TUTORAT UE4 2011-2012 – Biostatistiques
Séance n°3 – Semaine du 17/10/2011

Lois de probabilités continues, estimations – Sabatier
Tests statistiques – Molinari

Séance préparée par Emmanuelle FAROUZ et Léo BARDOT

QCM n°1 : **L'inférence statistique :**

a) L'inférence statistique permet, connaissant les paramètres de la population, d'estimer ceux de notre échantillon
b) La base de toute estimation repose sur la réalisation d'un échantillonnage adéquat.
c) Le risque d'erreur que l'on s'octroie sur l'estimation du paramètre à calculer est fixé avant le calcul, c'est-à-dire a postériori
d) Un estimateur E(T) d'un paramètre θ est qualifié de « bon » dès lors qu'il est non biaisé (E(T) > θ) et convergent ($\lim_{n->+\infty} E[(T_n - \theta)^2] = 0$)
e) Par exemple : $\hat{S}_n^{\,2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ est un bon estimateur de la variance estimée dans la population
f) Toutes les propositions précédentes sont fausses.

QCM n°2 : **Intervalle de confiance :**

a) Un intervalle de confiance donné au risque α comporte dans 1-2 α pourcent des cas (le risque étant choisi en symétrique dans un intervalle de confiance) la valeur calculée.
b) La largeur de l'intervalle de confiance choisi dépend de certains paramètres de l'échantillon
c) La largeur de l'intervalle de confiance choisi est strictement indépendante de tous les paramètres de l'échantillon
d) Le calcul d'un intervalle de confiance ne repose sur l'utilisation de la seule table de la loi normale centrée réduite.
e) Le calcul de l'intervalle de confiance d'une loi normale n'est possible que si n>30.
f) Toutes les propositions précédentes sont fausses.

QCM n°3 : **Une enquête est effectuée auprès de 10 étudiants de Montpellier ne vivant pas chez leurs parents afin de savoir quel budget chacun alloue à son alimentation quotidienne.**
On trouve ainsi ces 10 réponses : 6€ ; 7,5€ ; 10€ ; 5€ ; 4,5€ ; 5,5€ ; 8€ ; 7€ ; 8€ ; 5,5€

a) La moyenne de l'échantillon est de 6,7 €.
b) L'écart-type estimé dans la population est de 1,62 €.
c) Il est nécessaire de supposer que le budget quotidien alloué aux dépenses quotidiennes de nourriture suit une Loi Normale afin de déterminer un intervalle de confiance dans la population.
d) L'intervalle de confiance de la moyenne estimée de ce budget quotidien dans la population, au seuil de 5%, est [5,48 ; 7,92].

e) L'intervalle de confiance de la moyenne estimée de ce budget quotidien dans la population, au seuil de 10%, est [5,81 ; 7,58].
f) Toutes les propositions précédentes sont fausses.

QCM n°4 : **A l'entrée du stade de La Mosson, accueillant la rencontre de football Montpellier-Paris, on interroge 80 spectateurs sur leur pronostic pour le match à venir. Ainsi sur les 80 personnes interrogées 47 prédisent une victoire de Montpellier** (*malheureusement à tord…)* **:**

a) Afin de calculer l'intervalle de confiance de la proportion de gens prédisant une victoire de Montpellier dans le stade (sur la population) il faut au préalable supposer que le pronostic de chaque personne suit une Loi Normale.
b) Au seuil de 5% l'intervalle de confiance de la proportion de spectateurs prédisant une victoire de Montpellier est [0,4796 ; 0,6953].
c) Au seuil de 5% l'intervalle de confiance de la proportion de spectateurs prédisant une victoire de Montpellier est [0,4896 ; 0,6953].
d) Au seuil de 10% l'intervalle de confiance de la proportion de spectateurs prédisant une victoire de Montpellier est [0,4979 ; 0,6780].
e) Au seuil de 10% l'intervalle de confiance de la proportion de spectateurs prédisant une victoire de Montpellier est [0,4969 ; 0,6780].
f) Toutes les propositions précédentes sont fausses.

