Fiche de Statistique et Econométrie
Licence L3 d'Economie, Université Paris I Panthéon Sorbonne
TD n° 1 : Rappels

Références :

Chapitre I : Rappels : Notions de probabilité ; variable aléatoire, lois statistiques usuelles, vecteurs aléatoires, types de convergence, notion de modèle économétrique

Lecoutre, Chapitres 1 à 4
Kauffmann, Chapitres 1, 2, 3

I. Probabilités :

Exercice 1 : Problème du Chevalier de Méré
Résoudre le problème du *Chevalier de Méré* qui demandait lequel des évènements suivants est le plus probable :
A = "obtenir au moins une fois le 6 en quatre lancers de dé" ou
B = "obtenir au moins une fois un double 6 en vingt-quatre lancers de deux dés".

Exercice 2 : Fiabilité de l'alcotest
Un laboratoire a mis au point un alcootest dont les propriétés sont décrites ci-après :
- il se révèle positif pour quelqu'un qui n'est pas en état d'ébriété dans 2% des cas ;
- il se révèle positif pour quelqu'un qui est en état d'ébriété dans 96% des cas.

Dans un département donné, on sait que 3% des conducteurs sont en état d'ébriété. Si le contrôle se révèle positif, quelle est la probabilité que ce conducteur ne soit pas malgré tout en état d'ébriété?

Exercice 3 : Evènements indépendants
Les évènements A et B étant indépendants, montrer que les évènements A et $\overline{B}$ ainsi que $\overline{A}$ et $\overline{B}$ sont aussi indépendants.
Si l'évènement C est indépendant de A et de B, est-il aussi indépendant de $A \cup B$? L'évènement $\overline{C}$ est-il indépendant de $A \cap B$?

Exercice 4 : Formule de Bayes
The book of Risks (1994) contient des informations probabilistes sur les risques que courent les individus dans leurs activités quotidiennes. Par exemple la probabilité qu'un homme ait un accident de voiture au cours d'une année est, semble-t-il, deux fois plus élevée que la probabilité qu'une femme ait un accident de voiture au cours de la même période. Les probabilités indiquées sont de 0,113 pour un homme et de 0,057 pour une femme. Supposez que 55% des conducteurs dans le comté de Lucas soient des hommes. En remplissant un questionnaire sur la conduite automobile, une personne originaire du comté de Lucas indique qu'elle a eu un accident de la route au cours de l'année passée. Quelle est la probabilité que cette personne soit une femme?

Exercice 5 :
Dans un sondage d'opinion, la question suivante était posée aux Américains : "Etes-vous satisfaits ou insatisfaits de l'état actuel de l'économie américaine?" (The Wall Street Journal, 27 juin 1997). Les réponses de tous les adultes, regroupées par tranches d'âge, sont présentées dans le tableau suivant.

	Satisfaits (%)	Insatisfaits (%)	Autre (%)
Tous les adultes	61	37	2
18-34	64	35	1
35-49	58	41	1
50-64	57	40	3
65+	70	26	4

1) Quelle est la probabilité qu'un adulte sélectionné aléatoirement soit satisfait?
2) Quelle est la tranche d'âge qui a un niveau de satisfaction plus élevé que la moyenne?
3) Quelle est la probabilité qu'un adulte de 65 ans ou plus n'ait pas répondu qu'il était satisfait?

Exercice 6 :
On considère les familles de deux enfants. X est le nombre de garçons, Y le nombre de filles. Les évènements $\{X = 2\}$ et $\{Y = 2\}$ sont supposés équiprobables et la probabilité d'avoir un garçon et une fille est égale à 0,5.

1) Présentez le tableau de la loi conjointe. Les variables X et Y sont-elles indépendantes ?
2) Déterminez la loi marginale de X et celle de Y.
3) Déterminez la distribution de probabilité de $Z = \frac{|X - Y|}{2}$.

