Université Paul Valéry
Arts - Lettres - Langues
Sciences humaines & sociales

PSYCHOLOGIE - Licence 3
E56XPM
Statistique appliquée à la psychologie 4
Examen du 30 janvier 2008

NOM	**:**
PRÉNOM	**:**
N° Carte d'étudiant	**:**
Signature	**:**

Durée : 2 heures.
• Matériel autorisé : une page manuscrite recto/verso, calculatrice et tables de lois.
• Téléphone portable interdit.
• Les calculatrices ne doivent en aucun cas circuler ou être échangées entre les étudiants.

Consignes pour le choix des risques : dans tout l'examen
Pour la construction d'un intervalle de dispersion on choisira un risque $\alpha = 1.2$ %.
Pour la construction d'un intervalle de confiance on choisira un risque $\alpha = 2.8$ %.
Tous les tests statistiques seront construits avec une erreur de 1re espèce de 1 %.

Les exercices 1 et 2 sont 2 questions de cours. Donner une réponse claire et succincte en précisant chacun des objets que vous introduisez.

Exercice 1 : Dans une séquence ADN de longueur 3500, formée des lettres A, C, G, T, on compte le nombre d'occurrences de chacune des lettres. Et on cherche à éprouver l'hypothèse que chaque lettre apparaît de façon équiprobable.
Construire le test approprié en spécifiant clairement les hypothèses, la statistique du test et la zone de rejet.

Exercice 2 : On dispose de 200 observations $x_1, x_2, ..., x_{200}$ de la variable aléatoire continue X.
Donner l'expression de l'intervalle de confiance pour le paramètre de l'espérance mathématique : $E(X)$.

Exercice 3 : On suppose que la distribution de la variable Y est $\mathcal{N}(2, \sigma^2 = .0625)$ mais on a de bonnes raisons de penser qu'en réalité le paramètre de l'espérance mathématique a diminué.

1. Construire le test statistique qui permette d'éprouver cette idée avec un échantillon de taille n : donner les hypothèses, la zone de rejet.
2. Calculer la borne de rejet associée à un échantillon de 36 individus.

Exercice 4 : On a mesuré sur 380 individus 2 variables Z et T.
On obtient les résultats suivants : $\text{cov}(Z, T) = 136$ et $\sigma_Z = 14$ et $\sigma_T = 19$.
Peut on considérer qu'il existe une liaison linéaire entre Z et T ?

Exercice 5 : Dans la séquence ADN formée des lettres A, C, G et T, on suppose que ces lettres apparaissent de façon équiprobable et indépendante.

1. Exprimer par une variable aléatoire (dont on donnera la loi) le fait qu'une lettre est un "A".
2. Donner l'expression de l'intervalle de dispersion du nombre de "A" dans un séquence de longueur n.
3. Calculer les bornes de cet intervalle pour une séquence de longueur 350.

Exercice 6 : Dans un service hospitalier, on a l'habitude de déclarer aux patients atteints d'une certaine pathologie que l'examen qu'ils doivent subir comporte un risque de complication de 15 %. Cette valeur fait référence dans le milieu médical depuis plusieurs années.

1. Il y a en France de l'ordre de 1500 personnes qui doivent subir cet examen chaque année. Donner sous forme d'intervalle, le taux de complications auquel on peut s'attendre au cours d'une année.

2. Une nouvelle procédure vient récemment d'être mise au point de sorte à diminuer ce risque de complication. On envisage donc de procéder à une expérimentation sur 200 patients afin de confirmer cette hypothèse. Proposer une modélisation et une règle de décision qui permettent de répondre à la question.

3. Quelle devrait être la conclusion à apporter à cette expérimentation lorsque on observe 22 patients (sur 200) atteints de complication.

4. Quelle devrait être la conclusion à apporter à cette expérimentation lorsque on observe 15 patients (sur 200) atteints de complication.

5. À partir des conclusions de la question précédente, discuter de la pertinence de construire un intervalle pour la valeur du nouveau risque de complication que l'on pourrait annoncer aux patients. Calculer alors les bornes de cet intervalle.