

MAT7381 Chapitre 7 Comparaisons multiples

Supposons qu'on veuille tester plusieurs hypothèses, $H_1; \dots; H_k$. Si on effectue un test pour chacune des hypothèses, chacun au niveau α , la conclusion finale n'est pas de niveau α , mais d'un niveau supérieur à α . Car si R_i est une région critique pour H_i choisie de telle sorte que $P(R_i | H_i) = \alpha$, alors la probabilité de rejeter au moins l'une des hypothèses est $P(R_1 \cup \dots \cup R_k)$, ce qui est en général supérieur à α . Les méthodes de comparaisons multiples visent à assurer un niveau *global* qui ne dépasse pas α . Plus précisément, nous cherchons une procédure pour laquelle la probabilité de rejeter *au moins* l'une des hypothèses lorsqu'elles sont *toutes* vraies soit α , ou inférieures à α :

$$P(R_1 \cup \dots \cup R_k | H_0) \leq \alpha$$

où H_0 est l'hypothèse que H_1, \dots, H_k sont toutes vraies :

$$H_0 = H_1 \wedge \dots \wedge H_k$$

L'inégalité de Bonferroni

La propriété suivante concernant la probabilité d'une union d'événements R_1, \dots, R_k est à la base d'une méthode de comparaisons multiples :

$$P(R_1 \cup \dots \cup R_k) \leq \sum_{i=1}^k P(R_i).$$

Elle servira à établir des tests d'hypothèses simultanés. L'*inégalité de Bonferroni*, pour des événements C_1, \dots, C_k , soit

$$P(C_1 \cap \dots \cap C_k) \geq 1 - \sum_{i=1}^k P(C_i^c),$$

qui découle de la règle de de Morgan, $P(C_1 \cap \dots \cap C_k) = 1 - P(C_1^c \cup \dots \cup C_k^c)$, servira à déterminer des intervalles de confiance simultanés.

Tests d'hypothèses simultanés

Considérons les régions critiques R_1, \dots, R_k pour tester, respectivement, les hypothèses H_1, \dots, H_k , et supposons que chacune est de niveau α/k , c'est-à-dire, si $P(R_i | H_i) \leq \alpha/k$. Alors la probabilité de rejeter *au moins l'une* des hypothèses H_1, \dots, H_k lorsqu'elles sont *toutes vraies* (H_0) est

$$P(R_1 \cup \dots \cup R_k | H_0) \leq \sum_{i=1}^k P(R_i | H_0) = \sum_{i=1}^k \alpha / k = \alpha.$$

Autres formulations de la procédure La procédure consiste donc à rejeter une hypothèse H_i si p_i , sa valeur p , est inférieure à α/k . Soit p_1, \dots, p_k les niveaux de signification des k tests. La procédure consiste donc à

$$\begin{aligned} & \text{rejeter } H_i \text{ si } p_i \leq \alpha/k \\ & \text{rejeter } H_0 \text{ si } p_i \leq \alpha/k, \text{ pour au moins une valeur de } i = \{1, \dots, k\} \end{aligned}$$

Ou encore, on peut énoncer la procédure comme ceci : soit $p_{(1)} \leq \dots \leq p_{(k)}$ les valeurs ordonnées des p_i et $H_{(1)}, \dots, H_{(k)}$ les hypothèses correspondantes. La procédure consiste alors à

$$\begin{aligned} & \text{rejeter les hypothèses } H_{(1)}, \dots, H_{(i)} \text{ si et seulement si } p_{(i)} \leq \alpha/k. \\ & \text{rejeter } H_0 \text{ si et seulement si } p_{(i)} \leq \alpha/k. \end{aligned}$$

Intervalles de confiance simultanés

L'inégalité de Bonferroni peut également servir à déterminer des intervalles de confiance simultanés. Soit I_i in intervalle de confiance de niveau $1 - \alpha/k$ pour un paramètre β_i , $i = 1, \dots, k$ et soit C_i l'événement $\beta_i \in I_i$. On a

$$P(C_i) \geq 1 - \alpha/k, i = 1, \dots, k$$

La probabilité que les k intervalles recouvrent simultanément leur paramètre est

$$P(C_1 \cap \dots \cap C_k) = 1 - P(C_1^c \cup \dots \cup C_k^c) \geq 1 - \sum_{i=1}^k P(C_i^c) = 1 - \sum_{i=1}^k \alpha / k = 1 - \alpha$$

La méthode de Scheffé

Lemme. Soit $\mathbf{b} \in \mathbb{R}^n$, \mathbf{A} une matrice $n \times n$ définie positive. Alors

$$\mathbf{b}'\mathbf{A}\mathbf{b} = \max_m \frac{(\mathbf{m}'\mathbf{b})^2}{\mathbf{m}'\mathbf{A}^{-1}\mathbf{m}}$$

Démonstration: Puisque \mathbf{A} est définie positive, il existe une matrice symétrique \mathbf{B} telle que $\mathbf{A} = \mathbf{B}\mathbf{B}$. Par l'inégalité de Cauchy-Schwarz, $(\mathbf{m}'\mathbf{b})^2 = (\mathbf{m}'\mathbf{B}^{-1}\mathbf{B}\mathbf{b})^2 \leq (\mathbf{m}'\mathbf{B}^{-2}\mathbf{m})(\mathbf{b}'\mathbf{B}^2\mathbf{b}) = \mathbf{m}'\mathbf{A}^{-1}\mathbf{m} \mathbf{b}'\mathbf{A}\mathbf{b}$. Nous concluons que pour tout \mathbf{m} , $(\mathbf{m}'\mathbf{b})^2 / \mathbf{m}'\mathbf{A}^{-1}\mathbf{m} \leq \mathbf{b}'\mathbf{A}\mathbf{b}$, et puisque la borne supérieure $\mathbf{b}'\mathbf{A}\mathbf{b}$ est atteignable (pour $\mathbf{m} = \mathbf{A}\mathbf{b}$, par exemple), on a

$$\mathbf{b}'\mathbf{A}\mathbf{b} = \max_m \frac{(\mathbf{m}'\mathbf{b})^2}{\mathbf{m}'\mathbf{A}^{-1}\mathbf{m}} \quad \blacksquare$$