QCM n°5 : (suite du 4): **Toujours dans ce même stade mais cette fois ci à la sortie on interroge 15 spectateurs sur leur âge, noté X.**
Dans ce QCM on supposera que l'âge du spectateur suit une Loi Normale.
On donne : $\sum_{i=1}^{15} x_i = 375$ et $\sum_{i=1}^{15} x_i^{\,2} = 10635$

a) La moyenne d'âge dans cet échantillon est de 25 ans.
b) La variance dans l'échantillon est égale à 84 ans.
c) L'écart-type estimé de la population est alors égal à 9,49 ans à 10^{-2} près.
d) L'intervalle de confiance à 80% de la variance dans la population est [50,785 ; 170,785].
e) L'intervalle de confiance à 80% de l'écart-type dans la population est [7,73 ; 12,72].
f) Toutes les propositions précédentes sont fausses.

QCM n°6 : **Soit un échantillon représentatif de 100 hommes jardiniers-paysagistes. On s'intéresse à la répartition de la durée du travail (en heures) lors de la rénovation d'un jardin. On se propose de calculer la moyenne expérimentale m et la variance expérimentale S² de la durée du travail de ces hommes.**
On donne $\sum xi=700$ et $\sum xi^2=5790$:

a) m est un estimateur ponctuel de la moyenne
b) $6 \le m \le 9$ et $2 \le s^2 \le 3$
c) $7 \le m \le 8$ et $2 \le s \le 3$
d) $7 \le m \le 9$ et $8 \le s^2 \le 9$
e) $8 \le m \le 9$ et $7 \le s^2 \le 9$
f) Toutes les propositions précédentes sont fausses.

QCM n°7 : **Avant l'élection des Z'élus on souhaite effectuer un sondage d'opinion auprès des PACES. Il n'y a que deux candidats, et on suppose que tous les participants au sondage voteront pour l'un ou l'autre des candidats. On note p la proportion de votes pour un des candidats. Si on réalise une approximation gaussienne, combien faut-il interroger de PACES pour qu'ils aient 95% de chances d'estimer p à 2% près :**

a) 996
b) 1010
c) 1563
d) 2401
e) 2542
f) Toutes les propositions précédentes sont fausses.

QCM n°8 : **Concernant le principe général des tests statistiques:**

a) L'hypothèse nulle H_0 est définie a priori et l'hypothèse complémentaire alternative H_1 est définie a posteriori
b) Les hypothèses ne portent que sur les paramètres d'une ou plusieurs distributions.
c) Le risque de première espèce est le risque de ne pas rejeter l'hypothèse H_0 sachant qu'elle est fausse. On l'appelle le risque bêta (manque de puissance).
d) La puissance est définie comme la probabilité complémentaire à 1 de bêta, c'est-à-dire la probabilité de rejeter l'hypothèse H_0 sachant que H_0 est fausse.
e) On rejette H_0 (Hypothèse nulle) si p-value (inférieure ou égale) à alpha ou si t_α (inférieur) à t_{obs}
f) Toutes les propositions précédentes sont fausses.

QCM n°9 : **Concernant le principe général des tests statistiques:**

a) La statistique de test et le choix de test statistique sont intimement liés.
b) Les risques de première et de deuxième espèce ne sont pas maitrisables simultanément. Le risque de première espèce est choisi comme règle de décision.
c) t_{obs} est la valeur de la statistique de test calculée sur l'échantillon, alors t_α que se lit dans une table statistique.
d) Dans un test statistique, la statistique de test est une variable aléatoire dont on connaît la loi sous H_0.
e) Cette statistique de test dépend du risque alpha.
f) Toutes les propositions précédentes sont fausses.

