

Séance de travaux dirigés 3

Démarche statistique

Exercice 1 : Test d'égalité de moyennes

Sur une feuille blanche, tracez, **sans l'aide d'une règle**, deux points distants de 12 cm. Mesurez ensuite la longueur entre ces deux points et reportez la longueur sur une feuille.

Une fois toutes les données recueillies, considérons deux groupes : les longueurs évaluées par les garçons et celles évaluées par les filles. On suppose que la longueur d'un segment suit une loi normale d'espérance μ_1 et de variance σ^2 lorsqu'elle est évaluée par un garçon et une loi normale d'espérance μ_2 et de variance σ^2 lorsqu'elle est évaluée par une fille.

On cherche dans un premier temps à savoir si les données recueillies confirment ou non une évaluation juste de la longueur demandée du segment (12 cm) par une fille d'une part, et par un garçon d'autre part.

1. *Exprimez la problématique ci-dessus sous la forme de deux tests d'hypothèses similaires, un pour les filles, l'autre pour les garçons.*
2. *Dans ces problématiques de test, quels sont les échantillons et les populations ?*
3. *A l'aide des données recueillies, mettez en œuvre les tests de la question 1 avec un seuil de 5% pour l'erreur de type I.*
4. *Donnez, aux niveaux de confiance 95% les intervalles de confiance de μ_1 et μ_2 .*
5. *Tester l'égalité entre μ_1 et μ_2 au seuil de 10% pour l'erreur de type I.*

Exercice 2 : Comparaison de deux chocolats noirs

On compare la saveur en cacao de 2 chocolats noirs appelés A et B. Pour cela, $n_A = 24$ personnes ont noté la saveur en cacao du chocolat A et $n_B = 24$ autres personnes la saveur en cacao du chocolat B. La notation s'effectue à l'aide d'une échelle de 0 (pas de saveur) à 10 (très forte saveur).

Le chocolat A a une teneur en cacao annoncée de 78 % tandis que le chocolat B a une teneur annoncée de 66 %. On se demande si le chocolat A est perçu comme ayant une saveur de cacao plus intense que celle du chocolat B.

Dans la suite, μ_A et μ_B désignent les notes de saveurs théoriques des chocolats A et B respectivement.

1. *Montrez que ce problème relève d'un test d'hypothèses dont vous préciserez l'hypothèse nulle H_0 et choisirez H_1 , en justifiant votre choix, parmi les possibilités suivantes :*
 - a) $H_1 : \mu_A \neq \mu_B$,
 - b) $H_1 : \mu_A < \mu_B$,
 - c) $H_1 : \mu_A > \mu_B$.

Dans la suite, on note \bar{x}_A (respectivement \bar{x}_B) la note de saveur moyenne pour le chocolat A (respectivement pour le chocolat B) et s'_A (respectivement s'_B) l'écart-type des notes de saveur pour le chocolat A (respectivement pour le chocolat B).

2. *Donnez, en fonction de \bar{x}_A , \bar{x}_B , s'_A , s'_B , n_A et n_B , l'expression de la statistique de test qu'il faut calculer pour décider de rejeter ou non l'hypothèse H_0 ?*

Dans la suite, la statistique de test de la question précédente est notée T (comme Test). Pour finaliser le test, il reste à définir la région de rejet de H_0 , c'est-à-dire l'ensemble des valeurs de T pour lesquelles on rejette H_0 , pour

un risque de 1ère espèce $\alpha = 0.05$. On définit en général cette région de rejet à partir d'un seuil k , dont la valeur dépend de α .

3. Parmi les formes suivantes de la région de rejet, donnez celle qui est la plus adaptée à la situation posée dans la question 1 (justifiez) :

- a) $|T| > k$,
- b) $|T| < k$,
- c) $T > k$,
- d) $T < k$,

4. Donnez la loi de T sous l'hypothèse H_0 .

Dans la suite, on appelle cette loi \mathcal{L} .

5. Parmi les choix suivants, donnez la valeur de k , pour garantir un risque de 1ère espèce inférieur à 0.05 :

- a) k est le quantile de la loi \mathcal{L} d'ordre 0.975,
- b) k est le quantile de la loi \mathcal{L} d'ordre 0.95,
- c) k est le quantile de la loi \mathcal{L} d'ordre 0.05,
- d) k est le quantile de la loi \mathcal{L} d'ordre 0.025.

La fonction `t.test` du logiciel R permet de mettre en œuvre le test d'hypothèse ci-dessus. Les résultats sont donnés dans le tableau 1.

6. Finalement, au seuil $\alpha = 0.05$, quelle décision prenez-vous ? (justifier votre choix). Si le seuil était $\alpha = 0.001$, quelle serait votre décision ?

```
t.test(saveurA,saveurB,var.equal=TRUE,alternative="greater")

Two Sample t-test

data: saveurA and saveurB
t = 3.4469, df = 46, p-value = 6.1120e-04
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.5152      Inf
sample estimates:
mean of x mean of y
 7.3375     6.3333
```

Table 1: Saveur en cacao de chocolats noirs : test de comparaison de moyennes (fonction `t.test` de R)