Soit $\hat{\boldsymbol{\beta}}$ l'estimateur de $\boldsymbol{\beta}$ dans le modèle linéaire général. On sait que la variable

$$F = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) / q \hat{\sigma}^2 \sim F_{q, n-q}$$

où q est le nombre de colonnes de \mathbf{X} et $\hat{\sigma}^2 = \text{MCR}$. On a alors le théorème suivant

Théorème. Des intervalles de confiance simultanés de niveau de confiance $100(1-\alpha)\%$ pour toutes fonctions linéaires $\mathbf{m}'\boldsymbol{\beta}$ sont donnés par

$$\mathbf{m}'\hat{\boldsymbol{\beta}} - \hat{\sigma} \sqrt{q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m}F_{q; n-q; \alpha}} \leq \mathbf{m}'\boldsymbol{\beta} \leq \mathbf{m}'\hat{\boldsymbol{\beta}} + \hat{\sigma} \sqrt{q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m}F_{q; n-q; \alpha}}$$

Démonstration. Puisque $P[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) / q \hat{\sigma}^2 \leq F_{q; n-q; \alpha}] = 1 - \alpha$, on a

$$P[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq q \hat{\sigma}^2 F_{q; n-q; \alpha}] = 1 - \alpha$$

La forme quadratique à gauche est de la forme $\mathbf{b}'\mathbf{A}\mathbf{b}$ avec $\mathbf{b} = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ et $\mathbf{A} = \mathbf{X}'\mathbf{X}$. Du dernier lemme, on déduit que

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \max_m \frac{[\mathbf{m}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]^2}{\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m}}$$

Nous avons alors

$$\begin{aligned} 1 - \alpha &= P \left[\max_m \frac{[\mathbf{m}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]^2}{\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m}} \leq q \hat{\sigma}^2 F_{q; n-q; \alpha} \right] \\ &= P \left[\frac{[\mathbf{m}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]^2}{\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m}} \leq q \hat{\sigma}^2 F_{q; n-q; \alpha} \quad \forall \mathbf{m} \right] \\ &= P \left[[\mathbf{m}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]^2 \leq q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m} \hat{\sigma}^2 F_{q; n-q; \alpha} \quad \forall \mathbf{m} \right] \\ &= P \left[|\mathbf{m}'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})| \leq \sqrt{q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m} \hat{\sigma}^2 F_{q; n-q; \alpha}} \quad \forall \mathbf{m} \right] \\ &= P \left[\mathbf{m}'\hat{\boldsymbol{\beta}} - \hat{\sigma} \sqrt{q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m} F_{q; n-q; \alpha}} \leq \mathbf{m}'\boldsymbol{\beta} \leq \mathbf{m}'\hat{\boldsymbol{\beta}} + \hat{\sigma} \sqrt{q\mathbf{m}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{m} F_{q; n-q; \alpha}} \quad \forall \mathbf{m} \right] \end{aligned}$$

Ceci conclut la démonstration du théorème. ■

La méthode de la différence la moins significative (méthode LSD)

Considérons une série d'hypothèses de la forme $H_i: \gamma_i = h_i, i = 1, \dots, k$. Par exemple, si $\mu_1, \mu_2, \dots, \mu_{k+1}$ sont les moyennes de $k+1$ échantillons indépendants, on s'intéresserait aux hypothèses $H_i: \mu_i - \mu_{i+1} = 0, i = 1, \dots, k$. Il existe normalement un test global F de l'hypothèse $H_0 = H_1 \wedge \dots \wedge H_k$, ainsi que k tests t individuels des hypothèses H_i . La procédure LSD (*least significant difference*) consiste à rejeter H_i seulement si le test individuel que le test global F mènent à un rejet. Si α est le niveau de chacun de ces tests, alors la probabilité de rejeter au moins l'une des hypothèses H_i si elles sont toutes vraies est inférieure ou égale α . Car si R_i est la région critique pour tester individuellement $H_i, i = 1, \dots, k$ et R_0 est la région critique pour tester H_0 , alors

$$P(\text{rejeter au moins l'une des hypothèses } H_i | H_0) = P[(R_1 \cup \dots \cup R_k) \cap R_0 | H_0] \leq P[R_0 | H_0] \leq \alpha, \text{ et}$$

$$P(\text{rejeter } H_i | H_0) = P(R_i \cap R_0 | H_0) \leq P(R_0 | H_0) \leq \alpha.$$

Dans le modèle linéaire général $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, cette approche peut s'appliquer directement à des tests de la forme $\beta_i = h_i$, mais typiquement l'intérêt porte sur certaines fonctions linéaires $\boldsymbol{\gamma} = \mathbf{L}'\boldsymbol{\beta}$, où \mathbf{L} est $q \times r$ de rang r , q étant le nombre de colonnes de \mathbf{X} . Les hypothèses à tester seraient donc les suivantes :

$$H_0: \boldsymbol{\gamma} = \mathbf{c}$$

et puis, séparément,

$$H_i: \gamma_i = c_i, i = 1, \dots, r$$

où γ_i et c_i sont, respectivement, le i^{e} élément de $\boldsymbol{\gamma}$ et de \mathbf{c} .

Normalement, on rejette H_0 si et seulement si

$$\frac{(\hat{\boldsymbol{\gamma}} - \mathbf{c})' \mathbf{V}^{-1} (\hat{\boldsymbol{\gamma}} - \mathbf{c})}{r \hat{\sigma}^2} > \mathcal{F}_{r, n-q; \alpha} \quad (*)$$

où $\mathbf{V} = \mathbf{L}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}$, alors que si on testait une seule des hypothèses H_i , la région critique serait

$$c_i \notin (\hat{\gamma}_i - t_{n-q; \alpha/2} \sqrt{v^{ii}}; \hat{\gamma}_i + t_{n-q; \alpha/2} \sqrt{v^{ii}}), \text{ où } v^{ii} \text{ est l'élément } [i; i] \text{ de } (\hat{\sigma}^2 \mathbf{V})^{-1}. \quad (**)$$

La procédure LSD consiste à rejeter H_i si et seulement si les inégalités (*) et (**) sont toutes deux vérifiées.

