

# Statistique Inférentielle - Examen sur machine du 11/05/2015

Durée : 2h - Tous documents autorisés

## Consignes

- l'ensemble de vos réponses devront être intégrées dans un fichier word, dans lequel vous inclurez notamment vos analyses, des graphiques illustratifs, les codes R utilisés...
- en fin de séance, vous exporterez ce fichier en pdf sous le nom : **prenom-nom.pdf**. Une clef USB circulera en fin de séance pour récupérer ces documents pdf, et uniquement ceux-ci, non compressés. C'est cet unique fichier que je corrigerez. Pensez également à en sauvegarder une copie par vos propres moyens.
- pensez à **enregistrer régulièrement** votre travail.
- sauf indication contraire, on utilisera un risque de première espèce de 5%

## Exercice 1

Un laboratoire spécialisé en médecine du sport réalise une étude sur les capacités physiologiques développées par les sportifs. L'étude se focalise sur deux sports : la natation et l'athlétisme. Pour réaliser cette étude, le laboratoire sélectionne 20 candidats, 10 dans chaque sport. De sorte à avoir des candidats de même niveau dans les deux sports, le laboratoire sélectionne les 10 premiers du championnat national : 400m nage libre pour la natation, et 800m pour l'athlétisme.

Le laboratoire organise un test d'effort pour chacun de ces sportifs, et mesure les informations physiologiques suivantes :

- VO2 max : la consommation maximale d'oxygène (en litres d'oxygène par minute),
- temps de récupération : le temps (en secondes) que met le sportif, à la fin du test, pour retrouver un rythme cardiaque au niveau de son rythme avant l'effort.

sport		natation									
classement		1	2	3	4	5	6	7	8	9	10
temps de récupération (s)		186	197	233	180	231	185	188	181	182	208
VO2 max (L/min)		87.3	82.7	89.51	81.8	74.0	76.6	81.5	83.7	69.3	79.5
sport		athlétisme									
classement		1	2	3	4	5	6	7	8	9	10
temps de récupération (s)		162	172	178	163	167	166	162	164	172	178
VO2 max (L/min)		79.0	80.5	84.1	76.1	82.4	76.4	77.4	73.7	84.9	79.6

Pour chacune des questions suivantes posées par le laboratoire, apporter une réponse à l'aide d'une **procédure statistique appropriée**, et illustrez vos propos à l'aide de **graphiques** :

1. Un des deux sports permet-il de développer mieux que l'autre les capacités de récupération (récupération plus rapide) ?
2. Un des deux sports développe-t-il plus que l'autre la VO2 max ?
3. Y-a-t'il un lien entre le classement lors du championnat national et la VO2 max, pour l'athlétisme puis pour la natation ? Même question pour le temps de récupération.

## Exercice 2

Le même laboratoire que dans l'exercice précédent a suivi 12 nageurs de haut niveau, tous âgés de 18 ans en 2010, sur une période de 4 années. Les tests d'efforts réalisés chaque année ont permis de répertorier les VO2 max suivantes.

Année	Nageur											
	1	2	3	4	5	6	7	8	9	10	11	12
2010	76.5	77.3	75.1	79.0	79.7	76.1	77.6	71.8	71.8	76.6	76.3	71.8
2011	80.8	74.2	79.6	77.2	74.0	80.4	72.7	80.7	81.3	73.5	80.1	83.4
2012	84.0	79.5	76.2	78.6	79.6	77.2	80.2	79.2	83.1	86.5	83.5	82.9
2013	81.9	80.4	83.3	83.0	77.3	82.8	78.1	83.0	82.2	77.4	82.0	80.4

La question que se pose ce laboratoire est si ces nageurs continuent à augmenter leur capacités physiologiques après l'âge de 18 ans, ou autrement dit, si la VO2 max de ces sportifs croît significativement au cours des 4 années d'étude ?

1. Représenter graphiquement les données de sorte que l'on puisse se faire une idée de la réponse à la question.
2. A quelle question vous permettrait de répondre un test de Friedman sur ces données ? Réaliser ce test.
3. Pourquoi le test de Page est-il approprié ici pour répondre à la question du laboratoire ?  
En utilisant les informations ci-dessous (et non celles de votre polycopié de Statistique Inférentielle qui comporte une erreur), implémentez dans une fonction R ce test (utilisant l'approximation par la loi du  $\chi^2$  puisqu'ici nous avons 12 nageurs observés sur 4 ans).

*La statistique du test de Page est*

$$L = \sum_{k=1}^K k \sum_{j=1}^n R_{jk}$$

où  $R_{jk}$  est le rang intra-individu  $j$  de l'observation  $X_{jk}$  parmi  $(X_{j1}, \dots, X_{jK})$ . Lorsque  $n \geq 12$  et  $K \geq 4$ , la statistique suivante suit sous  $H_0$  une loi du  $\chi^2$  à 1 degré de liberté :

$$\frac{a^2}{nK^2(K^2 - 1)(K + 1)} \quad \text{avec} \quad a = 12L - 3nK(K + 1)^2$$

L'hypothèse  $H_0$  sera rejetée au profit de  $H_1 : F_1 > \dots > F_K$  si  $a > 0$  et si la valeur de la statistique est supérieure au quantile d'ordre  $1 - \alpha$  de la loi  $\chi^2_1$ . Si la valeur de la statistique est supérieure au quantile d'ordre  $1 - \alpha$  de la loi  $\chi^2_1$  mais avec  $a < 0$ , l'hypothèse  $H_0$  sera alors rejetée au profit de  $H_1 : F_1 < \dots < F_K$ .

4. Répondez alors à la question du laboratoire à l'aide du test de Page.