QCM n°10 : **Le risque de première espèce :**

a) Est noté α
b) Est défini a posteriori
c) Correspond au risque que l'on accepte de prendre quand on rejette à tort l'hypothèse nulle.
d) Est le risque que la conclusion soit fausse si l'on rejette l'hypothèse nulle
e) Est la probabilité d'accepter l'hypothèse alternative sachant que H_0 est vraie.
f) Toutes les propositions précédentes sont fausses.

QCM n°11 : **On a un groupe de 1000 patients tirés au sort. Ce groupe est composé de 528 patients et 472 patientes. On se demande s'il y a significativement plus d'hommes que de femmes dans la population d'où ce groupe est issu :**

a) Il faut réaliser un test bilatéral.
b) Il faut réaliser un test unilatéral.
c) L'hypothèse nulle est qu'il y a autant de femmes que d'hommes.
d) On doit comparer 52,8% à 50% par un test statistique.
e) Les données sont appariées.
f) Toutes les propositions précédentes sont fausses.

QCM n°12 : **On espère qu'un nouvel antiviral diminuera le nombre de jours avec symptômes cliniques chez les sujets infectés par le virus de la grippe. On étudie un échantillon de 100 personnes non traités tirés au hasard atteints d'une grippe confirmée par un test virologique. On trouve que sur cet échantillon, le nombre moyen de jours avec symptômes est de 4,75 jours et que la variance vaut 1.**
Après avoir observé les résultats précédents, il est décidé de choisir au hasard à partir de la même population de malades 100 nouveaux malades et de les traiter avec l'antiviral .On observe chez ces malades traités une durée moyenne de signes cliniques m1 de 4,2 jours et un écart type 1,732 jours:

a) Le test statistique de comparaison des moyennes m=4,75 et m1=4,2 traitées permet de rejeter l'hypothèse nulle au risque alpha à 5%.
b) Le test statistique de comparaison des moyennes m=4,75 et m1=4,2 ne permet pas de rejeter l'hypothèse nulle
c) On ne peut pas conclure que l'antiviral est efficace à cause du risque de première espèce fixé à 5%
d) On ne peut pas conclure que l'antiviral est efficace car il ne s'agit pas d'un essai randomisé
e) On peut conclure statistiquement que l'antiviral a une efficacité sur la durée moyenne des signes cliniques.
f) Toutes les propositions précédentes sont fausses.

QCM n°13 : (suite du 12) **On réalise maintenant un essai randomisé, en double aveugle, comportant un bras placebo sur deux groupes de 225(=15²) sujets. Dans le groupe placebo, on trouve m_A=4,7 et s^2_A=1,5 ; dans le groupe traité, on trouve m_B=4,7 et s^2_B=2,5:**

a) La valeur $|z|$ calculée dans le test de comparaison est inférieure à 2.
b) La valeur $|z|$ calculée dans le test de comparaison est comprise entre 2 et 4.
c) Pour cet essai on peut utiliser un test du Chi² de MacNémar.
d) On rejette l'hypothèse nulle. L'antiviral diminue le nombre de jours avec symptômes
e) On ne rejette pas l'hypothèse nulle. On ne peut pas dire que l'antiviral diminue le nombre de jours avec symptômes.
f) Toutes les propositions précédentes sont fausses.

QCM n°14 : **Un statisticien soupçonne un dé d'être pipé. Plus précisément, il pense que ce dé fait « 6 » plus souvent qu'un dé ordinaire, c'est-à-dire plus souvent qu'une fois sur 6. Pour le prouver il va lancer le dé suspect 10 fois de suite, noter la proportion de « 6 » et construire un test unilatéral destiné à prouver que cette proportion est supérieure à 1/6. En notant p la probabilité d'obtenir « 6 » avec le dé, et p_0 la proportion de « 6 » qu'il va obtenir en lançant 10 fois le dé, comment va-t-il formuler les hypothèses du test ?**

a) $H_0 : p > 1/6$
b) $H_0 : p_0 = 1/6$
c) $H_1 : p_0 > 1/6$
d) $H_1 : p=1/6$
e) $H_1 : p> 1/6$
f) Toutes les propositions précédentes sont fausses.