Si H_0 est vraie (et donc tous les H_i le sont), la probabilité de rejeter au moins l'une des hypothèses H_i est

$$P(c_i \notin (\hat{\gamma}_i - t_{n-q; \alpha/2} \sqrt{v^{ii}}; \hat{\gamma}_i + t_{n-q; \alpha/2} \sqrt{v^{ii}}) \text{ pour } i = 1, \dots, r \text{ et } \frac{(\hat{\boldsymbol{\gamma}} - \mathbf{c})' \mathbf{V}^{-1} (\hat{\boldsymbol{\gamma}} - \mathbf{c})}{r \hat{\sigma}^2} > \mathcal{F}_{r, n-q; \alpha})$$

et cette probabilité est inférieure à $\alpha = P(\frac{(\hat{\boldsymbol{\gamma}} - \mathbf{c})' \mathbf{V}^{-1} (\hat{\boldsymbol{\gamma}} - \mathbf{c})}{r \hat{\sigma}^2} > \mathcal{F}_{r, n-q; \alpha})$.

La méthode de Tukey et la loi de l'étendue de Student

Une autre méthode, due à Tukey, est particulièrement appropriée pour les comparaisons de moyennes. Elle est basée sur une loi appelée loi de l'étendue de Student (loi de l'écart absolu normé).

Définitions. Soit X_1, \dots, X_k k variables aléatoires indépendantes, de loi commune $\mathcal{N}(\mu; \sigma^2)$. Soit

$$R = \max\{X_1, \dots, X_k\} - \min\{X_1, \dots, X_k\}$$

l'écart absolu entre la plus grande et la plus petite valeur. Soit S^2 un estimateur de σ^2 indépendant de R et supposons que

$$\frac{vS^2}{\sigma^2} \sim \chi_v^2$$

Alors la variable R/S suit une loi appelée loi de l'étendue de Student, notée $q_{k, v}$. Soit $q_{k, v; \alpha}$ le point tel

que

$$P\left(\frac{R}{S} > q_{k;v;\alpha}\right) = \alpha$$

Théorème. Soit $\bar{y}_1, \dots, \bar{y}_k$ les moyennes de k échantillons indépendants, chacun de taille r , provenant de k populations normales de moyennes μ_1, \dots, μ_k et de variance commune σ^2 :

$$\bar{y}_i \sim \mathcal{N}(\mu_i; \sigma^2/r);$$

et soit S^2 un estimateur indépendant de σ^2 tel que

$$\frac{vS^2}{\sigma^2} \sim \chi_v^2$$

Alors la probabilité est $1-\alpha$ que les $k(k-1)/2$ différences $\mu_i - \mu_j$ satisfassent simultanément les inégalités

$$\bar{y}_i - \bar{y}_j - q_{k;v;\alpha} \frac{S}{\sqrt{r}} \leq \mu_i - \mu_j \leq \bar{y}_i - \bar{y}_j + q_{k;v;\alpha} \frac{S}{\sqrt{r}}$$

Démonstration. Les variables $x_i = \bar{y}_i - \mu_i$ sont toutes de moyenne nulle et S^2/r est un estimateur de leur variance commune σ^2/r . Soit

$$R = \max\{X_1, \dots, X_m\} - \min\{X_1, \dots, X_m\}$$

Alors

$$1 - \alpha = P\left(\frac{R}{S/\sqrt{r}} \leq q_{k;v;\alpha}\right) = P\left(R \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}}\right)$$

Mais

$$R = \max_i\{X_i\} - \min_i\{X_i\} = \max_{i,j}\{|X_i - X_j|\}$$

et donc

$$P\left(R \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}}\right) = P\left(\max_{i,j}\{|X_i - X_j|\} \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}}\right) = P\left(|X_i - X_j| \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}} \quad \forall i, j\right) = 1 - \alpha$$

$$\begin{aligned} \text{Donc } 1-\alpha &= P\left(-q_{k;v;\alpha} \frac{S}{\sqrt{r}} \leq X_i - X_j \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}} \quad \forall i, j\right) \\ &= P\left(-q_{k;v;\alpha} \frac{S}{\sqrt{r}} \leq (\bar{y}_i - \mu_i) - (\bar{y}_j - \mu_j) \leq q_{k;v;\alpha} \frac{S}{\sqrt{r}} \quad \forall i, j\right) \\ &= P\left(\bar{y}_i - \bar{y}_j - q_{k;v;\alpha} \frac{S}{\sqrt{r}} \leq \mu_i - \mu_j \leq \bar{y}_i - \bar{y}_j + q_{k;v;\alpha} \frac{S}{\sqrt{r}} \quad \forall i, j\right) \end{aligned}$$

C'est la conclusion du théorème ■

Améliorations de la méthode de Bonferroni

Procédure de Holm [Holm, S. (1979) A simple sequentially rejective multiple testing procedure based on a modified Bonferroni test, *Biometrika* **75**, 2, 383-6]

Une procédure due à Holm améliore grandement la performance des tests de Bonferroni. Soit $p_1; \dots; p_k$ les niveaux de signification des k tests correspondant aux hypothèses $H_1; \dots; H_k$. Soit $p_{(1)} \leq \dots \leq p_{(k)}$ les statistiques d'ordre des p_i et $H_{(1)}; \dots; H_{(k)}$ les hypothèses correspondantes. Pour $i = 1, \dots, k$, la procédure consiste à rejeter $H_{(i)}$ si

$$p_{(i)} \leq \frac{\alpha}{k}; p_{(2)} \leq \frac{\alpha}{k-1}; \dots; \text{et } p_{(i)} \leq \frac{\alpha}{k-i+1}.$$

Il est évident que cette procédure est plus puissante que l'approche par l'inégalité de Bonferroni qui rejette $H_{(1)}, \dots, H_{(i)}$ si et seulement si $p_{(1)} \leq \frac{\alpha}{k}; p_{(2)} \leq \frac{\alpha}{k}; \dots; \text{et } p_{(i)} \leq \frac{\alpha}{k}$.

Le niveau global est α , c'est-à-dire,

$$P(\text{rejeter au moins l'une des hypothèses } H_1, \dots, H_k | H_0) \leq \alpha$$

On peut démontrer ceci à partir d'un théorème plus général dû à Hommel [Hommel, G. (1986) Multiple test procedures for arbitrary dependence structures. *Metrika* **33**, 321-36].

