

EXERCICES SUR LE CHAPITRE 3 – 1^{ÈRE} PARTIE

L'estimation d'une moyenne et la comparaison de deux moyennes

(Échantillons indépendants et données pairées)

FORMULES

- $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$
- $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \left(\frac{1}{n} \sum_{i=1}^n y_i^2 \right) - \bar{y}^2$
- $s^2 = \frac{n}{n-1} \hat{\sigma}^2$

• Estimation d'une moyenne isolée :

→ $(\mu | \text{données}) \sim t_{n-1} \left(\bar{y}, \frac{s^2}{n} \right)$

IC (95%) pour μ : $\bar{y} \pm t_{n-1}(0.975) \frac{s}{\sqrt{n}}$

→ Si $n \rightarrow \infty$: $(\mu | \text{données}) \sim N \left(\bar{y}, \frac{s^2}{n} \right)$

• Comparaison de deux moyennes (échantillons indépendants) : $\theta = \mu_1 - \mu_2$

→ 1^{er} cas : si n_1 et n_2 sont $> 20 \rightarrow$ Approximation normale

$(\theta = \mu_1 - \mu_2 | \text{données}) \approx N \left(\bar{y}_1 - \bar{y}_2, \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)$

IC (95%) pour $\theta = (\bar{y}_1 - \bar{y}_2) \pm 1.96 \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

→ 2^{ème} cas : si $\sigma_1^2 = \sigma_2^2$

$(\theta = \mu_1 - \mu_2 | \text{données}) \sim t_{n_1+n_2-2} \left(\bar{y}_1 - \bar{y}_2, s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right)$

avec $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

IC (95%) pour $\theta = (\bar{y}_1 - \bar{y}_2) \pm t_{n_1+n_2-2}(0.975) \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$

→ 3^{ème} cas : si $\sigma_1^2 \neq \sigma_2^2$

$(\theta = \mu_1 - \mu_2 | \text{données}) \approx t_{df} \left(\bar{y}_1 - \bar{y}_2, \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)$

avec $df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{1}{n_1-1} \left(\frac{s_1^2}{n_1} \right)^2 + \frac{1}{n_2-1} \left(\frac{s_2^2}{n_2} \right)^2}$ (Méthode de Satterthwaite)

NB : arrondir df à l'unité inférieure par "conservatisme"

• Comparaison de deux moyennes (données pairées) :

→ $(\mu_D | \text{données}) \sim t_{n_D-1} \left(\bar{d}, \frac{s_D^2}{n_D} \right)$

IC (95%) pour μ_D : $\bar{d} \pm t_{n_D-1}(0.975) \frac{s_D}{\sqrt{n_D}}$

EXERCICES À FAIRE À LA SÉANCE DE TP

Estimation d'une moyenne isolée

1. Une machine automatique remplit des paquets dont le poids théorique doit être de 40 grammes. Les poids observés dans un échantillon au hasard de quinze paquets sont les suivants:

41 ; 40,2 ; 39,5 ; 40,3 ; 40,5 ; 40 ; 39,8 ; 40,1 ; 40 ; 39,2 ; 39,2 ; 39,5 ; 40 ; 39,5 ; 39,5

Si l'on suppose la normalité des poids,

- Déterminez la densité a posteriori du poids moyen (μ).
 - Déterminez si la machine ne lèse pas l'entreprise en remplissant trop les paquets.
 - Déterminez un ensemble de valeurs plausibles pour le poids moyen.
2. Le conseil d'administration de la section nourriture et nutrition de l'académie nationale des sciences a établi que la dose journalière recommandée en fer pour les femmes de moins de 51 ans est de 18mg. Les apports en fer, en milligrammes, durant une période de 24 heures ont été obtenus pour un échantillon de 45 femmes de moins de 51 ans. Voici les données:

15	18,1	14,4	14,6	10,9	18,1	18,2	18,3	15
16	12,6	16,6	20,7	19,8	11,6	12,8	15,6	11
15,3	9,4	19,5	18,3	14,5	16,6	11,5	16,4	12,5
14,6	11,9	12,5	18,6	13,1	12,1	10,7	17,3	12,4
17	6,3	16,8	12,5	16,3	14,7	12,7	16,3	11,5

