations dans le i ème traitement;
Et la valeur critique contre C_u qu'on compare est:

$$S_{\alpha,u} = S_{C_u} \sqrt{(a-1)F_{\alpha,a-1,N-a}}$$

Si $|C_u| > S_{\alpha,u}$, on a rejet de l'hypothèse $H_0: \Gamma_u = 0$, i.e. le contraste Γ_u est significativement non nul.

Par la méthode de Scheffé, on définit aussi les Intervalles de confiances de tous les contrastes possibles sur les traitements moyens. Les $IC^{(1-\alpha)}(.)$ sont simultanés, définis par

$$C_u - S_{\alpha,u} \leq \Gamma_u \leq C_u + S_{\alpha,u}$$

6. Comparaison de paires traitements-moyens:

La comparaison de toutes les paires de a traitements moyens est définie sous l'hypothèse de base $H_0: \mu_i = \mu_j, \forall i \neq j$. Pour cette comparaison paire, on définit quatre méthodes:

- i. Test de Tukeys (1953): où on essaye de tester toutes les comparaisons moyennes par paires sous le corps d'hypothèses:

$$\begin{cases} H_0: \mu_i = \mu_j, & \forall i \neq j \\ H_a: \mu_i \neq \mu_j \end{cases}$$

La procédure utilise la distribution statistique de rang studentisé définie

$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{CM_{SSE}/n}}$$

où \bar{y}_{max} et \bar{y}_{min} définissent les moyennes max et min de l'échantillon.

** Pour des tailles égales d'échantillon, la Tukey-procédure déclare que deux moyennes sont significativement différentes si la valeur absolue de leurs différences d'échantillon excède la quantité:

$$T_\alpha = q_\alpha(a, f) \sqrt{CM_{SSE}/n}$$

où f : nombre de ddl associé à CM_{SSE}

$q_\alpha(a, f)$: valeur lue sur la table.

Et intervalle de confiance pour toutes paires de moyennes défini par:

$$IC^{(1-\alpha)}(\mu_i - \mu_j) = \left[\bar{y}_i - \bar{y}_j \pm q_\alpha(a, f) \sqrt{CM_{SSE}/n} \right], \forall i \neq j$$

** Lorsque les tailles d'échantillon diffèrent, on a la procédure de Tukey-Kramer suivante:

$$T_\alpha = \frac{q_\alpha(a, f)}{\sqrt{2}} \sqrt{CM_{SSE} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

$$IC^{(1-\alpha)}(\mu_i - \mu_j) = \left[\bar{y}_i - \bar{y}_j \pm \frac{q_\alpha(a, f)}{\sqrt{2}} \sqrt{CM_{SSE} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \right], \forall i \neq j$$

- ii. Méthode LSD – Fisher: C'est l'approche des différences les moins significatives de Fisher qui utilise le Fisher-test statistique pour tester l'hypothèse de base,

$$H_0: \mu_i = \mu_j, \quad \forall i \neq j.$$

* La paire de moyennes traitements, μ_i et μ_j , est déclarée différente significativement si

$$|\bar{y}_i - \bar{y}_j| > t_{\alpha/2, N-a} \sqrt{CM_{SS_E} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

* La quantité différence moins significative est définie par

$$LSD = t_{\alpha/2, N-a} \sqrt{CM_{SS_E} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

Si le plan est balancé, n_i égales, alors

$$LSD = t_{\alpha/2, N-a} \sqrt{\frac{2CM_{SS_E}}{n}}$$

- iii. Test de comparaisons multiples de Duncan (1955): défini sur le rang des a moyennes traitements qui doivent être arrangées en ordre ascendant.

* L'erreur standard pour chaque moyenne est définie par

$$(1) \quad S_{\bar{y}_i} = \sqrt{\frac{CM_{SS_E}}{n}} \text{ pour tailles égales}$$

Pour tailles inégales, l'expression (1) sera définie en fonction de la moyenne Harmonique $n_h = \frac{a}{\sum_{i=1}^a \frac{1}{n_i}}$

** Sur la table de Duncan des rangs significatifs, on lit les valeurs

$$r_{\alpha}(p, f) \text{ pour } p = 2, 3, \dots, a$$

où α définit le niveau de signification et f est le ddl. de l'erreur.

*** Pour un ensemble de rangs moins significatifs, on calcule:

$$R_p = r_{\alpha}(p, f) S_{\bar{y}_i} \quad \text{pour } p = 2, 3, \dots, a$$

**** Les différences observées entre moyennes seront testées (en commençant par le plus grand vs. le plus petit nombre des moyennes) et seront comparées avec le rang le moins significatif R_a . Puis, la différence du large vs. 2ème petit est calculée et comparée avec le rang le moins significatif R_{a-1} , etc.

Et, on continue les comparaisons jusqu'à que les différences entre toutes paires de moyennes possibles, $\frac{a(a-1)}{2}$ sont considérées.

On en déduit ainsi que si une différence observée est supérieure à son rang le moins significatif alors on conclut que les paires de moyennes sont différentes significativement.

- iv. Test de Newman-Keuls (1939): de procédure presque similaire à celle de Duncan-test où les différences critiques entre moyennes sont calculées un peu différemment.

En effet, on calcule un ensemble de valeurs critiques comme suit:

$$K_p = q_\alpha(p, f) S_{\bar{y}_i} \quad \text{pour } p = 2, 3, \dots, a$$

$q_\alpha(p, f)$ définit le rang studentisé du groupe de moyennes de taille p ,
 f : ddl. de l'erreur.

Et, on compare les paires de moyennes extrêmes des groupes de taille p à la valeur critique K_p comme dans le Duncan-test.

A suivre.