Théorème de Hommel Soit k hypothèses H_1, \dots, H_k et soit $H_I = \bigcap_{i \in I} H_i$ où I est un membre de l'ensemble \mathcal{P} des $M = 2^k - 1$ sous-ensembles non vides de $\mathbf{A} = \{1, \dots, k\}$. (Dans cette notation $H_I = H_0$ lorsque $\mathbf{I} = \mathbf{A}$.) Supposons que pour tout $\mathbf{I} \in \mathcal{P}$ il existe un test T_I basé sur une région critique R_I telle que $P(R_I | H_I) \leq \alpha$. Ces tests ne sont pas simultanés. Mais des tests simultanés peuvent être créés à partir de ceux-là par la procédure suivante : On rejettera H_I si le test T_I le rejette et si les tests T_J rejettent H_J pour tout $\mathbf{J} \supseteq \mathbf{I}$.

Démonstration La région critique pour H_I du test proposé est $S_I = \bigcap_{\mathbf{J} \supseteq \mathbf{I}} R_J$. Mais l'un des R_J est toujours R_A . Donc $S_I \subseteq R_A$ pour tout \mathbf{I} , et par conséquent la probabilité de rejeter au moins l'une des hypothèses H_I est $P\left(\bigcup_{\mathbf{I} \in \mathcal{P}} S_I | H_0\right) \leq P(R_A | H_0) \leq \alpha$. ■

Le test de Holm est une application de ce théorème. Soit T_I le test de Bonferroni de l'hypothèse H_I :

$$T_I : P(R_I | H_I) \leq \alpha/|\mathbf{I}|, \quad |\mathbf{I}| = \text{nombre d'éléments de } \mathbf{I}.$$

T_I est équivalent à

$$\text{rejeter } H_I \text{ si } p_{(\ell)} \leq \alpha/|\mathbf{I}|, \text{ où } \ell = \min(\mathbf{I})$$

La procédure de Holm est basée sur le fait qu'il existe un test (Bonferroni) de niveau global α de l'hypothèse H_I pour tout \mathbf{I} : On rejette $H_{(i)}$ si et seulement si H_J est rejetée pour tout \mathbf{J} qui contient i . Pour tout \mathbf{J} dont i est la plus petite valeur, il faut que $p_{(i)} \leq \alpha/|\mathbf{J}|$. $\alpha/|\mathbf{J}|$ atteint sa plus petite valeur lorsque $|\mathbf{J}| = k-i+1$, et donc il est nécessaire que

$$p_{(i)} \leq \alpha/(k-i+1) \Leftrightarrow (k-i+1)p_{(i)} \leq \alpha$$

Mais il faut aussi que pour tout $j < i$, $p_{(j)} \leq \alpha/|\mathbf{J}|$ pour tout \mathbf{J} contenant i et dont la plus petite valeur est j . La plus petite valeur de $\alpha/|\mathbf{J}|$ est $\alpha/(k-j+1)$. Il faut donc que

$$p_{(j)} \leq \alpha/(k-j+1) \text{ pour tout } j < i, \Leftrightarrow (k-j+1)p_{(j)} \leq \alpha \text{ pour tout } j < i$$

Résumons donc :

$$\text{On rejette } H(i) \text{ si et seulement si } (k-j-1)p_{(j)} \leq \alpha \text{ pour tout } j \leq i \Leftrightarrow \max\{(k-j+1)p_{(j)} \mid j \leq i\} \leq \alpha$$

Procédure On calcule

$$kp_{(1)}; (k-1)p_{(2)}; \dots; 2p_{(k-1)}; p_{(k)}$$

et puis

$$w_1 = kp_{(1)}; w_2 = \max\{kp_{(1)}; (k-1)p_{(2)}\} \leq \alpha; \dots; w_i = \max\{kp_{(1)}; \dots; (k-i+1)p_{(i)}\}; w_k = \max\{kp_{(1)}; \dots; p_{(k)}\}$$

On rejette toutes les hypothèses $H_{(1)}, \dots, H_{(i)}$ pour lesquelles

$$w_1 \leq \alpha; w_2 \leq \alpha; \dots; w_i \leq \alpha$$

La commande R `p.adjust` effectue ces calculs.

Exemple Considérons les 8 p -valeurs suivantes :

0,0103 ; 0,0306 ; 0,0626 ; 0,0774 ; 0,0848 ; 0,1553 ; 0,2374 ; 0,2916

Les valeurs $8p_{(1)} ; 7p_{(2)} ; \dots ; 2p_{(7)} ; p_{(8)}$ sont

0,0824 ; 0,2142 ; 0,3756 ; 0,3870 ; 0,3392 ; 0,4659 ; 0,4748 ; 0,2916

Nous devons évaluer les quantités suivantes :

$$L_1 = kp_{(1)}; L_2 = \max\{8p_{(1)}; 7p_{(2)}\}; L_3 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}\}; L_4 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}; 5p_{(4)}\};$$

$$L_5 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}; 5p_{(4)}; 4p_{(5)}\}; L_6 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}; 5p_{(4)}; 4p_{(5)}; 3p_{(6)}\};$$

$$L_7 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}; 5p_{(4)}; 4p_{(5)}; 3p_{(6)}; 2p_{(7)}\}; L_8 = \max\{8p_{(1)}; 7p_{(2)}; 6p_{(3)}; 5p_{(4)}; 4p_{(5)}; 3p_{(6)}; 2p_{(7)}; p_{(8)}\}.$$

On rejette $H_{(i)}$ si $L_1 \leq \alpha, \dots, \text{ et } L_i \leq \alpha$. Voici les calculs :

i	$p(i)$	$k-i+1$	$(k-i+1)p(i)$	Valeurs ajustées	
				$\max_{1 \leq j \leq i} \{(k-j+1)p_{(j)}\}$ (Holm)	$8p(i)$ (Bonferroni)
1	0,0103	8	0,0824	0,0824	0,0824
2	0,0306	7	0,2142	0,2142	0,2448
3	0,0626	6	0,3756	0,3756	0,5008
4	0,0774	5	0,3870	0,3870	0,6192
5	0,0848	4	0,3392	0,3870	0,6784
6	0,1553	3	0,4659	0,4659	1,2424
7	0,2374	2	0,4748	0,4748	1,8992
8	0,2916	1	0,2916	0,4748	2,3328