- Déterminez la densité a posteriori de la dose journalière moyenne en fer pour les femmes de moins de 51 ans.
 - Déterminez si les femmes de moins de 51 ans reçoivent en moyenne une dose journalière de fer moindre que la dose journalière recommandée.
 - Déterminer un ensemble de valeurs plausibles pour l'apport journalier moyen en fer chez de telles femmes. Que peut-on en conclure ?
3. Un échantillon de 145 personnes a été prélevé dans la population des touristes étrangers qui passent leurs vacances en France. Les dépenses moyennes par jour et par personnes pour ces 145 touristes s'élèvent à un montant de 22,5 euros, avec un écart-type de 5,95 euros.
- Ce résultat est-il en contradiction avec les 21€ mesurés l'année précédente lors d'une enquête de grande ampleur ?
 - Déterminez un ensemble de valeurs plausibles pour les dépenses journalières d'un touriste étranger passant ses vacances en France.
4. Une personne évalue chaque semaine son taux de cholestérol. Voici les mesures (en g/l) ainsi récoltées au cours des 12 dernières semaines :

1,5 ; 1,8 ; 1,7 ; 1,6 ; 2,1 ; 1,9 ; 1,8 ; 2,2 ; 1,7 ; 1,7 ; 1,5 ; 1,6

Si l'on suppose la normalité du taux de cholestérol,

- a. Déterminez un ensemble de valeurs plausibles pour la moyenne de son taux de cholestérol.
- b. Cette personne désire estimer son taux de cholestérol annuel moyen avec une précision de 0,1 g/l. Combien de mesures faut-il effectuer pour satisfaire à cette exigence de précision?

Comparaison de deux moyennes (échantillons indépendants)

1^{er} cas : si n_1 et n_2 sont > 20 → Approximation normale

5. Un échantillon aléatoire simple de 1231 divorces construit à partir des registres municipaux de l'année 1981 montre que la durée moyenne des mariages se terminant par un divorce est de 12,9 ans. Une étude similaire, conduite en 1985, à partir de 1743 observations a donné une moyenne de 12 ans. Il y a lieu de croire que l'écart-type de la durée de mariage n'a pas été modifié entre 1981 et 1985 et que sa valeur est 4,2 ans.
 - a. Déterminez si la durée moyenne des mariages se terminant par un divorce a diminué entre 1981 et 1985.
 - b. Déterminez un ensemble de valeurs plausibles pour cette différence de durées moyennes des mariages.

2^{ème} cas : si $\sigma_1^2 = \sigma_2^2$

6. Dans une fouille archéologique, une question quant à l'origine de certains vases excavés de deux chantiers différents a été posée. La fouille du premier chantier a permis de dégager 8 fragments de vases différents, alors que celle du second chantier 6 fragments seulement. Une thèse archéologique soutient l'hypothèse que les vases des deux chantiers proviennent du même atelier. Pour tester cette hypothèse, le diamètre de l'orifice supérieur de chaque fragment a été mesuré :

Chantier 1: 12 11 11 14 11 12 13 12

Chantier 2: 10 12 12 11 9 10

On suppose que les vases excavés constituent un échantillon représentatif de l'ensemble des vases anciennement produits dans les deux localités. De plus, on suppose que le diamètre des orifices suit une loi normale dont la variance est la même pour les deux localités.

- a. Calculez le diamètre moyen des vases excavés dans chaque chantier et obtenez une estimation de la variance de la différence des diamètres moyens.
- b. Déterminez si la thèse archéologique soutenant l'hypothèse selon laquelle les vases des deux chantiers proviennent du même atelier est plausible. (suggestion : utiliser la distribution a posteriori de la différence des diamètres moyens des vases excavés).
- c. Déterminez un ensemble de valeurs plausibles pour cette différence de diamètres moyens.

3^{ème} cas : si $\sigma_1^2 \neq \sigma_2^2$

7. On souhaite comparer l'âge auquel les jeunes quittent le domicile parental chez les hommes et chez les femmes. Dans la population globale, l'hypothèse selon laquelle les garçons quittent en moyenne le domicile des parents plus tardivement que les filles a déjà été confortée par plusieurs études. Toutefois, cet âge varie assez fortement suivant les situations, d'études notamment. C'est la raison pour laquelle on vous demande de déterminer, sur base d'un

échantillon de 16 femmes et de 13 hommes issus d'une promotion de Médecine si la différence d'âge moyen de départ du domicile parental entre hommes et femmes suit la même tendance dans ce groupe particulier par rapport à celle observée dans la population générale.

