STT 3500/STT 6230
Méthodes non paramétriques/Méthodes non paramétriques avancées
Devoir 2
Hiver 2020

À moins d'un avis contraire, les exercices proviennent de l'ouvrage *Nonparametric Statistical Inference*, cinquième édition, de J. D. Gibbons et S. Chakraborti. Assurez-vous d'avoir la bonne édition du livre.

(1) (STT-3500/STT-6230) Problème 5.2

(2) (STT-3500/STT-6230) Problème 5.4

(3) (STT-3500/STT-6230) Problème 5.5

(4) (STT-3500/STT-6230) Problème 5.6

(5) (STT-3500/STT-6230) Problème 5.7

(6) (STT-3500/STT-6230) Problème 5.14

(7) (STT-3500/STT-6230) On a demandé aux étudiants d'une classe combien ils ont de frères et de soeurs et si leur ville natale est urbaine ou rurale. Les résultats sont dans le tableau suivant:

Rural	3	2	1	1	2	1	3	2	2	2	2	5	1	4	1	1	1
	1	6	2	2	2	1	1										
Urbain	1	0	1	1	0	0	1	1	1	8	1	1	1	0	1	1	2

On vous invite à répondre aux questions suivantes.

(a) Tester s'il existe une différence significative entre les deux zones en utilisant le test de la somme des rangs de Wilcoxon. Utilisez le niveau de signification 5%.

(b) Tester s'il existe une différence significative entre les deux zones en utilisant le test-t pour deux échantillons. Utilisez le niveau de signification 5%.

(c) Utiliser l'approximation pour de grands échantillons pour la statistique de Wilcoxon afin de calculer la valeur-p. Énoncez votre conclusion.

(d) Comparer les résultats dans les parties (a), (b) et (c). Discuter.

(8) (STT-6230) On dispose de n paires de sujets (X_i, Y_i) indépendants de loi conjointe $F_{XY}(x, y)$, et l'on pose les différences $D_i = Y_i - X_i$. Les variables X_i forment le groupe contrôle, alors que les Y_i recoivent un traitement. Posons $F_D(d) = P(D_i \leq d) = L(d)$.

a) Montrer que sous l'hypothèse nulle d'absence de traitement la distribution F_D est symétrique autour de zéro, dans le sens que $L(d) = 1 - L(-d)$.

b) Une alternative importante est l'hypothèse d'additivité, qui fait en sorte que sous l'alternative le traitement ajoute une même constante Δ au contrôle. Soit $E(d)$ la loi sous l'hypothèse nulle de D_i. Supposons que l'effet de traitement est Δ. Montrer que la distribution de $(Y_i - \Delta) - X_i$ est E, et qu'en conséquence la loi de D_i satisfait:

$$P(D_i \leq d) = E(d - \Delta).$$

c) Posons $p = P(D > \Delta_0)$ et supposons que la loi de D satisfait $P(D_i \leq d) = E(d - \Delta)$. Supposons que E est continue et symétrique autour de zéro. Montrer que

$$p = E(\Delta - \Delta_0).$$

d) Si E est la loi normale avec moyenne zéro et variance σ^2, montrer que

$$p = \Phi\{(\Delta - \Delta_0)/\sigma\}.$$

(9) (STT-6230) Supposons que le modèle additif comme dans le problème précédent. Si nous disposons de $n = 15$ observations et que la loi du phénomène sous-jacent est normale, avec variance $\sigma^2 = 16$, déterminer la puissance approximative du test de rang signé de Wilcoxon du niveau $\alpha = .076$ pour le test de $H_0 : \Delta = 0$ contre l'hypothèse alternative $H_1 : \Delta > 0$ lorsque l'effet du traitement est $\Delta = 1.25$.