Voici la commande R et les résultats

```
> p.adjust(p, "holm")
[1] 0,0824 0,2142 0,3756 0,3870 0,3870 0,4659 0,4748 0,4748
```

Avec $\alpha = 0,05$ on ne rejette aucune hypothèse. Avec $\alpha = 0,10$, on rejette seulement $H_{(1)}$. Dans le cas présent, la méthode de Bonferroni aurait donné les mêmes résultats. Mais nous pouvons aisément imaginer des valeurs de p qui montrent que la méthode de Holm est plus puissante. En voici une série :

0,0043 0,0069 0,0081 0,0099 0,0126 0,0199 0,0245 0,0488

i	$p(i)$	$k-i+1$	$(k-i+1)p(i)$	Valeurs ajustées	
				$\max_{1 \leq j \leq i} \{(k-j+1)p_{(j)}\}$ (Holm)	$8p(i)$ (Bonferroni)
1	0,0043	8	0,0344	0,0344	0,0344
2	0,0069	7	0,0483	0,0483	0,0552
3	0,0081	6	0,0486	0,0486	0,0648
4	0,0099	5	0,0495	0,0495	0,0792
5	0,0126	4	0,0504	0,0504	0,1008
6	0,0199	3	0,0597	0,0597	0,1592
7	0,0245	2	0,049	0,0597	0,196
8	0,0488	1	0,0488	0,0597	0,3904

Avec $\alpha = 0,05$, la procédure de Holm rejette les 4 premières hypothèses, alors que celle de Bonferroni ne rejette que la première.

Procédure de Simes [R. J. Simes, An improved Bonferroni procedure for multiple tests of significance, *Biometrika* (1986), **73**,3, pp.751-4]

Simes a proposé la procédure suivante: Si pour l'un des entiers j de 1 à k est le plus grand entier pour lequel $p_{(j)} \leq j\alpha/k$, on rejette $H_0 = H_{(1)} \wedge H_{(2)} \wedge \dots \wedge H_{(k)}$. Il a démontré, sous l'hypothèse que les $p_{(i)}$ sont indépendants, que le niveau global est α , c'est-à-dire que $P(\text{rejeter au moins l'une des hypothèses } H_1; \dots; H_k | H_0) \leq \alpha$. Il n'a pas démontré la propriété plus forte établie pour la méthode de Holm, soit que $P(\text{rejeter } H_I | H_I) \leq \alpha$ pour tout H_I , $I \subseteq A = \{1, \dots, k\}$. Mais le théorème de Hommel permet de modifier la procédure de Simes de telle sorte que cette propriété soit vérifiée. Il s'agirait simplement de s'assurer que H_I ne soit rejetée que si le test de Simes rejette toute hypothèse H_J pour laquelle $J \supseteq I$.

Procédure de Hochberg [Yosef Hochberg, A sharper Bonferroni procedure for multiple tests of significance, *Biometrika* (1988), **75**, 4, pp. 800-2]

La procédure de Simes n'est pas facile à appliquer. Hochberg en a tiré cependant une procédure permettant de tester simultanément des hypothèses individuelles. Il a montré que si, pour un certain i , $(k-i+1)p_{(i)} \leq \alpha$, alors la procédure de Simes rejette toute hypothèse $H_{(j)}$ pour $j \leq i$. On en déduit la procédure suivante. On commence par la plus grande valeur, $p_{(k)}$. Si $p_{(k)} \leq \alpha$, on rejette toutes les hypothèses $H_{(1)}, \dots, H_{(k)}$; si $p_{(k)} > \alpha$, on ne peut pas rejeter $H_{(k)}$ et on s'interroge alors sur $H_{(k-1)}$. Si $p_{(k-1)} \leq \alpha/2$, on rejette les hypothèses $H_{(1)}, \dots, H_{(k-1)}$; si $p_{(k-1)} > \alpha/2$, on ne peut pas rejeter $H_{(k-1)}$. Ainsi de suite. Voici donc la procédure :

On commence par calculer

$$p_{(k)} ; 2p_{(k-1)} ; \dots ; kp_{(1)}$$

Ensuite on calcule

$$w_k = p_{(k)} ; w_{k-1} = \min(p_{(k)} ; 2p_{(k-1)}) ; \dots ; w_1 = \min(kp_{(1)} ; 2p_{(k-1)} ; \dots ; kp_{(1)})$$

Exemple Considérons les mêmes 8 p -valeurs du dernier exemple :

0,0103 0,0306 0,0626 0,0774 0,0848 0,1553 0,2374 0,2916

Voici le calcul des valeurs ajustées :

i	$p_{(i)}$	$k-i+1$	$(k-i+1)p_{(i)}$	Valeurs ajustées	
				$\min_{i \leq j \leq k} \{(k-j+1)p_{(j)}\}$ (Hochberg)	$\max_{1 \leq j \leq i} \{(k-j+1)p_{(j)}\}$ (Holm)
1	0,0103	8	0,0824	0,0824	0,0824
2	0,0306	7	0,2142	0,2142	0,2142
3	0,0626	6	0,3756	0,2916	0,3756
4	0,0774	5	0,387	0,2916	0,3870
5	0,0848	4	0,3392	0,2916	0,3870
6	0,1553	3	0,4659	0,2916	0,4659
7	0,2374	2	0,4748	0,2916	0,4748
8	0,2916	1	0,2916	0,2916	0,4748

On constate que la procédure de Hochberg est plus puissante que celle de Simes, bien que sans effet dans ce cas-ci à cause des valeurs élevées des $p_{(i)}$. La commande R est

```
> p<-c(.0103, .0306, .0626, .0774, .0848, .1553, .2374, .2916)
> p.adjust(p, "hochberg")
[1] 0.0824 0.2142 0.2916 0.2916 0.2916 0.2916 0.2916 0.2916
```

Avec $\alpha = 0,05$ on ne rejette aucune hypothèse.