Les données récoltées sont les suivantes :

- Chez les hommes :

Âge	18	23	24	25	26	27	28	29	32
Fréquence	1	2	1	2	2	2	1	1	1

- Chez les femmes :

Âge	22	23	24	25	26	27
Fréquence	4	3	3	4	1	1

Comparaison de deux moyennes (données pairées)

8. Une firme étudie l'influence des pauses-café sur la productivité de ses ouvriers. Ayant choisi 6 ouvriers au hasard, on évalue leur productivité (mesurée par le nombre de pièces produites) durant deux jours, le premier sans interruption, le deuxième avec interruption.

Ouvrier	1	2	3	4	5	6
Sans pause-café	23	35	29	33	43	32
Avec pauses-café	28	38	29	37	42	30

- Ces résultats suggèrent-ils que les pauses-café améliorent la productivité moyenne des ouvriers ?
- Déterminez un ensemble de valeurs plausibles pour cette différence de productivité moyennes.

EXERCICES COMPLÉMENTAIRES

- Dans un test de fabrication de composantes d'une chaîne Hifi, la baisse de puissance de sortie des circuits électriques après 2000 heures d'utilisation a été mesurée. Un essai sur 80 composantes identiques a donné une estimation de la baisse de puissance moyenne égale à 12 watts et une variance estimée de 5 watts-carré. Déterminer un ensemble de valeurs plausibles pour la baisse de puissance des composantes d'une chaîne Hifi.
- Un négociant en vin s'intéresse à la contenance des bouteilles d'un cru d'une année déterminée. Il se demande si la contenance moyenne n'est pas inférieure à la contenance légale de 75 cl. A cet effet, il mesure le contenu de 10 bouteilles (sélectionnées par échantillonnage aléatoire simple) et obtient les valeurs suivantes:
73,2 ; 72,6 ; 74,5 ; 75 ; 75 ; 73,7 ; 74,1 ; 75,1 ; 74,8 ; 74
 - En supposant la normalité de la distribution du contenu, peut-il conclure que le contenu moyen des bouteilles de cette production est inférieur à 75cl? Explicitiez vos hypothèses de travail.
 - Déterminer un ensemble de valeurs plausibles pour le contenu moyen des bouteilles de la production.

- c) Combien de mesures auraient été nécessaires pour obtenir une précision de 0,5 dans l'estimation du contenu moyen des bouteilles de cette production?
- d) Même question mais pour une précision de 0,1?
3. La compagnie Paul Lissier produit des feux de circulation; elle a décidé d'ajouter un microordinateur à l'équipement de contrôle de la production afin d'en augmenter l'efficacité. Les micro-ordinateurs de deux fabricants sont jugés adéquats pour remplir cette fonction. La compagnie Lissier achètera des micro-ordinateurs des deux fournisseurs s'il n'y a pas de différence significative de durabilité entre les deux marques. A partir d'un échantillon de 35 micro-ordinateurs de la marque A et d'un échantillon de 32 micro-ordinateurs de la marque B, on a obtenu les moyennes d'échantillonnage 2800 heures et 2750 heures respectivement. Les écart-types d'échantillonnage suivants ont été également calculés: 180 heures pour la marque B et 200 heures pour la marque A. Sur base d'un intervalle de crédibilité adéquat, peut-on mettre en évidence une différence de durabilité moyenne entre les deux marques?
4. Un peintre désire comparer deux types de peintures utilisées pour peindre les bandes blanches sur les routes. Dix bandes pour chaque peinture sont tracées sur une route principale. La durée (en mois) de persistance des bandes sont présentées dans la table suivante:

Type A	35,6	37	34,9	36	36,6	36,1	35,8	34,9	38,8	36,5
Type B	37,2	39,7	37,2	38,8	37,7	36,4	37,5	40,5	38,2	36,6

Peut-on mettre en évidence une différence entre les durées moyennes de persistance de la bande pour les deux types de peintures? Précisez vos hypothèses de travail.

