

TD N° 4 : Estimation et tests (suite)

EXERCICE 1. (Expérience de Mendel) Reprenons l'étude de Mendel sur 557 petits pois.

La couleur est codée par un gène présentant deux allèles, C et c , correspondant aux couleurs jaune et vert. Le jaune est dominant, le vert est récessif. La forme est portée par un autre gène à deux allèles : R pour rond, dominant et r pour ridé, récessif.

Il y a donc 16 génotypes probables (les petits pois sont diploïdes), mais on n'observe que 4 phénotypes : jaune rond, jaune ridé, vert rond et vert ridé.

- 1) Lister les 16 génotypes et trouver, pour chacun d'eux, les phénotypes correspondants.
- 2) On suppose que les 16 génotypes sont équiprobables. Calculer les probabilités de chacun des phénotypes. En déduire les effectifs théoriques sur un échantillon de 557 petits pois.
- 3) Les effectifs observés par Mendel, donnés dans le tableau ci-dessous, diffèrent-ils des effectifs théoriques ? On donne pour aide numérique que $\mathbb{P}(\chi_3^2 > 0.45) \approx 0.93$. Discuter alors de votre choix.

| Phénotype | jaune rond | jaune ridé | vert rond | vert ridé |
|-----------|------------|------------|-----------|-----------|
| Effectif | 315 | 102 | 108 | 32 |

```
# Code Python
import numpy as np
from scipy.stats import chisquare, chi2
obs = np.array([315 , 102 , 108 , 32])
th = np.array([313 , 104.5 , 104.5 , 35])
chi2_stat, pval = chisquare(obs, th)
1 - chi2.cdf(chi2_stat, df=3)
# pval = 0.9303804770054722
# chi2_stat = 0.44695590268147856
```

EXERCICE 2.

Sur 100 individus issus de croisements de deux hétérozygotes Aa , on a observé 36 phénotypes $[a]$, récessifs. Le but de l'étude est de tester si AA est mortel.

- a) Soit S le nombre de sujets de phénotype $[a]$, on suppose que les 4 génotypes sont équiprobables. Donner la loi de S .
- b) Les observations suivent-elles une loi de Mendel classique ?
- c) Écrire les hypothèses nulles et alternatives pour tester si AA est mortel.
- d) Conclure sur les données observées.

On donne pour aide numérique que $\mathbb{P}(\chi_1^2 > 0.321) \approx 0.5716$ et $\mathbb{P}(\chi_1^2 > 6.453) \approx 0.0111$

EXERCICE 3. On dispose d'un jeu de données composé de deux échantillons supposés gaussiens indépendants, dont le résumé est

| Échantillon 1 | Échantillon 2 |
|--------------------|--------------------|
| $n_1 = 55$ | $n_2 = 60$ |
| $\bar{x}_1 = 18.4$ | $\bar{x}_2 = 16.5$ |
| $s_1 = 8.6$ | $s_2 = 13.7$ |

Supposons que σ_1 et σ_2 sont connus tels que $\sigma_1 = 8.7$ et $\sigma_2 = 13.8$

- 1) Construire un intervalle de confiance pour la différence des moyennes au niveau 98%.
- 2) Tester $\mathcal{H}_0 : \mu_1 - \mu_2 = 2.5$ contre $\mathcal{H}_1 : \mu_1 - \mu_2 \neq 2.5$ avec $\alpha = .02$.
- 3) Tester $\mathcal{H}_0 : \mu_1 - \mu_2 = 2.5$ contre $\mathcal{H}_1 : \mu_1 - \mu_2 < 2.5$ avec $\alpha = .02$.

4) Refaire l'exercice en supposant σ_1 et σ_2 inconnus.

On donne $q_{0.99} = 2.3263$, $q_{0.98} = 2.0537$, $t_{0.99}(54) = 2.3974$, $t_{0.98}(54) = 2.1045$.

EXERCICE 4. Une expérience de comparaison de deux traitements de plantes repose sur l'étude d'un échantillon de 88 plantes. Dans ce groupe, 40 plantes ont été sélectionnées au hasard pour subir le traitement 1 et les 48 plantes restantes ont subi le traitement 2. La moyenne et l'écart-type du poids des plantes dans les deux échantillons sont

| | Traitement 1 | Traitement 2 |
|------------|--------------|--------------|
| Moyenne | 16.21 | 27.84 |
| Écart-type | 2.88 | 4.32 |

- 1) Donner un intervalle de confiance, au niveau 95%, de la différence des moyennes.
- 2) Le traitement 1 est un placebo, le traitement 2 est un nouvel insecticide. Cet insecticide ne sera mis sur le marché que si le poids des plantes ayant subi ce traitement est d'au moins 10 unités plus élevé qu'en l'absence de traitement. Écrire les hypothèses nulles et alternatives.
- 3) Quelle est la statistique de test ? Quelle est la région de rejet pour $\alpha = 0.05$?
- 4) Effectuer le test. Trouver la p -valeur et commenter.

On donne $q_{0.95} \approx 1.645$, $q_{0.975} \approx 1.96$, $t_{0.975}(39) \simeq 2.023$, $t_{0.95}(39) \approx 1.685$.

EXERCICE 5. Un chercheur en pharmacie veut déterminer si le médicament qu'il a développé a comme effet secondaire de faire baisser la pression artérielle. L'étude commence par l'enregistrement de la pression chez 15 jeunes étudiantes avant le traitement. Ensuite, ces 15 femmes ont pris ce médicament quotidiennement pendant 6 mois et leur pression artérielle a été mesurée à la fin de cette période.

Les données sont

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|-------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Avant | 70 | 80 | 72 | 76 | 76 | 76 | 72 | 78 | 82 | 64 | 74 | 92 | 74 | 68 | 84 |
| Après | 68 | 72 | 62 | 70 | 58 | 66 | 68 | 52 | 64 | 72 | 74 | 60 | 74 | 72 | 74 |

- 1) S'agit-il d'un échantillon apparié ou indépendant ?
- 2) Quelle hypothèse nulle et alternative doit-on poser ?
- 3) Calculer un intervalle de confiance au niveau 95% de la différence de pression artérielle.
- 4) Les données montrent-elles une baisse significative de la pression ?