Voici les calculs avec la deuxième série considérée :

i	p(i)	k-i+1	(k-i+1)p(i)	Valeurs ajustées	
				$\min_{i \leq j \leq 8} \{(k-j+1)p_{(j)}\}$ (Hochberg)	$\max_{1 \leq j \leq i} \{(k-j+1)p_{(j)}\}$ (Holm)
1	0,0043	8	0,0344	0,0344	0,0344
2	0,0069	7	0,0483	0,0483	0,0483
3	0,0081	6	0,0486	0,0486	0,0486
4	0,0099	5	0,0495	0,0488	0,0495
5	0,0126	4	0,0504	0,0488	0,0504
6	0,0199	3	0,0597	0,0488	0,0597
7	0,0245	2	0,0490	0,0488	0,0597
8	0,0488	1	0,0488	0,0488	0,0597

On constate que le test de Hochberg est plus puissant que le test de Holm. Au niveau $\alpha = 0,05$, toutes les hypothèses seraient rejetées avec Hochberg alors qu'avec Holm les hypothèses 5 à 8 ne pourraient l'être.

Méthode de Tukey dans une analyse de variance

Une adaptation approximative de l'approche de Tukey existe pour des effectifs inégaux. Le logiciel R permet de déterminer des tests et des intervalles de confiance simultanés pour une classe choisie de contrastes. Voici des données sur la consommation d'essence de plusieurs types de voiture :

Exemple Données sur plusieurs modèles de voitures

```
> fuel.frame
      Weight Disp. Mileage   Fuel   Type
Eagle Summit 4   2560   97    33 3.030303 Small
Ford Escort  4   2345  114   33 3.030303 Small
Ford Festiva 4   1845   81   37 2.702703 Small
Honda Civic  4   2260   91   32 3.125000 Small
Mazda Protege 4  2440  113   32 3.125000 Small
Mercury Tracer 4 2285   97   26 3.846154 Small
Nissan Sentra 4 2275   97   33 3.030303 Small
Pontiac LeMans 4 2350   98   28 3.571429 Small
Subaru Loyale 4 2295  109   25 4.000000 Small
Subaru Justy  3  1900   73   34 2.941176 Small
Toyota Corolla 4 2390   97   29 3.448276 Small
Toyota Tercel 4 2075   89   35 2.857143 Small
Volkswagen Jetta 4 2330  109   26 3.846154 Small
Chevrolet Camaro V8 3320 305   20 5.000000 Sporty
Dodge Daytona 2885  153   27 3.703704 Sporty
Ford Mustang V8 3310 302   19 5.263158 Sporty
Ford Probe 2695  133   30 3.333333 Sporty
Honda Civic CRX Si 4 2170   97   33 3.030303 Sporty
Honda Prelude Si 4WS 4 2710 125   27 3.703704 Sporty
Nissan 240SX 4 2775  146   24 4.166667 Sporty
Plymouth Laser 2840  107   26 3.846154 Sporty
Subaru XT 4 2485  109   28 3.571429 Sporty
Audi 80 4 2670  121   27 3.703704 Compact
Buick Skylark 4 2640  151   23 4.347826 Compact
Chevrolet Beretta 4 2655  133   26 3.846154 Compact
Chrysler Le Baron V6 3065  181   25 4.000000 Compact
Ford Tempo 4 2750  141   24 4.166667 Compact
Honda Accord 4 2920  132   26 3.846154 Compact
Mazda 626 4 2780  133   24 4.166667 Compact
Mitsubishi Galant 4 2745  122   25 4.000000 Compact
Mitsubishi Sigma V6 3110  181   21 4.761905 Compact
Nissan Stanza 4 2920  146   21 4.761905 Compact
Oldsmobile Calais 4 2645  151   23 4.347826 Compact
Peugeot 405 4 2575  116   24 4.166667 Compact
Subaru Legacy 4 2935  135   23 4.347826 Compact
Toyota Camry 4 2920  122   27 3.703704 Compact
Volvo 240 4 2985  141   23 4.347826 Compact
Acura Legend V6 3265  163   20 5.000000 Medium
```

Buick Century 4	2880	151	21	4.761905	Medium
Chrysler Le Baron Coupe	2975	153	22	4.545455	Medium
Chrysler New Yorker V6	3450	202	22	4.545455	Medium
Eagle Premier V6	3145	180	22	4.545455	Medium
Ford Taurus V6	3190	182	22	4.545455	Medium
Ford Thunderbird V6	3610	232	23	4.347826	Medium
Hyundai Sonata 4	2885	143	23	4.347826	Medium
Mazda 929 V6	3480	180	21	4.761905	Medium
Nissan Maxima V6	3200	180	22	4.545455	Medium
Oldsmobile Cutlass Ciera 4	2765	151	21	4.761905	Medium
Oldsmobile Cutlass Supreme V6	3220	189	21	4.761905	Medium
Toyota Cressida 6	3480	180	23	4.347826	Medium
Buick Le Sabre V6	3325	231	23	4.347826	Large
Chevrolet Caprice V8	3855	305	18	5.555556	Large
Ford LTD Crown Victoria V8	3850	302	20	5.000000	Large
Chevrolet Lumina APV V6	3195	151	18	5.555556	Van
Dodge Grand Caravan V6	3735	202	18	5.555556	Van
Ford Aerostar V6	3665	182	18	5.555556	Van
Mazda MPV V6	3735	181	19	5.263158	Van
Mitsubishi Wagon 4	3415	143	20	5.000000	Van
Nissan Axxess 4	3185	146	20	5.000000	Van
Nissan Van 4	3690	146	19	5.263158	Va

ANOVA simple liant la consommation (fuel) au type de voiture :

```
> a<-lm(fuel~type)
> summary(a)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.1677       0.1090  38.248 < 2e-16 ***
typeLarge    0.8001       0.2669   2.998  0.00410 **
typeMedium   0.4338       0.1599   2.712  0.00894 **
typeSmall   -0.8943       0.1599  -5.592  7.62e-07 ***
typeSporty  -0.2100       0.1779  -1.180  0.24299
typeVan      1.1456       0.1932   5.931  2.20e-07 ***
Residual standard error: 0.422 on 54 degrees of freedom
Multiple R-squared: 0.7159,    Adjusted R-squared: 0.6896
F-statistic: 27.22 on 5 and 54 DF,  p-value: 1.220e-13
```

Intervalle de confiance pour les coefficients :

```
> confint(a)
              2.5 %      97.5 %
(Intercept)  3.9491955  4.3861153
typeLarge    0.2650234  1.3352538
typeMedium   0.1131465  0.7543694
typeSmall   -1.2148865 -0.5736636
typeSporty  -0.5667931  0.1466939
typeVan      0.7583401  1.5329160
```

Ces intervalles ne sont pas simultanés. Pour des tests d'hypothèses et des intervalles de confiance simultanés, il faut d'abord charger le package `multcomp`. Ensuite, on utilise la commande `glht` (**g**eneral **l**inear **h**ypothesis **t**est).

