



Licence3 en Génie Electrique
Examen de Traitement Statistique de Données d'Energie
Durée : 2h00 mn

NB : Tout document et ordinateur autorisés sauf la copie du voisin

Enseignant : Dr Ir. Codjo Emile AGBANGBA

Exercice 1.

a. Répondez par vrai ou faux

1. Un test statistique permet entre autres de comparer une population à une population de référence.
2. Le but d'un test statistique est de déterminer si la différence observée est due au hasard ou à une vraie différence entre les échantillons dont sont issus les populations.
3. Une hypothèse alternative contient forcément un signe différent (\neq)
4. H_0 et H_1 concernent toujours la (les) valeurs vraie(s) dans la population.
5. Plus la P-value est faible, moins H_0 est crédible.

1 point * 5 = 5 points

1. Vraie
2. Vraie
3. Faux
4. Faux
5. Vraie

b. La largeur (en cm) d'une fente dans une pièce fabriquée en aluminium est distribuée selon une loi normale de paramètre $\mu = 2$ et $\alpha = 0,012$. Les limites de tolérance sont données comme étant $2,000 \pm 0,012$.

Quel sera le pourcentage de pièces défectueuses ?

$$X \rightarrow N(2,0,012) ; \text{ soit } U \rightarrow (0,1) ; U = \frac{X - \mu}{\alpha}$$

$$P = 1 - P(2,000 - 0,012 < X < 2,000 + 0,012) = 1 - P(1,988 < X < 2,012)$$

$$P(1,988 < X < 2,012) = P\left(\frac{1,988 - 2}{0,012} < U < \frac{2,012 - 2}{0,012}\right) = P(-1 < U < 1)$$

$$P(-1 < U < 1) = 0.8185946$$

$$P = 1 - 0.8185946 = 0.1814054$$

$$P = 0,18 \quad \mathbf{3 \text{ points}}$$

Exercice 2.

On veut étudier la consommation en énergie des ménages d'un village. On choisit de façon aléatoire simple et indépendante $n = 30$ ménages. Les données collectées sont consignées dans le tableau ci-dessous :

Ménages	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Energie consommée (Kwh)	40	100	25	80	10	2	40	20	9	1	50	60	80	47	6
Ménages	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Energie consommée (Kwh)	23	79	100	16	64	40	25	21	45	18	20	46	32	30	38

- Calculer la consommation moyenne en énergie dans le village
 $M = 39 \text{ Kwh}$ 1pt
- En supposant que la variable suit une distribution normale, dans quel intervalle varie la consommation à un seuil de signification de $\alpha = 0,05$?

$IC 95\% = [28,57 - 49,23]$ 2 pts

- Peut-on affirmer à un seuil $\alpha = 0,05$ que la consommation moyenne en énergie est de 60 Kwh ?

Ici l'étudiant peut utiliser trois méthodes :

Hypothèse :

$H_0 = m=60$ contre $H_1 : m \neq 60$ (1pts)

- L'utilisation de l'intervalle de confiance :
 $M_0 = 60$ n'appartient pas à l'intervalle de confiance, donc H_0 est rejetée (2pts)
 et on conclut que la consommation énergétique moyenne est différente de 60 Kwh (2pts).
 Oubien :
- Utilisation de la statistique du test
 $T_{obs} = 5.7236 > t_{1-\alpha/2}$ à $n-1$ ddl donc on rejette H_0 (2pts)
 et on conclut que la consommation énergétique moyenne est différente de 60 Kwh (2pts)
 Oubien :
- Calcul du P-value
 $P\text{-value} = P(R_{Ho}/H_0) = 0.0002457 < 0.05$, on Rho (2pts)
 et on conclut que la consommation énergétique moyenne est différente de 60 Kwh (2pts)

- Le type d'énergie consommé est varié : énergie renouvelable et énergie non renouvelable. On veut en outre savoir si les types d'énergie consommée dépendent du niveau de scolarité des chefs ménages. On réalise de ce fait, une enquête impliquant 100 ménages choisis de façon aléatoire simple et indépendante.

Type d'énergie	Illettré	Primaire	Secondaire	Supérieur
Energie renouvelable	2	5	40	68
Energie non renouvelable	50	20	10	4

- a. Peut-on affirmer au seuil de 0,05 que le type d'énergie utilisée dépend du niveau de scolarité du chef ménage ?
- b. Peux-ton affirmer au seuil de 0,05 que la proportion de ménages utilisant les énergies renouvelables est d'au moins de 30% ?

- a. H_0 : Indépendance ($P_1=P_2=P_3=...=P_8$) contre H_1 : dépendance (au moins l'une au moins des proportions différent)

Ici la méthode à utiliser ici est le texte exact de fisher car moins de 80 % des n_{ij} sont sup à 5.

```
> fisher.test(data)
```

Fisher's Exact Test for Count Data

```
data: data
p-value < 2.2e-16
alternative hypothesis: two.sided
```

p-value < 0.05 donc l'hypothèse H_0 est rejetée et on conclut que le type d'énergie utilisé dépend du niveau de scolarisation du chef ménage. ? (1.5pt)

On peut aussi accepter les gens qui ont utilisé simplement le test d'indépendance de χ^2 ? (1.5pts)

Si quelqu'un n'utilise pas le test exact de fisher, et utilise simplement le test d'indépendance de χ^2 et tire bien la conclusion, tu peux lui accord **1,5 points au total** (hypothèse, conclusion)

Deux méthodes

-Soit calcul de $\chi^2_{obs} = 126.44 > \chi^2_{1-\alpha}$ à ddl = $(2-1)(4-1) = 3$; donc on RHO

- soit calcul du P-value

P-value < 2.2e-16 donc < 0.05 donc on RHO.

b.

c'est un test de conformité d'une proportion :

H_0 : $p > 0.30$ contre H_1 : $p < 0.30$ (0.5 pt)

p-value = 1 > 0.05 alors on accepte H_0 et on conclut que la proportion de personnes utilisant les énergies renouvelables est d'au moins de 30% au seuil de 0.05. (1.5 pts)

```
> binom.test(115,199,0.30, alternative ="less")
```

Exact binomial test

```
data: 115 and 199
```

```
number of successes = 115, number of trials = 199, p-value = 1
```

```
alternative hypothesis: true probability of success is less than 0.3
```

95 percent confidence interval:

0.000000 0.636823

sample estimates:

probability of success

0.5778894