

Durée de l'examen : **1h30**

Ref : *ING1-MAIN-STA-DES-EXAM-S4*

A l'intention de : Etudiants des ING1-MAIN

Date : 27/ 05/2014

*Les documents papiers et électroniques sont autorisés. Votre PC est **autorisé seulement en local.***

### Exercice 1.

Choisissez la bonne réponse et justifiez-la :

**Q1)** Un professeur exigeant annonce à ses étudiants qu'au prochain examen, il veut que chacun d'eux ait une note supérieure à la moyenne du groupe. Est-ce possible ?

- A) Oui                      B) Non                      c) peut-être

**Q2)** Un professeur décide d'augmenter la note de chacun de ses étudiants de cinq points. Parmi les mesures suivantes, laquelle ne changera pas de valeur alors que les autres vont augmenter de dix ?

- A) La moyenne                      B) la médiane                      C) le mode                      D) l'écart-type

**Q3)** La catégorie socioprofessionnelle est :

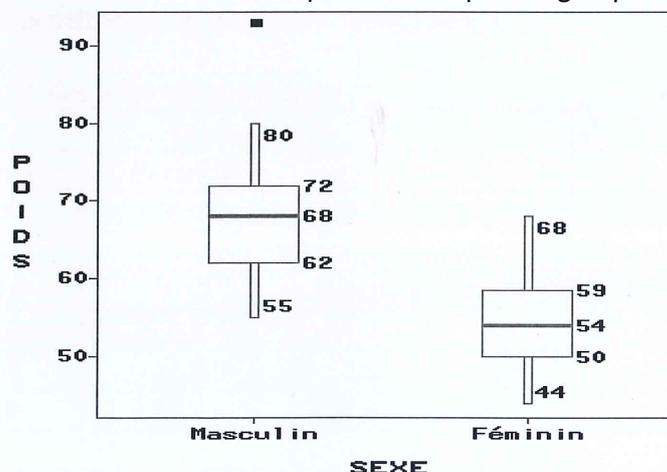
- A) Une variable qualitative nominale    B) Une variable qualitative ordinale

### Exercice 2.

J'ai sept enfants : l'âge modal (c'est à dire de plus grand effectif) est 5 ans, Paul a précisément l'âge médian : 7ans. Les jumeaux (Vladimir et Napoléon) ont l'âge moyen : 8 ans. Quel âge a mon aîné ? (considérez comme âge uniquement des valeurs entières).

### Exercice 3.

Dans une classe mixte d'effectif 59, nous avons collecté pour chaque élève : son poids (en kilogrammes) et son sexe. Pour comparer la distribution de la variable POIDS selon le sexe : Masculin/Féminin, nous avons juxtaposé sur le même graphique les 2 boîtes à moustaches définies respectivement pour le groupe Masculin et le groupe Féminin.



*Distributions de la variable POIDS selon le sexe.*

- Quelles sont les valeurs maximales et minimales (pour les deux groupes) ?
- Commenter, en donnant le maximum d'information, la distribution représentant le poids masculin.
- Comparer les deux distributions selon le sexe.

# ING1-MAIN : STATISTIQUE DESCRIPTIVE : EXAMEN S4

## Exercice 4.

Le tableau ci-dessous, résume la distribution des 2589 milliers de tonnes d'équivalent-pétrole utilisés par les principaux pays occidentaux en 1974 suivant les pays et les sources d'énergie:

**Tableau 1 des effectifs observés**

Pays (X)	Sources d'énergie (Y)				Total
	charbon	électricité	gaz	Pétrole	
Allemagne Fed	82	3	31	135	251
Canada	18	16	40	74	148
Etats-Unis	345	31	570	754	1700
France	30	5	15	110	160
Italie	8	4	18	88	118
Royaume Uni	88	3	27	94	212
Total	571	62	701	1255	2589

**Tableau 2 des ... ?**

Pays (X)	Sources (Y)				Total
	charbon	électricité	gaz	Pétrole	
Allemagne Fed	0,327	0,012	0,124	??	1
Canada	0,122	0,108	0,270	0,500	1
Etats-Unis	0,203	0,018	0,335	0,444	1
France	??	0,031	0,094	0,688	1
Italie	0,068	0,034	0,153	0,746	1
Royaume Uni	0,415	0,014	0,127	0,443	1
Total	0,221	0,024	0,271	0,485	1

- 1) Que représente le tableau 2 ? Compléter les cases vides et expliquer ce que signifient ces cases. Que représente la dernière ligne ?
- 2) Laquelle des deux interprétations suivantes de ce tableau vous paraît le mieux traduire la situation : les variables X et Y sont "plutôt liées" ou les variables X et Y sont "plutôt Indépendantes". Justifier votre réponse.
- 3) Décrire brièvement la méthode du test du Chi deux ( $\chi^2$ ) qui peut confirmer ou infirmer votre jugement.

