L2 Psychologie **Statistique** **2005/2006**

TD 1: généralités sur les tests

1 Bref rappels de cours

Les tests statistiques (dont le test du khi^2, mais aussi les tests de comparaisons de moyennes) fonctionnent tous sur le même principe.

On oppose deux *hypothèses* :

- L'hypothèse nulle H_0
- L'hypothèse alternative H_1

On vérifie les conditions d'applicaton.

On calcule une statistiques T connue sur l'échantillon : c'est la *variable de décision.*

On détermine une *zone de rejet*, pour un risque α fixé.

Si la valeur trouvée pour T, disons t, est dans la zone de rejet, alors on rejette H_0 et on conclut H_1. Dans le cas contraire, on ne peut pas conclure.

2 Exercices

Exercice 1 *"Les tests d'hypothèses sont de simples* aides à la décision *et ne démontrent rien".*

Discutez.

Exercice 2 *"Les tests d'hypothèses sont des méthodes de* démonstration risquée*".*

Discutez.

Exercice 3 *"On ne peut jamais accepter H_0".*

Discutez.

Exercice 4 *Si un test effectué au risque de 1% est concluant, on peut dire que la probabilité que H_1 (hypothèse alternative) soit vraie est de 99% au moins.*

Vrai ou faux ?

Exercice 5 *Les rôles de H_0 et H_1 sont-ils symétriques ?*

Exercice 6 *Exprimez en fonction de α et β, risques de première et deuxième espèce*

1. *la probabilité de conclure H_1 si H_1 est vraie*
2. *et celle de ne pas rejeter H_0 si H_0 est vraie.*

Exercice 7 *Les valeurs obtenues sur un échantillon peuvent nous aider à choisir un test adapté, mais il faut refaire toute l'expérience avec un autre échantillon.*

Vrai ou faux ?

Exercice 8 *On veut tester une hypothèse H_0 contre une hypothèse H_1, et l'on dispose pour cela de 5 tests différents. Puisqu'il est difficile de choisir entre les 5 tests, on décide de faire les 5 tests (avec un risque de 1%). Certains de ces tests sont concluants. On peut donc dire au risque de* 1% *que H_1 est vraie contre H_0.*

Vrai ou faux ?

Exercice 9 *On veut tester une hypothèse H_0 contre une hypothèse H_1, et l'on dispose pour cela de 5 tests différents. Puisqu'il est difficile de choisir entre les 5 tests, on décide de faire les 5 tests (avec un risque de 1%). On utilise d'abord le premier test, qui échoue. Puis le second, qui est concluant. On peut donc dire au risque de* 2% *que H_1 est vrai contre H_0.*

Vrai ou faux ?

Exercice 10 *Un test d'hypothèse est une forme affaiblie de* démonstration par l'absurde.

Expliquez.

3 Un peu plus concret

Exercice 11 *On dispose d'une variable X dont on sait qu'elle est* constante. *Proposez un test permettant de décider entre*

$$H_0 : X = 0$$

et

$$H_1 : X \neq 0.$$

Quelles sont les caractéristiques de ce test ?

Exercice 12 *On dispose d'une variable X* bornée. *On veut tester l'hypothèse*

$$H_0 : X \leq 10 \text{ pour tout individu}$$

contre

$$H_1 : \max X > 10.$$

Proposez un test.

Exercice 13 *On dispose d'un dé dont on veut savoir s'il est truqué ou non. On propose de tester*

$$H_0 : \text{ le dé n'est pas truqué}$$

contre

$$H_1 : \text{ le dé est truqué}$$

de la manière qui suit. On lance deux fois le dé. S'il tombe deux fois de suite sur la même valeur 6, on dit qu'il est truqué. Dans le cas contraire, on ne rejette pas H_0. Quel est le risque associé à ce test ?

Exercice 14 *On reprend l'idée de l'exerice précédent , mais avec deux 5 au lieu de deux 6. On conclut donc si le dé tombe deux fois sur 5. Quel est le risque associé à ce* "test des deux 5" *? Même question avec le "test des deux 3", etc.*

Exercice 15 *Pourquoi le* test des deux 6 *est-il plus convaincant que celui des* deux 4 *? Ce deuxième test est-il valable ? Peut-on calculer sa puissance ?*

Exercice 16 *On propose un test combiné, utilisant les tests des deux 6, des deux 5, etc.. On procède comme suit : on choisit arbitrairement, et avant le début de l'expérience, un nombre entre 1 et 6. Si par exemple le 3 est choisi, alors on décide d'utiliser le test des deux 3 pour tester le dé douteux. On lance ensuite deux fois le dé douteux, et on conclut selon le test des deux 3.*

1. *Cette méthode est-elle acceptable ?*
2. *Quel est son risque ?*
3. *Quelle est sa puissance si l'on suppose qu'un dé truqué tombe systématiquement sur 6 ?*

Exercice 17 *On procède un peu comme à l'exercice précédent, mais en lançant d'abord deux fois le dé douteux avant de choisir le nombre entre 1 et 6. Si par exemple on choisit, a posteriori, le 3, on utilisera le test des deux 3 avec les deux tirages initiaux du dé.*

1. *Cette méthode est-elle acceptable ?*
2. *Quel est son risque ?*
3. *Concluez.*

Exercice 18 *Un philosophe s'étonne : "Il est stupéfiant que parmi toutes les planètes possibles, l'homme se soit développé précisément sur la terre !" Critiquez cette remarque du point de vue des tests statistiques.*

4 Pour aller plus loin

Les exercices de cette section sont réservés à ceux qui souhaitent s'initier aux statistiques mathématiques. Il n'est pas utile de connaître les points développer ici pour réussir !

Exercice 19 *On souhaite tester sur une variable X de paramètre θ l'hypothèse*

$$H_0 : \ \theta = \theta_0$$

contre

$$H_1 : \theta = \theta_1.$$

Pour une observation

$$x_1, .., x_n,$$

on notera

$$L_0 (x) = P (X_1 = x_1 \wedge ... \wedge X_n = x_n | H_0)$$

et

$$L_1 (x) = P (X_1 = x_1 \wedge ... \wedge X_n = x_n | H_1),$$

qui sont les vraissemblances *de θ_0 et θ_1. Pour quelles valeurs de L_0 et L_1 considèrera-t-on que l'expérience est concluante ?*

(Répondez de manière intuitive en français courant)

Exercice 20 *On se place à nouveau dans les hypothèses de l'exercice 19. La* méthode de Neyman et Pearson *consiste à prendre comme zone de rejet :*

$$W = \left\{ x, \frac{L_1 (x)}{L_0 (x)} > k \right\}$$

où k est une constante.

Cela fait-il sens ?

Exercice 21 *Si l'on a une hésitation sur H_0 et H_1, prendra-t-on plutôt pour H_0 une hypothèse simple*[1] *ou composite ?*

[1] Une hypothèse simple est de la forme

$$\theta = 0$$

par exemple. Au contraire, une hypothèse comme

$$\theta > 0$$

est composite, parce qu'elle renvoit à plusieurs valeurs (toutes les valeurs positives).