La sortie suivante présente des tests simultanés pour les coefficients. Étant donné la paramétrisation `treatment`, ces coefficients comparent toutes les catégories à la catégorie `Compact`.

```
> b<-glht(a)
> summary(b)
Simultaneous Tests for General Linear Hypotheses
Fit: lm(formula = fuel ~ type)
Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) == 0  4.1677       0.1090  38.248 <0.001 ***
typeLarge == 0   0.8001       0.2669   2.998  0.0207 *
typeMedium == 0  0.4338       0.1599   2.712  0.0438 *
typeSmall == 0  -0.8943       0.1599  -5.592 <0.001 ***
```

```

typeSporty == 0    -0.2100    0.1779   -1.180    0.7150
typeVan == 0      1.1456    0.1932    5.931   <0.001 ***
(Adjusted p values reported -- single-step method)

```

La commande `TukeyHSD()` permet d'effectuer toutes les comparaisons deux à deux des paramètres :

```

> TukeyHSD(aov(lm(fuel~type), ordered=T))
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = lm(fuel ~ type), ordered = T)

$type
      diff      lwr      upr    p adj
Large-Compact  0.8001386  0.01156890  1.5887083  0.0448425
Medium-Compact  0.4337579 -0.03870944  0.9062253  0.0892733
Small-Compact -0.8942751 -1.36674244 -0.4218077  0.0000110
Sporty-Compact -0.2100496 -0.73576275  0.3156635  0.8441531
Van-Compact    1.1456280  0.57490318  1.7163529  0.0000032
Medium-Large   -0.3663807 -1.16499627  0.4322349  0.7528040
Small-Large    -1.6944137 -2.49302927 -0.8957981  0.0000009
Sporty-Large   -1.0101882 -1.84141367 -0.1789628  0.0088451
Van-Large      0.3454894 -0.51491065  1.2058895  0.8413999
Small-Medium   -1.3280330 -1.81708317 -0.8389828  0.0000000
Sporty-Medium  -0.6438075 -1.18447283 -0.1031422  0.0109200
Van-Medium     0.7118701  0.12734337  1.2963969  0.0086513
Sporty-Small   0.6842255  0.14356017  1.2248908  0.0056917
Van-Small      2.0399031  1.45537637  2.6244299  0.0000000
Van-Sporty    1.3556777  0.72733028  1.9840250  0.0000006

```

On peut également effectuer des comparaisons personnalisées:

```

> dif<-glht(lm(fuel~type), linfct=c("typeLarge-typeMedium=0", "typeMedium-
typeSmall=0", "typeSmall-typeSporty=0", "typeSporty-typeVan=0"))
> summary(dif)
Simultaneous Tests for General Linear Hypotheses
Fit: lm(formula = fuel ~ type)
Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
typeLarge - typeMedium == 0  0.3664    0.2703   1.355  0.50604
typeMedium - typeSmall == 0  1.3280    0.1655   8.023  < 0.001 ***
typeSmall - typeSporty == 0 -0.6842    0.1830  -3.739  0.00183 **
typeSporty - typeVan == 0   -1.3557    0.2127  -6.374  < 0.001 ***
(Adjusted p values reported -- single-step method)

```

Ou encore, on peut donner une *matrice* de contrastes :

```

> M
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    0    1   -1    0    0    0
[2,]    0    0    1   -1    0    0
[3,]    0    0    0    1   -1    0
> summary(glht(lm(fuel~type), linfct=M))
Simultaneous Tests for General Linear Hypotheses
Fit: lm(formula = fuel ~ type)
Linear Hypotheses:
              Estimate Std. Error t value Pr(>|t|)
1 == 0  0.3664    0.2703   1.355  0.42272
2 == 0  1.3280    0.1655   8.023  < 0.001 ***
3 == 0 -0.6842    0.1830  -3.739  0.00137 **
(Adjusted p values reported -- single-step method)

```

On obtient le même résultat à l'aide d'intervalles de confiance simultanés. Le niveau des intervalles est de 95 % :

```

> confint(glht(lm(fuel~type), linfct=M))
      Simultaneous Confidence Intervals
Fit: lm(formula = fuel ~ type)
Quantile = 2.4446
95% family-wise confidence level
Linear Hypotheses:
      Estimate lwr      upr
1 == 0  0.3664 -0.2944  1.0272
2 == 0  1.3280  0.9234  1.7327
3 == 0 -0.6842 -1.1316 -0.2369

```

On peut réduire le niveau de confiance afin d'obtenir des intervalles plus courts. Voici des intervalles à 80 % :

```

> confint(glht(lm(fuel~type), linfct=M, level=.8))
      Simultaneous Confidence Intervals
Fit: lm(formula = fuel ~ type)
Quantile = 2.4435
95% family-wise confidence level
Linear Hypotheses:
      Estimate lwr      upr
1 == 0  0.3664 -0.2941  1.0269
2 == 0  1.3280  0.9236  1.7325
3 == 0 -0.6842 -1.1314 -0.2371

```

Exemple Plan en blocs

Voici un plan en quatre blocs avec quatre traitements dans chaque bloc:

	y	bloc	traitement
1	133	1	1
2	128	1	2
3	129	1	3
4	134	1	4
5	132	2	1
6	127	2	2
7	130	2	3
8	125	2	4
9	127	3	1
10	127	3	2
11	124	3	3
12	126	3	4
13	132	4	1
14	128	4	2
15	129	4	3
16	131	4	4

Voici les résultats d'une analyse de variance à deux facteurs avec une observation par cellule :

```

> summary(b)
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  133.125      1.433  92.892 9.84e-15 ***
bloc2         -2.500      1.532  -1.632 0.13716
bloc3         -5.000      1.532  -3.264 0.00978 **
bloc4         -1.000      1.532  -0.653 0.53026
traitement2   -3.500      1.532  -2.284 0.04821 *
traitement3   -3.000      1.532  -1.958 0.08189 .
traitement4   -2.000      1.532  -1.305 0.22413

Residual standard error: 2.167 on 9 degrees of freedom
Multiple R-squared: 0.6693, Adjusted R-squared: 0.4488
F-statistic: 3.036 on 6 and 9 DF, p-value: 0.06574