II. Variables aléatoires :

Exercice 7 :
Vous participez à un jeu où vous avez la probabilité p de remporter une partie. Si vous gagnez deux parties consécutives le jeu s'arrête et vous remportez un gain de $40-4N$ francs, N étant le nombre total de parties jouées. Le nombre maximum de parties jouées est fixé à quatre et vous donnez à votre adversaire dans ce jeu la somme de 25 euros en cas de perte. Ce jeu vous parait-il équitable?

Exercice 8 :
La fonction de répartition F d'une v.a. est définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 - 1/(1*2) & \text{si } 1 \leq x < 2 \\ 1 - 1/(2*3) & \text{si } 2 \leq x < 3 \\ \dots & \dots \\ 1 - 1/n(n+1) & \text{si } n \leq x < n+1 \\ \dots & \dots \end{cases}$$

1) Calculez les probabilités $p_n = P(X = n), \ n \in \mathbb{N}^*$.
2) Calculez $E(X)$ et $V(X)$.

Exercice 9 :
Nous avons envoyé 600 propositions à des clients potentiels tirés dans une très large population et nous avons en retour obtenu 78 réponses favorables. Nous nous intéressons à la valeur de la probabilité p de succès pour chaque envoi et nous pouvons supposer que les comportements des individus sont indépendants les uns des autres.
Décrire avec précision le modèle statistique correspondant à ces observations : quelle est la variable aléatoire observée, quel est le nombre d'observations, quelle est la loi de ces observations?

III. Lois statistiques usuelles :

Exercice 10 :
Soit U est une v.a. de loi normale standard, calculez $P(U < -2)$, $P(-1 < U < 0{,}5)$ et $P(4U \geq -3)$ puis déterminez u_0 et v_0 tels que $P(|U| < u_0) = 0{,}82$ et $P(U < -v_0) = 0{,}61$.

Exercice 11 :
Si X est une v.a. de loi normale telle que $P(X < 2) = 0{,}0668$ et $P(X \geq 12) = 0{,}1587$, calculez la valeur de a telle que $P\left\{ [X - E(X)]^2 < a \right\} = 0{,}95$.

IV. Vecteurs aléatoires :

Exercice 12 :
On considère le couple de variables (X, Y) de fonction de densité conjointe :

$f_{X,Y} = c.\exp(-b(x + y))$ si $0 < x < y < +\infty$ et $f_{X,Y} = 0$ sinon,

avec c, b réels positifs non nuls.

1) Les variables X et Y sont-elles indépendantes ?
2) Quelle condition doit-on imposer à b et c ?
3) On suppose que $c = 2$. Déterminer les densités marginales de X et de Y.
4) Déterminer les densités conditionnelles de $X/ Y = y$ et de $Y/ X = x$.
5) Déterminer $E(Y/X = x)$.

Exercice 13 :
Si X est un vecteur de loi $\mathrm{N}_n(\mu, \Sigma)$, montrez que la v.a. réelle ${}^t(X - \mu)\,\Sigma^{-1}(X - \mu)$ suit une loi du Khi-deux à n degrés de liberté.

