

# Test Tests Non-Paramétriques - Niveau 1 (Intermédiaire)

## UniBank Haiti - Data Analyst

**Durée:** 40 minutes

**Questions:** 25

**Niveau:** Intermédiaire

**Sujets:** Mann-Whitney, Wilcoxon, Kruskal-Wallis, Chi-carré, Spearman

---

**Q1. Quand devez-vous utiliser un test non-paramétrique au lieu d'un test paramétrique?**

- A) Toujours, ils sont meilleurs
- B) Quand les données sont normalement distribuées
- C) Quand les hypothèses de normalité ou homogénéité des variances sont violées
- D) Uniquement pour les grandes bases de données

**Réponse:** C) Quand les hypothèses de normalité ou homogénéité des variances sont violées

*Les tests non-paramétriques ne font pas d'hypothèses sur la distribution des données. Ils sont appropriés quand les données sont asymétriques, ordinales, ou de petite taille.*

---

**Q2. Vous comparez les montants de transactions entre clients Retail et Premium. Le test de Shapiro-Wilk rejette la normalité ( $p < 0.01$ ). Quel test utiliser?**

- A) t-test pour échantillons indépendants
- B) Mann-Whitney U
- C) ANOVA
- D) Chi-carré

**Réponse:** B) Mann-Whitney U

*Mann-Whitney U est l'alternative non-paramétrique au t-test pour 2 groupes indépendants. Il compare les rangs plutôt que les moyennes, sans hypothèse de normalité.*

---

**Q3. Quelle est l'hypothèse nulle du test de Mann-Whitney U?**

- A) Les moyennes des deux groupes sont égales
- B) Les distributions des deux groupes sont identiques (mêmes tendances centrales)
- C) Les variances sont égales
- D) Les données sont normales

**Réponse:** B) Les distributions des deux groupes sont identiques (mêmes tendances centrales)  
*Mann-Whitney teste si les deux groupes proviennent de la même distribution. En pratique, cela revient souvent à comparer les médianes.*

---

**Q4. Le test de Wilcoxon Signed-Rank est utilisé pour:**

- A) Comparer deux groupes indépendants
- B) Comparer des mesures appariées (avant/après) sur les mêmes sujets
- C) Comparer trois groupes ou plus
- D) Tester la normalité

**Réponse:** B) Comparer des mesures appariées (avant/après) sur les mêmes sujets

*Wilcoxon Signed-Rank est l'alternative non-paramétrique au t-test apparié. Il compare les différences intra-sujet (ex: satisfaction avant/après formation).*

---

**Q5. Vous analysez la satisfaction client (échelle 1-5) avant et après une campagne. Les données sont ordinales. Quel test choisir?**

- A) t-test apparié
- B) Wilcoxon Signed-Rank
- C) Mann-Whitney U
- D) ANOVA

**Réponse:** B) Wilcoxon Signed-Rank

*Les données ordinales (échelle 1-5) ne peuvent pas être traitées comme continues. Le Wilcoxon Signed-Rank est approprié pour les comparaisons appariées sur données ordinales.*

---

**Q6. Vous comparez les montants de prêts accordés entre 4 agences. Les données sont asymétriques. Quel test utiliser?**

- A) ANOVA one-way
- B) t-test
- C) Kruskal-Wallis
- D) Chi-carré

**Réponse:** C) Kruskal-Wallis

*Kruskal-Wallis est l'alternative non-paramétrique à l'ANOVA pour comparer 3+ groupes indépendants sans hypothèse de normalité.*

---

**Q7. Le test de Kruskal-Wallis donne  $p = 0.02$ . Que concluez-vous?**

- A) Tous les groupes sont différents
- B) Au moins un groupe diffère significativement des autres
- C) Aucune différence
- D) Le test est invalide

**Réponse:** B) Au moins un groupe diffère significativement des autres

*Comme l'ANOVA, Kruskal-Wallis est un test omnibus - il indique une différence globale. Des tests post-hoc (Dunn) sont nécessaires pour identifier quels groupes diffèrent.*

---

**Q8. Quel test post-hoc est approprié après un Kruskal-Wallis significatif?**

- A) Tukey HSD
- B) Test de Dunn
- C) Test de Student
- D) Chi-carré

**Réponse:** B) Test de Dunn

*Le test de Dunn est le post-hoc standard pour Kruskal-Wallis. Il compare les rangs moyens entre paires de groupes avec correction pour comparaisons multiples.*

---

**Q9. Vous voulez tester si le type de compte (Épargne, Courant, DAT) est indépendant du segment client (Retail, Premium). Quel test?**

- A) t-test
- B) ANOVA
- C) Chi-carré d'indépendance
- D) Corrélation de Pearson

**Réponse:** C) Chi-carré d'indépendance

*Le Chi-carré d'indépendance teste l'association entre deux variables catégorielles. Il compare les fréquences observées aux fréquences attendues sous indépendance.*

