

Université Paris 1, UFR 02, Licence de Sciences Economiques
STATISTIQUE, cours de Mme PRADEL
Partiel 14 juin 2001

Sont autorisés, outre une calculette et les tables statistiques : deux feuilles (recto verso) manuscrites remplies par l'étudiant, à sa convenance et de sa main, des formules qu'il a choisies.

*Il sera tenu compte de la justesse et de la pertinence des arguments justifiant les réponses
Les exercices sont indépendants les uns des autres*

Exercice 1

Soient $(Y_1, Y_2, \dots, Y_{25})$ des variables aléatoires indépendantes et identiquement distribuées selon une loi Normale d'espérance m et de variance σ^2 :

$$Y_1, Y_2, \dots, Y_{25} \approx i.i.d.N(m, \sigma^2)$$

1. Quel estimateur proposez vous pour l'espérance m ?
2. Quelle est la loi de cet estimateur ?
3. Si nous pouvons assurer que la variance est connue, $\sigma^2 = 2.25$, calculez un intervalle de confiance 95% pour m .
4. La variance est en fait inconnue. Comment est modifié l'intervalle de confiance 95% précédent ?
5. Application numérique : donner l'estimation de m et les deux intervalles de confiance correspondant aux observations faites, qui ont pour moyenne et variance empiriques :

$$\bar{y} = \frac{1}{25} \sum_{i=1}^{25} y_i = 13.5 \text{ et } s^2 = \frac{1}{24} \sum_{i=1}^{25} (y_i - \bar{y})^2 = 1.69$$

Exercice 2

On désire tester l'efficacité d'un nouveau médicament censé corriger l'hypertension avec moins d'effets secondaires gênants que ceux existant sur le marché. Pour les médicaments déjà sur le marché, la probabilité d'apparition de ces effets gênants, notée p_a , est connue : $p_a = 0.60$. Le gain apporté par le nouveau médicament ne sera considéré comme décisif que si la probabilité d'apparition qui lui correspond est inférieure à $p_e = 0.54$. Pour simplifier, nous supposons que la probabilité d'apparition des effets secondaires avec le nouveau médicament ne peut prendre que l'une des deux valeurs p_a ou p_e .

1. Vous êtes employé par le laboratoire pharmaceutique qui a mis au point le médicament, et vous mettez en place un essai thérapeutique en administrant ce médicament à 80 patients, chez qui vous notez l'apparition ou l'absence des effets secondaires. Décrivez le problème de test qui se pose à vous, en précisant quelle est votre hypothèse de base.
2. Construisez le test optimal de seuil 5%.
3. Calculez la puissance de votre test
4. Quelle est la conclusion que vous tirez de l'observation des 80 patients de l'essai thérapeutique sachant que 47 d'entre eux ont éprouvé des effets secondaires gênants ?
5. Vous présentez votre dossier à un responsable de l'Agence pour le Médicament, afin d'obtenir l'autorisation de mise sur le marché et le remboursement de votre produit par la Sécurité Sociale (ce qui ne sera accepté que si le gain apporté par votre produit est décisif). Quelle est l'hypothèse de base de ce responsable ? Est-il d'accord avec votre conclusion ?

6. Calculez la règle d'acceptation du médicament par l'administration (qui fixe un test de seuil 5%)
7. Chez combien de patients au maximum aurait-il fallu constater d'effets secondaires pour que le médicament soit accepté ?
8. Que comptez vous faire ?

Exercice 3

Nous cherchons à expliquer les variations d'une variable macroéconomique annuelle Y en fonction de deux variables explicatives x et z . Nous avons pu observer $T = 25$ tirages précédents de Y correspondant à différentes valeurs des variables explicatives et nous nous plaçons dans le cadre d'un modèle linéaire standard normal :

$Y_1 \approx N(ax_1 + bz_1 + c; \sigma^2), Y_2 \approx N(ax_2 + bz_2 + c; \sigma^2), \dots, Y_{25} \approx N(ax_{25} + bz_{25} + c; \sigma^2)$
indépendantes.

La régression par les Moindres Carrés Ordinaires (MCO) de y sur x et z avec constante fournit les résultats suivants :

AJUSTEMENT	1	DEPENDENT VARIABLE	y
SMPL	1 25	TOTAL OBSERVATIONS	25
FISHER global	233.970	SSR	0.1834374
DURBIN-WATSON	2.382		

VARIABLE	COEFFICIENT	STAND. ERROR	T-STATISTIC	PROB
x	0.34222	0.3393	1.0087	0.3241
z	0.10551	0.1695	0.6224	0.5401
constante	1.05657	0.0783	13.493	< 0.0001

1. L'hypothèse d'absence d'autocorrélations des observations est-elle acceptable (précisez le test utilisé et le seuil choisi) ?
2. Les deux variables x et z sont-elles globalement explicatives au seuil de 5% ?
3. Effectuer les tests individuels de significativité des variables du modèle : quelles sont les ou la variable significativement explicative(s) au seuil de 5% ?
4. Comment expliquez vous les résultats apparemment contradictoires des questions 2 et 3?

Exercice 4

Une enquête effectuée auprès de 808 entreprises françaises du secteur “industries” nous a permis de construire le tableau de répartition de ces entreprises selon deux variables. La première variable, notée X, traduit l’activité innovante de l’entreprise selon trois modalités :

- X = 0 : “pas d’innovation”,
- X = 1 : “innovante à la création”
- X = 2 : “innovante non à la création mais durant son activité”.

La seconde, notée Y, traduit le capital humain du créateur selon 6 modalités :

- Y = 0 : “expérience professionnelle de moins de 3 ans et sans diplômes”,
- Y = 1 : “expérience professionnelle de moins de 3 ans et BEPC ou CEP”,
- Y = 2 : “expérience professionnelle de moins de 3 ans et Bac, BTS ou plus”,
- Y = 3 : “expérience professionnelle ≥ 3 ans et sans diplômes”
- Y = 4 : “expérience professionnelle ≥ 3 ans et BEPC ou CEP”,
- Y = 5 : “expérience professionnelle ≥ 3 ans et Bac, BTS ou plus”.

Le tableau des effectifs observés est le suivant :

$x \backslash y$	0	1	2	3	4	5	Total
0	101	121	64	78	98	57	519
1	11	15	13	9	14	14	76
2	26	37	26	37	58	29	213
Total	138	173	103	124	170	100	808

Afin de tester, l’indépendance des deux variables, on effectue un test du chi-deux.

1. Calculer l’effectif théorique correspondant à la case “X = 1, Y = 2”. En déduire que tous les effectifs théoriques seront assez grands pour pouvoir utiliser la loi de chi-deux dans le calcul du test.
2. Calculer la contribution δ_{12} de l’effectif observé dans cette case à la statistique du chi-deux.
3. Tous calculs faits, la statistique observée est $\Delta = 18.65$, et la p-value associée est $PROB = 0.0550$.

Au seuil de 10% , quelle conclusion tirez-vous de ces observations ?

1. Extrait de la table du chi-deux :

d.liberté	2	3	4	5	6	7	8	9
p=0,10	4.605	6.251	7.779	9.236	10.645	12.017	13.362	14.684
d.liberté	10	11	12	13	14	15	16	17
p=0,10	15.987	17.275	18.549	19.812	21.064	22.307	23.542	24.769