Partiel Statistiques Appliquées

Mardi 16 janvier 2007 : 8h30 - 11h30
Cours de F. GARDES

Sont autorisées les calculatrices. Les six parties sont indépendantes les unes des autres.

Partie I : Question de cours (1,5 points)

Traitez au choix l'une des questions A ou B.

A. Pourquoi dit-on que les estimateurs par les MCO d'un modèle économique sont des variables aléatoires ?

B. Enoncez le théorème Central Limite. Donnez-en une application.

Partie II : Question de cours (2,5 points)

Traitez au choix l'une des questions A ou B.

A. Expliquez :

1. Le principe de la méthode des MCO pour l'estimation d'un modèle linéaire : $y = X\beta + u$
2. Le principe de la méthode du Maximum de Vraisemblance.

B.

1. Justifiez la distance utilisée dans le tests du χ^2 :

$$D_n = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n_{i.}.n_{.j}/n)^2}{n_{i.}.n_{.j}/n} = n\left(\sum_{i=1}^{r}\sum_{j=1}^{s} \frac{{n_{ij}}^2}{n_{i.}n_{.j}} - 1\right) = n\left(\sum_{i=1}^{r}\sum_{j=1}^{s} \frac{{p_{ij}}^2}{p_{i.}p_{.j}} - 1\right)$$

   pour n_{ij} le nombre d'individus correspondant aux items de réponse i et j de deux caractères qualitatifs (à respectivement r et s modalités), $n_{i.}$ et $n_{.j}$ les sommes marginales, (avec $p_{ij} = \frac{n_{ij}}{n}$, $p_{i.} = \frac{n_{i.}}{n}$ et $p_{.j} = \frac{n_{.j}}{n}$) et n la taille de la population enquêtée.
2. Donnez un exemple économique où ce test peut être utilisé.

Partie III : Exercice (5 points)

1. Définissez la convergence en Probabilité et la convergence en Moyenne Quadratique.
2. Quelle relation existe-t-il entre elles ?
3. On considère un estimateur T_n sans biais de l'élasticité de la consommation permanente C^p par rapport au revenu permanent Y^p, obtenue par les MCO sur une enquête de consommation auprès de n ménages. Expliquez ce que signifie cette hypothèse d'absence de biais.
4. Supposons que la valeur vraie de cette élasticité soit 1. Démontrez que l'estimateur T_n est convergent en faisant l'hypothèse que la variance de T_n tend vers 0 quand la taille de l'échantillon tend vers l'infini.

5. Rappelez la formule matricielle qui permet d'obtenir cet estimateur des MCO par l'estimation de l'équation :
$$lnC_i^p = \beta_1 lnY_i^p + \beta_2 D_i + \beta_3 + u_i = X_i\beta + u_i$$
du logarithme de la consommation permanente lnC_i^p (de la famille i) par rapport au logarithme du revenu permanent de cette même famille lnY_i^p, à la taille de la famille D_i et à une constante β_3, ces trois variables explicatives étant rassemblées dans le vecteur X_i, avec un résidu u_i (on donnera la formule pour le vecteur de paramètres β, vecteur colonne contenant β_1 dans sa première ligne).
6. Quelles sont les propriétés de cet estimateur sous les hypothèses habituelles ?

Partie IV : Question de cours (2 points)

Traitez les deux questions ci-dessous :
1. Définissez les deux risques d'erreurs dans la théorie des tests. Pourquoi privilégie-t-on le risque de première espèce ?
2. On veut tester l'hypothèse que l'élasticité de la consommation permanente par rapport au revenu est égale à 1. En reprenant les notations de la partie précédente, énoncez l'hypothèse nulle et l'hypothèse alternative et caractérisez l'erreur de première espèce.

Partie V : Exercice (6 points)

Nous disposons d'un échantillon de 25 étudiants tiré de manière aléatoire dans la population étudiante ayant souscrit un emprunt. La dette moyenne dans cet échantillon est de 10290 euros. On considère que la dette d'un étudiant ayant souscrit un emprunt suit une loi normale $N(m, \sigma^2)$. L'écart-type théorique de cette dette, sur l'ensemble de la population étudiante ayant souscrit un emprunt, est supposé connu : $\sigma = 2500$ euros.
1. Construisez un intervalle bilatéral de confiance à 90% pour estimer la dette moyenne m de l'ensemble de la population étudiante ayant souscrit un emprunt.
2. Même question pour un intervalle bilatéral de confiance à 99%.
3. Expliquez l'effet de l'augmentation du niveau de confiance sur la longueur de l'intervalle.
4. On suppose dans cette question (et uniquement dans cette question) qu'on ne connaît pas l'écart-type théorique σ de la dette des étudiants ayant souscrit un emprunt. Par contre, on dispose de l'écart-type empirique s obtenu à partir du même échantillon : $s = 2000$ euros. Déterminez un intervalle bilatéral de confiance à 90% et un intervalle bilatéral de confiance à 95% pour la dette moyenne m.
5. On veut tester l'hypothèse que la dette moyenne de la population étudiante qui souscrit un emprunt est inférieure ou égale à 9300 euros. Ecrivez les hypothèses nulle et alternative et, connaissant l'écart-type théorique de 2500 euros, effectuez le test pour les deux seuils $\alpha = 5\%$ et $\alpha = 1\%$.
6. Pour un seuil $\alpha = 1\%$, quelle taille d'échantillon est nécessaire pour obtenir une réponse négative à un test d'une dette moyenne (de la population étudiante ayant souscrit un emprunt) inférieure ou égale à 9300 euros ?

Partie VI : Exercice (3 points)

1. Soit (X_i) un échantillon tiré de façon i.i.d dans une loi normale $N(m, \sigma^2)$. Calculez l'estimateur du Maximum de Vraisemblance du paramètre m. Comment s'appelle l'estimateur que vous trouvez ?
Indication : Fonction de densité de la loi normale $\mathcal{N}\left(m, \sigma^2\right)$:
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} \qquad x \in R$$
2. On dispose d'un échantillon i.i.d tiré dans la loi de densité :
$$f(y; a) = ay^{a-1} \qquad\qquad avec\ \ 0 < y < 1$$
Donnez un estimateur de a suivant la méthode des moments.