---

**Q10. Le test Chi-carré nécessite que les fréquences attendues soient:**

- A) Exactement égales aux fréquences observées
- B) Supérieures ou égales à 5 dans chaque cellule (règle générale)

- C) Inférieures à 1
- D) Normalement distribuées

**Réponse:** B) Supérieures ou égales à 5 dans chaque cellule (règle générale)

*La règle des 5 stipule que chaque cellule doit avoir une fréquence attendue  $\geq 5$  pour que l'approximation Chi-carré soit valide. Sinon, utiliser Fisher exact.*

---

**Q11. Quelle est la différence entre la corrélation de Pearson et celle de Spearman?**

- A) Pearson mesure les relations linéaires, Spearman les relations monotones
- B) Spearman mesure les relations linéaires, Pearson les relations monotones
- C) Aucune différence
- D) Pearson est toujours plus précis

**Réponse:** A) Pearson mesure les relations linéaires, Spearman les relations monotones

*Pearson quantifie la relation linéaire entre variables continues normales. Spearman (basé sur les rangs) détecte toute relation monotone, même non-linéaire.*

---

**Q12. Les revenus bancaires sont souvent asymétriques à droite. Quelle corrélation privilégier pour analyser la relation revenu-montant du prêt?**

- A) Pearson
- B) Spearman
- C) Chi-carré
- D) Point-biserial

**Réponse:** B) Spearman

*L'asymétrie et les potentiels outliers des revenus bancaires violent les hypothèses de Pearson. Spearman, basée sur les rangs, est plus robuste.*

---

**Q13. Vous obtenez  $\rho$  (Spearman) = 0.75 entre l'ancienneté client et le nombre de produits détenus. Comment interpréter?**

- A) Corrélation négative forte
- B) Corrélation positive forte - clients plus anciens tendent à avoir plus de produits
- C) Pas de corrélation
- D) Relation causale confirmée

**Réponse:** B) Corrélation positive forte - clients plus anciens tendent à avoir plus de produits  
 $\rho = 0.75$  indique une corrélation positive forte. Plus l'ancienneté augmente, plus le nombre de produits augmente (en termes de rangs).

---

**Q14. Le coefficient Tau de Kendall diffère de Spearman par:**

- A) Aucune différence
- B) Kendall compare les paires concordantes vs discordantes, plus robuste pour petits échantillons
- C) Kendall ne fonctionne pas pour les données ordinales
- D) Spearman est toujours préférable

**Réponse:** B) Kendall compare les paires concordantes vs discordantes, plus robuste pour petits échantillons

*Kendall  $\tau$  compte les paires concordantes et discordantes. Il est plus robuste que Spearman pour les petits échantillons et interprétable en termes de probabilité.*

---

**Q15. Vous testez si la proportion de clients satisfaits diffère entre deux trimestres. Quel test non-paramétrique?**

- A) Mann-Whitney
- B) Test de McNemar
- C) Kruskal-Wallis
- D) Wilcoxon

**Réponse:** B) Test de McNemar

*Pour comparer des proportions appariées (mêmes clients à deux moments), McNemar est approprié. Il analyse les changements de catégorie.*

---

**Q16. Le test de Friedman est l'équivalent non-paramétrique de:**

- A) ANOVA one-way
- B) ANOVA à mesures répétées
- C) t-test
- D) Chi-carré

**Réponse:** B) ANOVA à mesures répétées

*Friedman compare 3+ conditions/temps sur les mêmes sujets sans hypothèse de normalité. C'est l'extension de Wilcoxon Signed-Rank à plus de 2 groupes appariés.*

---

**Q17. Vous analysez la performance de 10 gestionnaires de portefeuille sur 4 trimestres. Quel test non-paramétrique?**

- A) Kruskal-Wallis
- B) Friedman
- C) Mann-Whitney
- D) Chi-carré

**Réponse:** B) Friedman

*Les mêmes gestionnaires sont mesurés sur 4 trimestres (mesures répétées). Friedman est l'alternative non-paramétrique à l'ANOVA à mesures répétées.*

---

**Q18. Le V de Cramér après un Chi-carré est de 0.35. Comment l'interpréter?**

- A) L'effet est très faible
- B) Association modérée entre les deux variables catégorielles
- C) Association parfaite
- D) V de Cramér n'a pas d'interprétation

**Réponse:** B) Association modérée entre les deux variables catégorielles

*V de Cramér mesure la force d'association. Conventions: 0.1 = faible, 0.3 = modéré, 0.5 = fort. Donc 0.35 est une association modérée.*

---

**Q19. Quelle condition doit être vérifiée pour utiliser le test exact de Fisher au lieu du Chi-carré?**

- A) Grandes fréquences attendues
- B) Table 2x2 avec fréquences attendues < 5 dans au moins une cellule
- C) Données continues
- D) Échantillons appariés

