

T. D. n° 5

Intervalles de confiance.

Pour traiter ce T.D., il faut savoir que certains exercices sont à traiter intégralement avec le logiciel R.

D'autres exercices sont à résoudre avec une utilisation partielle du logiciel R.

Exercice 1 Intervalle de confiance pour une espérance, pages 246, 247, 248 et 249 du livre « Initiation à la statistique avec R ».

Considérons le fichier `Mesures` qui est dans le package `BioStatR`. Nous voulons construire un intervalle de confiance de l'espérance μ de la variable `taille` des glycines blanches. Pour cela, suivre la démarche suivante.

- Extraire les 54 données des glycines blanches des 252 données des quatre espèces en vous servant de la fonction `subset()`. Si vous ne savez pas vous servir de la fonction `subset()`, alors consultez l'aide en ligne. N'oubliez pas de donner un nom à votre jeu de données qui provient de l'extraction.
- Calculer un intervalle de confiance à 95% pour l'espérance de la variable `taille` des glycines blanches.
- Conclure.

Exercice 2 Intervalle de confiance pour une variance, pages 249, 250 et 251 du livre « Initiation à la statistique avec R ».

Considérons le fichier `Mesures` qui est dans le package `BioStatR`. Nous voulons construire un intervalle de confiance pour la variance σ^2 de la variable `taille` des glycines blanches dont l'espérance est inconnue. Pour cela, suivre la démarche suivante.

- Extraire les 54 données des glycines blanches des 252 données des quatre espèces en vous servant de la fonction `subset()`. Si vous ne savez pas vous servir de la fonction `subset()`, alors consultez l'aide en ligne. N'oubliez pas de donner un nom à votre jeu de données qui provient de l'extraction.
- Calculer un intervalle de confiance à 95% pour la variance de la variable `taille` des glycines blanches.
- Conclure.

Exercice 3 Intervalle de confiance pour une proportion, page 252 à page 256 du livre « Initiation à la statistique avec R ».

Considérons le fichier `Mesures` qui est dans le package `BioStatR`. Nous voulons construire un intervalle de confiance pour la proportion des gousses de glycine blanche qui ont moins de trois graines présentes dans une gousse. Pour cela, suivre la démarche suivante.

- a) Extraire les 54 données des glycines blanches des 252 données des quatre espèces en vous servant de la fonction `subset()`. Si vous ne savez pas vous servir de la fonction `subset()`, alors consultez l'aide en ligne. N'oubliez pas de donner un nom à votre jeu de données qui provient de l'extraction.
- b) Calculer un intervalle de confiance à 95% pour la proportion des gousses de glycine blanche qui ont moins de trois graines présentes dans une gousse.
- c) Conclure.

Exercice 4 Intervalle de confiance pour une espérance.

Une machine fabrique des billes métalliques dont le poids, mesuré en grammes, suit une loi normale. Nous prélevons au hasard 10 billes. Leurs poids sont

19,6; 20; 20,2; 20,1; 20; 19,9; 20; 20,3; 20,1; 19,8.

1. Quel est l'intervalle de confiance à 95% du poids des billes métalliques fabriquées ?
2. En réalité, l'écart-type σ de la population est connu et égal à 0,2. Quel est l'intervalle de confiance à 95% du poids des billes métalliques fabriquées ?

Exercice 5 Intervalle de confiance pour une espérance.

Voulant évaluer rapidement les résultats obtenus par ses 200 étudiants lors d'un partiel, un professeur décide de corriger quelques copies tirées au hasard. Il admet par ailleurs que les notes de ses élèves suivent une loi normale de variance 4.

1. Le professeur corrige un échantillon de 7 copies et trouve une moyenne de 11. Quel est l'intervalle de confiance à 95% de la moyenne des 200 copies ?
2. Combien de copies le professeur doit-il corriger s'il veut situer la moyenne générale de ses élèves dans un intervalle de confiance d'amplitude 2, avec un risque de 5% ?
3. En trouvant une moyenne égale à 11, combien de copies le professeur devrait-il corriger pour pouvoir dire, avec un risque de 1%, que la moyenne de tous les élèves est supérieure à 10 ?

Exercice 6 Intervalle de confiance pour une espérance.