Pour la mise en œuvre de cette méthode, on donne les deux tableaux :

**Tableau 3 des effectifs théoriques**

Pays (X)	Sources (Y)				Total
	charbon	électricité	gaz	Pétrole	
Allemagne Fed	55,4	6,0	68,0	121,7	251,0
Canada	32,6	3,5	40,1	71,7	148,0
Etats-Unis	374,9	40,7	460,3	824,1	1700,0
France	35,3	3,8	43,3	77,6	160,0
Italie	26,0	2,8	31,9	57,2	118,0
Royaume Unis	46,8	5,1	57,4	102,8	212,0
Total	571,0	62,0	701,0	1255,0	2589,0

# ING1-MAIN : STATISTIQUE DESCRIPTIVE : EXAMEN S4

**Tableau 4 des écarts entre les effectifs observés et les effectifs théoriques**

Pays (X)	Sources (Y)				Total
	charbon	électricité	gaz	Pétrole	
Allemagne Fed	-26,6	3,0	37,0	-13,3	0,0
Canada	14,6	-12,5	0,1	-2,3	0,0
Etats-Unis	29,9	9,7	-109,7	70,1	0,0
France	5,3	-1,2	28,3	-32,4	0,0
Italie	18,0	-1,2	13,9	-30,8	0,0
Royaume Unis	-41,2	2,1	30,4	8,8	0,0
Total	0,0	0,0	0,0	0,0	0,0

- 4) Qu'indiquent les écarts positifs et les écarts négatifs ? Comment se situe le caractère Etats-Unis ? Dans quel cas le tableau des écarts contient-il uniquement des 0 ?
- 5) Expliquer comment on calcule la distance du Chi-deux. Peut-on utiliser ce critère pour mesurer l'indépendance des variables ?

## Exercice 5

On envisage de prévoir la taille des œufs de coucou susceptibles d'être pondus dans un nid, à partir des dimensions de ce nid. La variable Y désigne la variable "longueur d'un œuf" et la variable X le "diamètre d'un nid" en mm. Pour chaque œuf d'un échantillon de taille  $n = 16$  œufs de coucou, on a relevé d'une part la réalisation de Y et d'autre part celle de X pour le nid où il a été trouvé. Les valeurs observées  $(x_i, y_i)$  du couple (X, Y) sont les suivantes :

$x_i$	100	113	110	106	112	105	107	108	122	126	121	122	110	116	118	120
$y_i$	19,8	22,1	21,5	20,9	22	20,8	21,2	21	23,8	24,9	24	23,8	21,7	22,8	23,1	23,5

diamètre nid  
longueur œuf

- 1) Préciser le modèle de régression linéaire proposé et les hypothèses nécessaires pour l'appliquer.
- 2) Calculer les estimations des paramètres a et b du modèle correspondant aux données (voir, indication numérique ci-dessous ou avec Excel), et donner l'équation de la régression.
- 3) Quel est le pourcentage de variabilité de la taille des œufs qui est expliqué par la relation obtenue ? Qu'en pensez-vous ?
- 4) On s'intéresse aux nids de 128 mm de diamètre et on cherche à prévoir la taille des œufs de coucou qu'on peut s'attendre à y trouver. Quelle est la valeur prédite de la longueur de tels œufs.

### Indications numériques :

$$\sum x_i = 1816, \quad \sum x_i^2 = 206952, \quad \sum y_i = 356.9, \quad \sum y_i^2 = 7992.07, \quad \sum x_i y_i = 40668.3.$$

$$\sum (x_i - \bar{x})^2 = 836, \quad \sum (y_i - \bar{y})^2 = 30,969375$$