TD3 : intervalles de confiance 2 (données pourcentage ou données gaussiennes)

Rappel :

– Loi des données : loi de Bernoulli (un “pourcentage”) de paramètre p.
Calcul du pourcentage estimé : $\hat{p} = \frac{\text{nombre de réponses "ok"}}{\text{nombre total de réponses}}$.
Calcul de l'écart-type estimé : $\hat{\sigma} = \sqrt{\hat{p}(1-\hat{p})}$.

– Loi des données : normale/gausienne de moyenne m (inconnue) et d'écart-type σ (inconnu).
Par la loi de Student : $P(-z \leq \frac{m-\hat{m}}{\hat{\sigma}/\sqrt{n}} \leq z) \simeq P(-z \leq \mathcal{T}(n-1) \leq z)$.

Exercice 1.
1. Trouver z tel que $P(-z \leq \mathcal{N}(0,1) \leq z) = 0.95$
2. Trouver z tel que $P(-z \leq \mathcal{T}(4) \leq z) = 0.95$.

Exercice 2. Trouver la valeur de z telle que

1. $P(-z \leq \mathcal{T}(8) \leq z) = 0.8$, $z = 1.397$.
2. $P(-z \leq \mathcal{T}(13) \leq z) = 0.95$, $z = 2.160$.
3. $P(-z \leq \mathcal{T}(1) \leq z) = 0.8$, $z = 3.078$.
4. $P(-z \leq \mathcal{T}(25) \leq z) = 0.99$, $z = 2.787$.
5. $P(-z \leq \mathcal{T}(76) \leq z) = 0.5$, $z = 0.674$.
6. $P(-z \leq \mathcal{N}(0,1) \leq z) = 0.99$, $z = 2.576$.

1 Exercices “pourcentage”

Exercice 3. Les résultats d'une enquête sur 200 personnes donnent que 164 personnes interrogées boivent de l'alcool dans un cadre festif.

1. Trouver un intervalle de confiance, au niveau 90%, sur le pourcentage de gens qui boivent de l'alcool dans un cadre festif.

$$\hat{p} = 0.82, \hat{\sigma} = \sqrt{0.82 \times 0.18} = 0.385.$$

Comme $200 \geq 30$, $164 \geq 5$ et $200 - 164 = 46 \geq 5$, on peut utiliser le TCL.

$$\left[0.82 - 1.645 \times \frac{0.385}{\sqrt{200}}; 0.82 + 1.645 \times \frac{0.385}{\sqrt{200}}\right] = [0.775; 0.865]$$

Exercice 4. Au cours d'une enquête d'opinion publique, on pose la question suivante : “êtes-vous satisfait(e) de votre salaire ?”. Sur 1500 personnes interrogées, 57% répondent être satisfaites.

1. Trouver un intervalle de confiance, au niveau 95%, sur le pourcentage de gens satisfaits par leur salaire.

$$\hat{p} = 0.57(\text{énoncé}), \hat{\sigma} = \sqrt{0.57 \times 0.43} = 0.495.$$

Comme $1500 \geq 30$, $1500 \times 0.57 = 855 \geq 5$ et $1500 \times 0.43 = 645 \geq 5$, TCL.

$$\left[0.57 - 1.96 \times \frac{0.495}{\sqrt{1500}}; 0.57 + 1.96 \times \frac{0.495}{\sqrt{1500}}\right] = [0.545; 0.595]$$

2 Exercices “loi des données gaussiennes”

Exercice 5. Pour mesurer la qualité de l’air sur 1 mois autour d’une usine, le service de contrôle de la qualité de l’air a prélevé 20 échantillons, supposés gaussiens, de la concentration en Dioxyde d’Azote (NO_2). L’estimation ponctuelle de ces échantillons a donné une moyenne de 93 μg/m^3 (μ = “micro”) de Dioxyde d’Azote avec un écart-type de 11 μg/m^3.

1. Trouver un intervalle de confiance, au niveau 95%, de la concentration en Dioxyde d’Azote autour de l’usine.

Comme $n = 20 \leq 30$, mais données gausiennes, donc loi de Student de degré de liberté 19.

$$\left[93 - 2.093 \times \frac{11}{\sqrt{20}}; 83 + 2.093 \times \frac{1}{\sqrt{20}}\right] = [87.8; 98.2]$$

Exercice 6 (Sujet de rattrapage 2021-2022)**.** Un artiste fabrique et vend des fèves pour galette. Il souhaite estimer sa vitesse de production moyenne. Pour cela, pendant 1 journée de 7h de travail, il note le nombre de fèves qu’il a fabriqués dans l’heure. Il obtient la liste suivante 4, 5, 8, 8, 5, 2 et 3.

1. Donner un intervalle de confiance au niveau 90% du nombre de fèves fabriquées en une heure. **n < 30 on ne peut pas faire (données de loi inconnue).**
2. En supposant que les données suivent une loi normale, donner un intervalle de confiance au niveau 90% du nombre de fèves fabriquées en une heure.