Une entreprise fabrique un certain type de composants électroniques dont la durée de vie X , exprimée en heures, est une variable aléatoire qui suit une loi normale. Des mesures effectuées sur un échantillon aléatoire de taille $n = 50$ ont donné les résultats suivants :

$$\sum_{i=1}^{50} x_i = 60\,000; \quad \sum_{i=1}^{50} x_i^2 = 74 \times 10^6.$$

1. Donner une estimation ponctuelle de la durée de vie moyenne des composants.
2. Donner une estimation ponctuelle de l'écart-type de cette durée de vie.
3. Donner l'intervalle de confiance à 95% de cette durée de vie moyenne.
4. Quelle aurait du être la taille de l'échantillon pour que l'intervalle de confiance à 95% de la durée de vie moyenne des composants ait une amplitude de 60 heures?

Exercice 7 Intervalle de confiance pour une proportion.

À la veille d'une consultation électorale, nous effectuons un sondage.

1. Dans un échantillon représentatif de 1000 personnes, 500 personnes déclarent vouloir voter pour Dupond, 250 pour Durand et 50 pour Duroc. Donner les intervalles de confiance à 95% et 99%, en utilisant la méthode de Wald, de la proportion de personnes ayant l'intention de voter Dupond, Durand ou Duroc.
2. Nous évaluons le pourcentage de personnes ayant l'intention de voter pour un quatrième candidat, Duval, à 17%. Combien faut-il interroger de personnes pour obtenir un intervalle de confiance à 95% de la proportion de personnes ayant l'intention de voter Duval, avec une précision de 1%?

Remarques :

1. Cette précision est aussi parfois appelée « marge d'erreur ».
2. La précision est égale à l'amplitude de l'intervalle divisée par deux.

Exercice 8 L'airbag dans les automobiles. D'après le CC3 de l'année 2014-2015.

L'airbag (ou coussin gonflable) est un système de sécurité de plus en plus souvent installé dans les automobiles. Son gonflement est assuré par un dispositif pyrotechnique dont les caractéristiques sont la moyenne et l'écart-type du délai entre la mise à feu et l'explosion. Lors de l'étude d'un certain dispositif d'allumage, les résultats des mesures qui proviennent d'une loi normale, effectués sur 30 exemplaires, ont été (en millisecondes) :

28,0	28,0	31	31,0	32	33	32,5	29	30,5	31
28,5	27,5	32	29,5	28	26	30,0	31	32,5	33
27,5	29,0	30	28,5	27	25	31,5	33	34,5	29

1. Calculer, au risque 5%, l'intervalle de confiance de la moyenne du délai si nous connaissons l'écart-type de la population de référence et qu'il est égal à 2.

2. Calculer, au même risque, ce même intervalle si nous ne connaissons pas l'écart-type de la population de référence.
3. Comparer les deux intervalles obtenus aux questions précédentes et expliquer pourquoi vous observez une différence entre les deux.
4. Calculer, au même risque, l'intervalle de confiance de la variance du délai, dont nous déduirons celui de l'écart-type dans le cas où nous ne connaissons pas l'écart-type de la population de référence.

Exercice 9 Les vers à soie. D'après le CC3 de l'année 2015-2016.

Un sériculteur a pesé 100 cocons provenant de son élevage de vers à soie. Il a obtenu les résultats suivants :

Poids en g	0,64	0,65	0,66	0,67	0,68	0,69	0,70	0,71
Effectif	3	5	2	6	6	10	12	10
Poids en g	0,72	0,73	0,74	0,75	0,76	0,77	0,78	0,79
Effectif	9	8	8	6	5	4	3	3

Nous appelons μ la moyenne des masses de tous les cocons de l'élevage et nous supposons que la variable « masse d'un cocon » notée X suit une loi normale de variance $0,0016g^2$.

1. Préciser les deux paramètres de cette loi normale.
2. Nous choisissons au hasard un échantillon de 100 cocons. Nous appelons X_i la masse du i -ème cocon de cet échantillon. Nous supposons que les variables X_1, \dots, X_{100} sont mutuellement indépendantes, de même loi que la variable X . Nous posons :

$$T_{100} = \frac{1}{100} \sum_{i=1}^{100} X_i.$$

Quelle loi suit la variable aléatoire T_{100} ? Préciser ses paramètres. Vous justifierez vos réponses.

3. Construire un intervalle de confiance pour μ au niveau de confiance de 95%.