V. Modélisation économétrique :

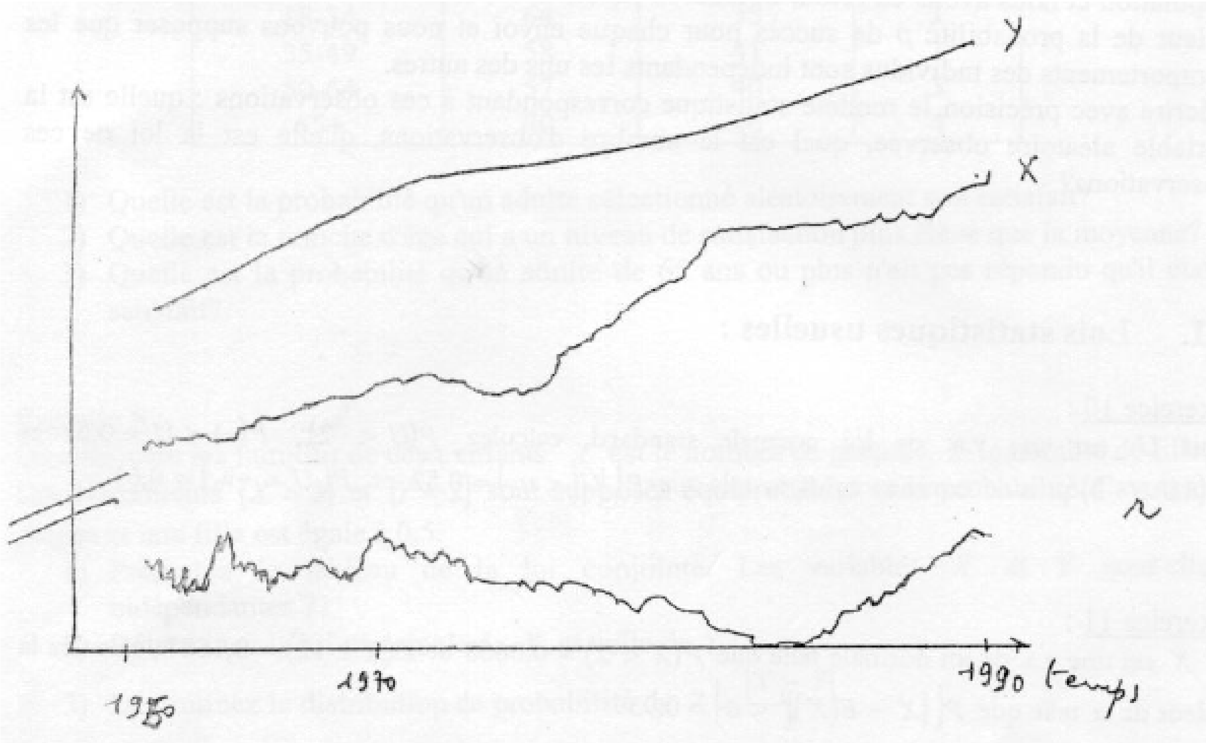
Exercice 15 :

Voici les graphiques, en fonction du temps (données annuelles de 1960 à 1990) de trois variables économiques :

X = consommation per capita des ménages (en francs constants)

Y = revenus per capita des ménages (en francs constants)

r = taux d'intérêt réel



1) Quelle relation théorique peut-on proposer entre ces trois variables?
2) Deux problèmes principaux vont se poser dans cette estimation :
    - Quels sont-ils ?
    - Comment les corriger ?

Exercice 16 :

Pour quelles raisons le R^2 d'une estimation effectuée sur données individuelles diffère-t-il de celui qu'on obtiendrait sur les mêmes données, mais après agrégation?

Donnez un exemple de ce phénomène.

Exercice 17 :

1) Quel problème pose l'existence de valeurs aberrantes dans les données?
2) Comment peut-on le corriger?
3) Donnez un exemple d'un tel phénomène.

Compléments conseillés :

LECOUTRE, Chapitre 1, exercices N° 2, 3, 13, 18.

LECOUTRE, Chapitre 3, exercice N° 16 (log-normale).

Exercice 18 : **Texte :** ***Le problème de la causalité***

James J. Heckman, 2000. "Causal Parameters And Policy Analysis In Economics: A Twentieth Century Retrospective," The Quarterly Journal of Economics, MIT Press, vol. 115(1), pages 45-97, February.

Questions sur le texte :

1) Pourquoi l'inter-relation des causes (voir page 47 lignes 163) rend-t-elle plus difficile la recherche des causalités ?
2) Définissez la notion d'identification (p. 47) en économétrie.
3) Quelle fut la « méthode de la Cowles Commission ». Quelles critiques en ont été faites ?
4) Pourquoi préférer l'interprétabilité des paramètres à la qualité de la prévision (« fit ») (page 50)?
5) Qu'est-ce qu'une « expérience naturelle » (page 51)?
6) Discutez la critique d'Heckman des conceptions classiques (Granger-Sims) de la causalité temporelle (page 60).