En supposant les données gaussiennes (ce qui est discutable),

$$\hat{m} = \frac{4+5+8+8+5+2+3}{7} = 5$$
$$\hat{\sigma} = \sqrt{\frac{(-1)^2 + 0^2 + 3^2 + 3^2 + 0^2 + (-3)^2 + (-2)^2}{6}} = \sqrt{\frac{32}{6}} = 2.31$$

Par la loi de Student à 6 degrés de liberté :

$$\left[5 - 1.943 \times \frac{2.31}{\sqrt{7}}; 5 + 1.943 \times \frac{2.31}{\sqrt{7}}\right] = [3.3; 6.7]$$

3 Exercices d’entrainement

Exercice 7 (Sujet de rattrapage 2021-2022)**.** Cette année, une grande surface tente une expérience pour booster la vente de cidre en période d’épiphanie : elle dispose à la vente, à côté des galettes des rois, des bouteilles de cidre. L’analyse des ventes montre que, sur les 2346 clients qui ont acheté une galette des rois, 485 ont aussi acheté une bouteille de cidre.

Donner un intervalle de confiance au niveau 92% du pourcentage de gens qui achète une bouteille de cidre avec leur galette dans cette configuration.

$$\hat{p} = \frac{485}{2346} = 0.2067, \hat{\sigma} = \sqrt{0.2067 \times 0.7933} = 0.4049.$$

Comme $2346 \geq 30$, $485 \geq 5$ et $2346 - 485 = 1861 \geq 5$, on peut utiliser le TCL.

$$\left[0.2067 - 1.75 \times \frac{0.4049}{\sqrt{2346}}; 0.2067 + 1.75 \times \frac{0.4049}{\sqrt{2346}}\right] = [0.1921; 0.2213]$$

Exercice 8 (Sujet 2021-2022)**.** Une boutique en ligne teste un nouveau design pour son site web afin d'améliorer son taux de conversion (pourcentage de visiteurs qui réalisent un achat). Sur le premier mois de test, 2965 sur les 95634 visiteurs du site ont effectué un achat.

Donner un intervalle de confiance au niveau 98% pour le taux de conversion avec le nouveau site web.

$$\hat{p} = \frac{2965}{95634} = 0.031, \hat{\sigma} = \sqrt{0.031 \times 0.969} = 0.173.$$

Comme $95634 \geq 30$, $2965 \geq 5$ et $95634 - 2965 = 92669 \geq 5$, on peut utiliser le TCL.

$$\left[0.031 - 2.33 \times \frac{0.173}{\sqrt{95634}}; 0.031 + 2.33 \times \frac{0.173}{\sqrt{95634}}\right] = [0.030; 0.032]$$

Exercice 9. Afin d'avoir une idée du marché, un agent immobilier a relevé le prix du m^2 d'appartements en vente à Sceaux.

9 610€, 9 774€, 9 800€, 10 067€, 9 486€

Sa longue expérience lui permet de supposer que les données sont gaussiennes.

1. Donner un intervalle de confiance au niveau 90% du prix au m^2 d'un appartement en vente à Sceaux.

$$\hat{m} = \frac{9610 + 9774 + 9800 + 10067 + 9486}{5} = 9747.4$$
$$\hat{\sigma} = \sqrt{\frac{(-137.4)^2 + 26.6^2 + 52.6^2 + 319.6^2 + (-261.4)^2}{4}} = 219.5$$

Par la loi de Student à 4 degrés de liberté :

$$\left[9747.4 - 2.132 \times \frac{219.5}{\sqrt{5}}; 9747.4 + 2.132 \times \frac{219.5}{\sqrt{5}}\right] = [9538.1; 9956.7]$$

2. Donner un intervalle de confiance au niveau 90% du prix d'un appartement de $80m^2$ en vente à Sceaux.

$$\hat{m}' = \frac{9610 + 9774 + 9800 + 10067 + 9486}{5} \times 80 = \hat{m} \times 80 = 779792$$
$$\hat{\sigma}' = \sqrt{\frac{((-137.4)^2 + 26.6^2 + 52.6^2 + 319.6^2 + (-261.4)^2) \times 80^2}{4}} = \hat{\sigma} \times 80 = 17564.8$$

Les valeurs de n et z ne changent pas.

$$\left[779792 - 2.132 \times \frac{17564.8}{\sqrt{5}}; 779792 + 2.132 \times \frac{17564.8}{\sqrt{5}}\right] = [763044.7; 796539.3]$$

Exercice 10. Un concessionnaire regarde le nombre de ventes de véhicules par jour dans sa concession sur 20 jours. Il compile ses données dans le tableau suivant

Véhicules vendus dans la journée	7	8	9	10	11	12
Nombre de jours	2	2	6	7	2	1

On suppose que le nombre de véhicules vendus en une journée suit une loi normale.

Donner un intervalle de confiance au niveau 95% du nombre moyen de véhicules vendus quotidiennement dans la concession.

$$\hat{m} = \frac{2 \times 7 + 2 \times 8 + 6 \times 9 + 7 \times 10 + 2 \times 11 + 1 \times 12}{20} = 9.4$$

$$\hat{\sigma} = \sqrt{\frac{(-2.4)^2 \times 2 + (-1.4)^2 \times 2 + (-0.4)^2 \times 6 + 0.6^2 \times 7 + 1.6^2 \times 2 + 2.6^2 \times 1}{19}} = 1.27$$

Par la loi de Student à 19 degrés de liberté :

$$\left[9.4 - 2.093 \times \frac{1.27}{\sqrt{20}}; 9.4 + 2.093 \times \frac{1.27}{\sqrt{20}}\right] = [8.81; 9.99]$$