5. Aux Etats-Unis, une étude statistique est conduite mensuellement dans le but d'estimer le salaire horaire d'employés indépendants dans diverses industries. A partir de deux échantillons de 14 travailleurs miniers et de 17 ouvriers du bâtiment, on a obtenu les moyennes horaires 12,93\$ et 13,42\$ respectivement. Les écart-types ont été également calculés: 2,17\$ pour les mineurs et 2,29\$ pour les ouvriers du bâtiment. Il y a lieu de supposer que la variance des salaires des ouvriers du bâtiment est supérieure à celle des travailleurs miniers au niveau de la population.
- a) Peut-on, sur cette base, affirmer que les travailleurs miniers ont un revenu horaire moyen inférieur? Expliquez vos hypothèses de travail.
- b) Répondre à la question en raisonnant à partir d'un ensemble de valeurs plausibles adéquat.
6. Un professeur désire comparer deux méthodes d'enseignement de l'algèbre dans le secondaire. La première repose sur la lecture tandis que la seconde est une méthode personnalisée d'enseignement (MPE). Des paires d'étudiants ayant des niveaux mathématiques similaires sont formées. Un échantillon aléatoire de 11 paires est sélectionné. Pour chaque paire, l'étudiant qui suivra le cours par la méthode de lecture est déterminé au hasard, l'autre suivant le cours par la MPE. Les deux cours d'algèbre sont

donnés par le même professeur. Les cotes finales (sur 100) des étudiants sont reprises dans la table suivante:

Lecture	66	93	36	84	60	66	80	73	74	83	52
MPE	67	93	35	85	64	57	79	70	67	79	50

- a) Les données suggèrent-elles de conclure à une différence de performances moyennes des étudiants selon la méthode d'enseignement (si l'on suppose qu'une plausibilité supérieure à 0,90 est une indication convaincante)? Explicitiez vos hypothèses de travail.
- b) Répondez à la question sur base d'un ensemble de valeurs plausibles adéquat.

7. Une compagnie américaine fabriquant des pneus a développé deux nouveaux procédés (A et B) pour rendre plus résistantes les ceintures radiales d'acier du pneu. Afin de comparer les deux procédés de fabrication, 50 pneus produits par chaque procédé sont choisis aléatoirement. Chacun des 50 pneus fabriqués par le procédé A (pneu A) est assigné aléatoirement à l'avant gauche ou l'avant droit de 50 voitures. Si, pour une voiture donnée, le pneu A est assigné à l'avant droit, un pneu fabriqué par le procédé B (pneu B) sera assigné à l'avant gauche et inversement. Les pneus arrière de chacune des 50 voitures proviendront de cette même compagnie mais seront de fabrication classique. La différence entre les durées de vie (durée du pneu A – durée du pneu B), exprimée en milliers de miles et arrondie au centième de miles, pour les 50 paires de pneus sont données dans la table ci-dessous:

-0,8	-2,4	-0,2	1,6	-3,6	-0,6	-1,7	-1,4	-2,5	-1,8
-0,9	0,9	2,8	-2,7	-0,6	0,0	-2,0	1,8	-1,5	1,9
4,2	-2,1	3,0	0,0	0,4	0,9	0,9	1,6	-2,7	0,9
1,0	-2,3	-0,4	0,5	2,4	-0,3	0,9	-0,6	0,8	3,6
-0,4	0,5	1,3	-4,0	2,4	-4,3	2,4	-0,5	1,5	3,1

Les données suggèrent-elles de conclure à une différence de durées de vie entre le procédé de fabrication A et le B? Explicitiez vos hypothèses de travail.

RÉPONSES EXERCICES COMPLÉMENTAIRES : CHAPITRE 3 – 1^{ÈRE} PARTIE

Exercice 1 : Estimation d'une moyenne isolée \rightarrow [11,51 ; 12.49]

Exercice 2 : Estimation d'une moyenne isolée

- a) $P \ni]0.99 ; 0.995[$
- b) [73.6 ; 74.8]
- c) $n = 14$
- d) $n = 274$

Exercice 3 : Comparaison de deux moyennes (échantillons indépendants)

1^{er} cas : si n_1 et n_2 sont $> 20 \rightarrow$ Approximation normale

[-40.99 ; 140.99]

Exercice 4 : Comparaison de deux moyennes (échantillons indépendants)

2^{ème} cas : si $\sigma_1^2 = \sigma_2^2$

$P \ni]0.995 ; 0.9975[$ ou [-2.92 ; -0.6]

Exercice 5 : Comparaison de deux moyennes (échantillons indépendants)

2^{ème} cas : si $\sigma_1^2 = \sigma_2^2$

- a. $P < 0.75$
- b. [-2.14 ; 1.16]

Exercice 6 : Comparaison de 2 moyennes (données pairées)

- a) $P \ni]0.9 ; 0.95[$
- b) [0.59 ; 4.41]

Exercice 7 : Comparaison de 2 moyennes (données pairées) \rightarrow [-0.54 ; 0.58]