

Le fichier « prêt_acceptation.xlsx » décrit les informations concernant $n = 944$ demandeurs de crédit. La variable « acceptance » joue un rôle particulier, elle indique si la demande de prêt a été couronnée de succès (oui) ou non.

1. Estimations et tests de comparaison à un standard sous Excel

Pour chaque question, créez une feuille spécifique étiquetée avec le numéro de la question.

1. Quelle est la proportion des « acceptance = oui » ?
2. Calculez les bornes basses et hautes de l'intervalle de confiance à 95% de cette proportion. Comment interpréter cet intervalle de confiance ?
3. Effectuer le test d'hypothèses suivant à 5% : H_0 : proportion des acceptations = 0.5 vs. H_1 : proportion > 0.5. Que constatez-vous ? Peut-on rapprocher ce résultat à celui de la question précédente ?
4. Calculez la moyenne et l'écart-type (échantillon, attention aux fonctions proposées par Excel) de la variable RevenuMenage.
5. Calculez l'intervalle de confiance à 95% de la moyenne. Comment interpréter cet intervalle ?
6. Calculez les effectifs dans chaque intervalle si l'on découpe RevenuMenage avec la méthode des intervalles de largeur égales. A partir de ces effectifs, créez le graphique « histogramme de fréquences ». Que constatez-vous ? A posteriori, que peut-on penser de l'intervalle calculé à la question précédente ?
7. Transformez la variable RevenuMenage avec l'opérateur logarithme (népérien). Appelez la variable LogRevMenage. Reconstituez l'histogramme de fréquence. Que constatez-vous maintenant ?
8. Calculez l'intervalle de confiance à 95% de la moyenne de LogRevMenage. Puis transformez les bornes de manière à les exprimer dans l'unité initiale de RevenuMenage. Comparez vos résultats avec l'approche développé à la question n°5. Quel commentaire peut-on faire ?
9. Calculez l'intervalle de confiance à 95% de la variance de l'âge.
10. Créez une variable (DifSalaire) qui fait la différence entre les salaires de l'homme et de la femme à l'intérieur de chaque ménage. Calculez l'intervalle de confiance à 95% de la moyenne de DifSalaire. Quelle conclusion peut-on tirer du résultat ?
11. Confrontez les hypothèses à 5% : H_0 : DifSalaire = 0 ; H_1 : DifSalaire \neq 0. Que peut-on conclure ? Le résultat est-il cohérent avec celui de la question précédente ?
12. Créez une variable SupSalaire qui comptabilise la proportion des ménages où le salaire de l'homme est supérieur à celui de la femme. S'il y avait parité globalement dans les ménages, à quelle proportion de référence devrait-on la comparer ? Utiliser l'outil adapté pour vérifier cette hypothèse de parité.
13. Quelle est la proportion des crédits acceptés chez les personnes qui ont pris une assurance ? Qui n'ont pas pris d'assurance ? Calculez les intervalles de confiance à 95 % de ces deux

proportions : peut-on considérer qu'elles sont significativement différentes à la lumière des résultats ?

14. Formez le tableau de contingence croisant l'acceptation de crédit et l'assurance. Peut-on conclure que les deux variables sont liées au sens du test d'indépendance du KHI-2 à 5 % (cf. par ex. [Wikipédia](#)) ? Ce résultat conforte-t-il celui obtenu dans la question précédente ?
15. Calculer les intervalles de confiance à 95% des moyennes de LogRevMenage selon l'acceptation de crédit. Peut-on considérer qu'elles sont significativement différentes à la lumière des résultats ?

2. Estimation ponctuelle et par intervalle sous R

Refaites le même exercice sous R. Programmez explicitement les calculs à l'aide des fonctions standards (mean, sd, table, etc.), ne passez pas par des packages spécifiques.

Remarque : Pour la construction des histogrammes de fréquences (n°6), il n'est pas nécessaire de discrétiser explicitement les variables, utilisez directement la fonction hist() de R.