



TD N° 7 : Comparaison d'échantillons

Exercice 1

Une entreprise fabrique un médicament sur deux chaînes de production. On s'intéresse aux variations de la quantité d'une certaine substance A contenue dans chaque médicament. On a contrôlé le dosage de la substance A avec un échantillon de 100 médicaments à la sortie de chacune des deux chaînes de fabrication. On a trouvé un dosage moyen de 10.75mg pour la première chaîne et 10.70mg pour la deuxième. On sait par ailleurs que l'écart-type des chaînes de production est le même et est égal à 0.2mg.

Construire un test à 1% permettant de savoir si la différence des moyennes observées est due à des fluctuations de l'échantillonnage ou bien si la chaîne de fabrication n°1 produit des médicaments contenant davantage de substance A que la chaîne n°2.

Exercice 2

Un sondage effectué auprès de 2000 personnes indique que 19% d'entre elles connaissent la marque de lessive Omopaic. Après une campagne publicitaire, un sondage analogue auprès de 1000 personnes montre que 230 d'entre elles connaissent cette marque. Peut-on considérer que la campagne a été efficace ?

Exercice 3

Une enquête sur la consommation annuelle des ménages est réalisée par l'INSEE régulièrement. Ces ménages sont répartis en 5 grandes catégories suivant leur localisation :

- C1 : ménages en zone rurale,
- C2 : ménages résidant dans une unité urbaine inférieure à 20000 habitants,
- C3 : ménages résidant dans une unité urbaine comprise entre 20000 habitants et 100000 habitants,
- C4 : ménages résidant dans une unité urbaine supérieure à 100000 habitants autre que l'agglomération parisienne,
- C5 : ménages résidant dans l'agglomération parisienne.

Un groupement commercial s'intéresse particulièrement à la consommation annuelle des produits contenus dans la nomenclature 17 de l'INSEE c'est-à-dire, la consommation annuelle en mouton, agneau et chevreau et il souhaite savoir s'il y a un effet "localisation" sur la consommation annuelle moyenne des ménages pour ces produits. Le groupement commercial interroge 5 ménages par catégories. Les résultats en euro sont :

C1	C2	C3	C4	C5
56	47	55	61	69
66	50	51	62	71
54	55	59	54	55
61	46	54	54	62
56	56	59	62	53

On suppose que, pour tout $i \in \{1, \dots, 5\}$, la consommation annuelle d'un ménage en euro de catégorie C_i peut être modélisée par une var X_i suivant la loi normale $N(\mu_i, \sigma^2)$, avec μ_i et σ inconnus.

1. On donne $sceT = 908,64$ et $sceR = 556,40$. Dresser le tableau ANOVA.

2. Effectuer, au risque 5%, le test ANOVA :

$$H_0: \mu_1 = \dots = \mu_5$$

contre

H_1 : "il existe au moins 2 moyennes différentes"

Interpréter le résultat.

3. Peut-on affirmer, au risque 5%, que $\mu_3 \neq \mu_4$?