**Réponse:** B) Table 2x2 avec fréquences attendues < 5 dans au moins une cellule

*Le test exact de Fisher est utilisé quand les hypothèses du Chi-carré sont violées (petits effectifs). Il calcule la probabilité exacte sans approximation.*

---

**Q20. Le test de Kolmogorov-Smirnov (K-S) compare:**

- A) Deux moyennes
- B) Deux distributions (si elles proviennent de la même population)
- C) Deux proportions
- D) Les rangs

**Réponse:** B) Deux distributions (si elles proviennent de la même population)

*K-S teste si deux échantillons proviennent de la même distribution. Il compare les fonctions de répartition cumulative (CDF).*

---

**Q21. Vous avez 15 observations par groupe. Pourquoi privilégier un test non-paramétrique?**

- A) Les tests paramétriques ne fonctionnent pas avec 15 observations
- B) Avec  $n < 30$ , la normalité est difficile à vérifier et les tests paramétriques moins robustes
- C) Les tests non-paramétriques sont toujours plus puissants
- D) Il n'y a pas de raison particulière

**Réponse:** B) Avec  $n < 30$ , la normalité est difficile à vérifier et les tests paramétriques moins robustes

*Le théorème central limite s'applique mieux pour  $n \geq 30$ . Avec de petits échantillons, les tests non-paramétriques sont souvent plus fiables.*

---

**Q22. Quel est le principal inconvénient des tests non-paramétriques par rapport aux paramétriques?**

- A) Ils ne fonctionnent jamais
- B) Ils sont moins puissants quand les hypothèses paramétriques sont satisfaites
- C) Ils nécessitent plus de données
- D) Ils sont plus difficiles à calculer

**Réponse:** B) Ils sont moins puissants quand les hypothèses paramétriques sont satisfaites

*Si les données sont vraiment normales, les tests paramétriques ont plus de puissance (probabilité de détecter un vrai effet). Les non-paramétriques "gaspillent" de l'information.*

---

**Q23. Le test de rang signé de Wilcoxon donne  $p = 0.08$  avec  $n = 12$  paires. Avec  $\alpha = 0.05$ , quelle conclusion?**

- A) Différence significative
- B) Pas de différence significative au seuil 5%
- C) Besoin de plus de données
- D) Le test est invalide

**Réponse:** B) Pas de différence significative au seuil 5%

*$p = 0.08 > \alpha = 0.05$ , donc on ne rejette pas  $H_0$ . Cependant, avec  $n = 12$ , la puissance est limitée. Le résultat pourrait devenir significatif avec plus de données.*

---

**Q24. Dans le contexte bancaire, pourquoi la corrélation de Spearman est-elle souvent préférée pour les données financières?**

- A) Elle est plus rapide à calculer
- B) Les données financières sont souvent asymétriques avec des outliers, Spearman est robuste
- C) Spearman fonctionne uniquement pour les données financières
- D) Pearson ne fonctionne pas avec l'argent

**Réponse:** B) Les données financières sont souvent asymétriques avec des outliers, Spearman est robuste

*Les revenus, montants de transactions, soldes ont généralement des distributions asymétriques (quelques très grandes valeurs). Spearman, basé sur les rangs, n'est pas affecté par ces extrêmes.*

---

**Q25. Vous effectuez 6 tests de Mann-Whitney pour comparer les montants entre différentes paires de segments. Comment corriger pour les comparaisons multiples?**

- A) Ne pas corriger
- B) Appliquer la correction de Bonferroni:  $\alpha/6 = 0.05/6 \approx 0.0083$
- C) Multiplier tous les  $p$  par 2
- D) Utiliser  $\alpha = 0.50$

**Réponse:** B) Appliquer la correction de Bonferroni:  $\alpha/6 = 0.05/6 \approx 0.0083$

*Avec 6 comparaisons, Bonferroni ajuste le seuil à  $0.05/6 \approx 0.0083$ . Un résultat n'est significatif que si  $p < 0.0083$ .*

---



## Résumé des Tests Non-Paramétriques

### Comparaison de 2 Groupes

Situation	Test Paramétrique	Test Non-Paramétrique
Indépendants	t-test	Mann-Whitney U
Appariés	t-test apparié	Wilcoxon Signed-Rank

### Comparaison de 3+ Groupes

Situation	Test Paramétrique	Test Non-Paramétrique
Indépendants	ANOVA	Kruskal-Wallis
Appariés	ANOVA mesures répétées	Friedman

### Corrélation

Type	Paramétrique	Non-Paramétrique
Linéaire	Pearson	-
Monotone	-	Spearman, Kendall

### Association Catégorielle

Situation	Test
2 variables catégorielles Force d'association	Chi-carré, Fisher exact V de Cramér

Score: \_\_/25