```

Tous les paramètres

Considérons le vecteur des valeurs p pour tous les paramètres du modèle (à l'exception du premier):

```
      b2      b3      b4      t2      t3      t4
0.13716 0.00978 0.53026 0.04821 0.08189 0.22413
```

Les voici ordonnées:

```
      b3      t2      t3      b2      t4      b4
0.00978 0.04821 0.08189 0.13716 0.22413 0.53026
```

Le paramètre b_3 , comparant le bloc 3 au bloc 1, et t_2 , comparant le traitement 2 au traitement 1, sont significatifs selon ces valeurs p non ajustées. Lorsqu'on les ajuste, (selon Holm ou Hochberg, il se trouve que les résultats sont les mêmes), on obtient les valeurs ajustées suivantes:

```
      b3      t2      t3      b2      t4      b4
0.05870 0.24103 0.32755 0.41148 0.44826 0.53026
```

Et là, hélas, on ne trouve rien de significatif à 5 %.

Et voici l'ensemble des intervalles de confiance simultanés :

```
> TukeyHSD(aov(b))
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = b)

$bloc
      diff      lwr      upr      p adj
2-1 -2.5 -7.2827976  2.2827976 0.4094263
3-1 -5.0 -9.7827976 -0.2172024 0.0403975
4-1 -1.0 -5.7827976  3.7827976 0.9119780
3-2 -2.5 -7.2827976  2.2827976 0.4094263
4-2  1.5 -3.2827976  6.2827976 0.7644984
4-3  4.0 -0.7827976  8.7827976 0.1074855

$traitement
      diff      lwr      upr      p adj
2-1 -3.5 -8.282798  1.282798 0.1728954
3-1 -3.0 -7.782798  1.782798 0.2714320
4-1 -2.0 -6.782798  2.782798 0.5821711
3-2  0.5 -4.282798  5.282798 0.9871995
4-2  1.5 -3.282798  6.282798 0.7644984
4-3  1.0 -3.782798  5.782798 0.9119780
```

Là l'approche de Tukey s'avère plus puissante dans la mesure où la différence entre le bloc 3 et le bloc 1 est maintenant significative. La différence entre le traitement 2 et le traitement 1 demeure non significative. Il faut souligner, cependant, que le nombre de contrastes est supérieur à ceux que nous avons considérés par la méthode de Hochberg et/ou Holm.

Contrastes deux à deux

La situation est encore moins reluisante lorsque le nombre de comparaisons augmente. Voici l'ensemble de valeurs p pour un nombre supérieur de comparaisons :

Voici le vecteur des valeurs p pour toutes les différences deux à deux (entre blocs et entre traitements) :

```
      b2      b3      b4      t2      t3      t4  b2-b3  b2-b4  b3-b4  t2-t3  t2-t4  t3-t4
0.1372 0.0098 0.5303 0.0482 0.0819 0.2241 0.1372 0.3531 0.0282 0.7516 0.3531 0.5303
```

Les voici ordonnées

```
      b3  b3-b4      t2      t3  b2-b3      b2      t4  t2-t4  b2-b4  t3-t4      b4  t2-t3
0.0098 0.0282 0.0482 0.0819 0.1372 0.1372 0.2241 0.3531 0.3531 0.5303 0.5303 0.7516
```

Les voici ajustées par la méthode de Holm

```
      b3  b3-b4      t2      t3  b2-b3      b2      t4  t2-t4  b2-b4  t3-t4      b4  t2-t3
```

0.1174 0.3105 0.4821 0.7370 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000

Les voici ajustées par la méthode de Hommel

	b3	b3-b4	t2	t3	b2-b3	b2	t4	t2-t4	b2-b4	t3-t4	b4	t2-t3
0.1174	0.3105	0.4821	0.7370	0.7516	0.7516	0.7516	0.7516	0.7516	0.7516	0.7516	0.7516	0.7516

Valeurs de p plus petites

La faible puissance de ces deux dernières méthodes est moins frappante lorsque les valeurs *p* sont plus petites. Supposons que les valeurs de *y* avaient été les suivantes :

146 179 206 241 145 178 208 225 140 178 198 227 145 179 206 236

Les valeurs *p* ordonnées pour toutes les différences deux à deux (entre blocs et entre traitements) auraient été plus petites :

	t4	t3	t2-t4	t2	t3-t4	t2-t3	b3	b3-b4	b2	b2-b3	b2-b4	b4
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0253	0.0627	0.1739	0.2608	0.3801	0.5932

Les voici ajustées par la méthode de Holm

	t4	t3	t2-t4	t2	t3-t4	t2-t3	b3	b3-b4	b2	b2-b3	b2-b4	b4
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1521	0.3137	0.6954	0.7825	0.7825	0.7825

Les voici ajustées par la méthode de Hommel

	t4	t3	t2-t4	t2	t3-t4	t2-t3	b3	b3-b4	b2	b2-b3	b2-b4	b4
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1521	0.3137	0.5216	0.5702	0.5932	0.5932

Comparons ces résultats avec ceux obtenus par la méthode de Tukey :

```
> TukeyHSD(aov(aaa))
  Tukey multiple comparisons of means
  95% family-wise confidence level

Fit: aov(formula = aaa)

$bloc
      diff          lwr          upr      p adj
2-1 -4.00 -12.455872   4.455872 0.4881408
3-1 -7.25 -15.705872   1.205872 0.0974930
4-1 -1.50  -9.955872   6.955872 0.9431804
3-2 -3.25 -11.705872   5.205872 0.6419809
4-2  2.50  -5.955872  10.955872 0.7936886
4-3  5.75  -2.705872  14.205872 0.2170323

$traitement
      diff          lwr          upr      p adj
2-1 34.50 26.04413 42.95587 2.30e-06
3-1 60.50 52.04413 68.95587 0.00e+00
4-1 88.25 79.79413 96.70587 0.00e+00
3-2 26.00 17.54413 34.45587 2.43e-05
4-2 53.75 45.29413 62.20587 0.00e+00
4-3 27.75 19.29413 36.20587 1.42e-05
```