

I. RÉGRESSION LOGISTIQUE POLYTOMIALE

Le modèle

- Variable réponse X (I catégories)
- Variables explicatives : Y (J catégories) et Z (K catégories)
→ choix des références pour i, j et k

(I-1) régressions :

$$\log \text{Cote}(X = i | Y = j, Z = k) = \mu^{(i)} + \alpha_j^{(i)} + \beta_k^{(i)} + \gamma_{jk}^{(i)} \quad \text{avec} \quad \alpha_1^{(i)} = \beta_1^{(i)} = \gamma_{1k}^{(i)} = \gamma_{j1}^{(i)} = 0$$

Etapas de l'analyse

- 1) Procéder à la sélection du modèle (uniquement sur base du type 3 avec « proc logistic »)
- 2) Interpréter les coefficients du modèle retenu

NB : Equivalents des coefficients dans le modèle de régression log-linéaire saturé :

$$\alpha_j^{(i)} = \lambda_{ij}^{XY} \quad ; \quad \beta_k^{(i)} = \lambda_{ik}^{XZ} \quad ; \quad \gamma_{jk}^{(i)} = \lambda_{ijk}^{XYZ}$$

Le code SAS

```
proc logistic data = SAS-dataset;  
  class Y (ref = 'ref') Z (ref='ref')/param = ref;  
  model X (ref = 'ref') = Y Z Y*Z/link = glogit;  
run;
```

Exercice

Refaites l'analyse de la table 4x2x3 (Evaldemo4 (X), polintr2 (Y), reg3 (Z) – dernier exercice du TP sur le chapitre 2) au moyen d'un modèle de régression logistique polytomiale plutôt que log-linéaire, en limitant l'analyse aux références suivantes :

- Evaldemo4 : « Very democratic »
- Polintr2 : « Interested »
- Reg3 : « Wal »

Vérifiez que les estimations obtenues sont bien les mêmes qu'en régression log-linéaire.

Présentez les résultats dans un rapport pdf (avec titres adéquats, etc.).

II. MODÈLE À COTES PROPORTIONNELLES

Le modèle

Variable réponse X ordinaire

$$\log \text{Cote}(X \leq i | Y = j, Z = k) = \mu_i + \alpha_j + \beta_k + \gamma_{jk} \quad \text{avec} \quad \alpha_1 = \beta_1 = \gamma_{1k} = \gamma_{j1} = 0$$

Étapes de l'analyse

- 1) Procéder à la sélection du modèle (cf. démo réalisée durant le TP)
- 2) Interpréter les coefficients du modèle retenu

NB : les paramètres s'interprètent indépendamment de i , ce qui simplifie considérablement l'interprétation des effets **si cette hypothèse est raisonnable** (d'où l'intérêt d'une démarche consistant à partir d'un tel modèle – simple – pour sélectionner les explicatives pertinentes puis d'affiner l'examen des effets au moyen d'une régression logistique polytomiale, par exemple).

Le code SAS

```
proc logistic data = SAS-dataset;  
  class Y Z(ref="ref") / param=ref ref=first;  
  model X = Y Z Y*Z / expb clparm=both;  
run;
```

OU

```
proc genmod data = SAS-dataset;  
  class Y Z(ref="ref") / param=ref ref=first;  
  model X = Y Z Y*Z /dist=mult type1 type3;  
run;
```

Exercice

Au terme de la démonstration sur la sélection des explicatives (réalisée avec la proc genmod), nous avons obtenu un modèle final avec 5 variables explicatives (reg3 educ3 plinsoc3 polintr2 hincfel) et 3 termes d'interactions (plinsoc3*hincfel polintr2*hincfel reg3*plinsoc3).

Ajustez ce modèle au moyen de la proc logistic, en prenant pour référence :

- Reg3 : « Wal »
- Educ3 : « Middle »
- Plinsoc3 : « Middle »
- Polintr2 : « Interested »
- Hincfel : « Coping on present income »

Intreprenez les paramètres. Quels sont les constats principaux qui se dégagent de cette analyse globale ?