

## Compléments sur la régression multiple

Frédéric Bertrand et Myriam Maumy<sup>1</sup>

<sup>1</sup>IRMA, Université Louis Pasteur  
Strasbourg, France

Master 2ème Année 28-03-2009

- Tester l'hypothèse nulle :  
$$H_0 : \beta_j = 0$$
 contre l'hypothèse alternative :  
$$H_1 : \beta_j \neq 0 \text{ pour un certain } j \text{ entre } 0 \text{ et } p - 1.$$
- Calculer la statistique :  
$$t_{obs} = \frac{\hat{\beta}_j}{s(\hat{\beta}_j)}.$$
- Comparer  $t_{obs}$  à la valeur théorique lue dans une table de Student à  $(n - p)ddl$  et à  $\alpha = 0, 05$ .

- Un intervalle de confiance au niveau  $(1 - \alpha)$  où  $\alpha$  est la probabilité d'erreur pour  $\beta_j$  est défini par :

$$\left[ \hat{\beta}_j - t_{\alpha/2, n-p} \times s(\hat{\beta}_j); \hat{\beta}_j + t_{\alpha/2, n-p} \times s(\hat{\beta}_j) \right].$$

- Cet intervalle de confiance est construit de telle sorte qu'il contienne le paramètre inconnu  $\beta_j$  avec une probabilité de  $(1 - \alpha)$ .

- Tester l'hypothèse nulle :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$$

contre l'hypothèse alternative :

$$H_1 : \exists j \text{ pour lequel } \beta_j \neq 0 \text{ où } j \text{ varie de } 1 \text{ à } p - 1.$$

- Si l'hypothèse nulle  $H_0$  est vérifiée alors le modèle s'écrit :

$$y_i = \beta_0 + \varepsilon_i.$$

Par conséquent :

$$F_{obs} = \frac{MC_{reg}}{MC_{res}}$$

suit une loi de Fisher avec  $(p - 1)$  et  $(n - p)$  ddl, où

$$MC_{reg} = \frac{SC_{reg}}{p - 1} = \frac{\sum (\hat{y}_i - \bar{y})^2}{p - 1}$$

et

$$MC_{res} = \frac{SC_{res}}{n - p} = \frac{\sum (y_i - \hat{y}_i)^2}{n - p}.$$

Source de variation	SC	ddl	MC	$F_{obs}$
Régression	$SC_{reg} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$p - 1$	$\frac{SC_{reg}}{p - 1}$	$\frac{MC_{reg}}{MC_{res}}$
Résiduelle	$SC_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	$n - p$	$\frac{SC_{res}}{n - p}$	
Totale	$SC_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2$	$n - 1$		

## Test F partiel :

la nullité d'un certain nombre  $r$  de paramètres dans un modèle de  $p$  paramètres.

$H_0$  : modèle réduit avec  $(p - r)$  paramètres

$H_1$  : modèle complet avec  $p$  paramètres.

Exemples :

- Tester la nullité d'un paramètre, par exemple :  $\beta_1$ .

$$H_0 : y_i = \beta_0 + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \text{ contre}$$

$$H_1 : y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i.$$

- Tester la nullité de plusieurs paramètres, par exemple les pairs :  $\beta_{2j}$ .

$$H_0 : y_i = \beta_1 x_{i1} + \beta_3 x_{i3} + \dots + \beta_{p-1} x_{i,p-1} + \varepsilon_i \text{ contre}$$

$$H_1 : y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \text{ avec } p \text{ pair.}$$

## La procédure :

- Calculer les valeurs estimées  $\hat{y}_i$  en utilisant la méthode des MC pour chacun des 2 modèles définis par  $H_0$  et  $H_1$ , notées :  $\hat{y}_i(H_0)$  et  $\hat{y}_i(H_1)$ .
- Calculer ensuite  $SC_{res}(H_0)$  et  $SC_{res}(H_1)$ .
- Calculer la statistique :

$$F_{obs} = \frac{SC_{res}(H_0) - SC_{res}(H_1)}{SC_{res}(H_1)} \times \frac{n - p}{r}$$

- Rejeter l'hypothèse nulle au seuil  $\alpha$  si

$$F_{obs} > F_{\alpha; r, n-p}$$

où  $F_{\alpha; r, n-p}$  est le  $(1 - \alpha)$ -quantile d'une loi de Fisher avec  $r$  et  $n - p$  ddl que l'on trouve dans une table de Fisher.