Statistique - Planche de TD n°3
L2 S2 - Economie et Finance - Economie Gestion, DU, et CMI*

1 Estimation

1.1 Cours*

1) A quoi sert un estimateur ?

2) Quelles sont les deux propriétés souhaitables d'un bon estimateur ?

3) Quelle est la différence entre un intervalle de pari et un intervalle de confiance?

4) Quelles sont les deux méthodes qui permettent de construire un estimateur?

5) Nombre ou variable aléatoire ? a) La variance empirique. b) L'espérance de la moyenne empirique. c) Une réalisation de la fréquence empirique. d) La fréquence empirique.

6) Variance Empirique Corrigée. La variance empirique d'une variable aléatoire X vaut 2427 pour un échantillon de taille 88. Combien vaut la variance empirique corrigée ? Donnez votre réponse accompagnée de l'expression formelle des 2 variances.

2 Marge d'erreur

2.1 Calcul 1*

La marge d'échantillonnage, ou marge d'erreur, ME, correspond à la demi-amplitude de l'intervalle de confiance. Pour un intervalle de confiance de niveau 98% sur une moyenne, dans le cas d'un échantillon gaussien, avec écart-type connu valant 15, quel doit être la taille de l'échantillon pour que la marge d'erreur soit égale à 5 ?

*Les exercices les plus importants sont marqués d'un *.

2.2 Calcul 2: élections

On vous demande de faire une enquête afin de déterminer le pourcentage des votants qui exerceront leur droit de vote lors des prochaines élections présidentielles. On exige de vous une estimation par intervalle avec un niveau de confiance de 95%, la marge d'erreur ne devant pas dépasser points de pourcentage.

1) Quel est le nombre minimal de personnes que vous devez interroger pour respecter les conditions imposées?

2) Quel serait ce nombre si on vous demandait une précision de 1,5% ?

3 Estimation par Intervalle

3.1 Maroquinerie

Un artisan qui fabrique des objets de maroquinerie veut estimer le nombre moyen m de porte-cartes qu'il vend quotidiennement. En notant ses ventes sur 36 jours, il arrive à une moyenne de ventes de 120. En supposant que l'écart-type des ventes quotidiennes est égal à 17 pour ce produit, construire un intervalle de confiance pour m au seuil de 99%.

3.2 Sushis*

Un petit commerce de fabrication et de livraison de sushis à domicile désire faire une étude sur le temps moyen qui s'écoule entre le moment où la commande est passée par téléphone et le moment où le client est livré. L'étude au départ porte sur les livraisons dans un rayon de 2km autour du restaurant. Une observation rapide faite sur 25 commandes fait ressortir un temps moyen de 27 min avec un écart-type de 1 min. On suppose que les temps de fabrication-livraison se distribuent normalement. Donner un intervalle de confiance de ce temps moyen, au niveau de confiance de 98%.

3.3 Salaires

On s'intéresse au salaire journalier moyen en euros de salariés de la branche metallurgie. Un échantillon de 500 salariés a donné les résultats suivants:

$$\Sigma_{i=1}^{500} s_i = 22500 \text{ et } \Sigma_{i=1}^{500} \left(s_i - \overline{s}\right)^2 = 12500$$

où S_i est la variable aléatoire salaire journalier en euro pour le salarié i. Donnez un intervalle de confiance à 95% du salaire moyen journalié distribué dans cette branche de l'industrie.

3.4 Web

Afin que les ressources informatiques disponibles d'un institut universitaire soient utilisées au mieux, une étude a porté sur le temps que passe en moyenne un étudiant sur le réseau Internet de l'institut par jour. On peut raisonnablement penser que ce temps suit une loi normale. 20 étudiants ont été observés, et il ressort qu'en moyenne, ces étudiants passent 1H30 par jour sur le Web. La dispersion mesurée par l'écart-type donne sur cet échantillon la valeur de 12 min. Quelle estimation peut-on donner au niveau de la population étudiante de cet institut pour le temps moyen quotidien passé sur Internet ? On donnera la réponse sous la forme d'un intervalle de confiance à 95%.

4 Construction d'estimateurs

4.1 Maximum de vraisemblance

Le total des ventes hebdomadaires d'un produit alimentaire dans un magasin i, $1 \leq n$ est une variable aléatoire X_i->$N(m_i, \sigma)$, où les valeurs m_i et σ sont supposées connues. Une campagne publicitaire de ce produit ont pour effet d'augmenter les ventes, de telle sorte que chaque moyenne est augmentée de la quantité a.

1) Déterminer un estimateur de a construit à partir d'observations indépendantes $(X_1, ..., X_n)$ des ventes après la campagne publicitaire, et étudier ses propriétés. Construire ensuite un intervalle de confiance de niveau 0,95.

2) Appliquez vos deux estimateurs aux données suivantes dans le cas où $\sigma = 3$:

m_i	98	101	104	99	100	102	95	97	105	103
x_i	109	105	110	106	110	114	108	104	115	118

4.2 Méthode des moments

Soit $(X_1, ..., X_n)$ un échantillon d'une variable aléatoire X de loi normale d'espérance et de variance égales à un paramètre inconnu $\theta > 0$.

1) Déterminer 2 estimateurs de θ par la méthode des moments, étudier leurs précisions respectives.

2) Construire un intervalle de confiance pour θ de niveau 0,95 ayant observé que $\Sigma_{i=1}^{25} x_i = 50,23$ et $\Sigma_{i=1}^{25} (x_i - \overline{x})^2 = 48,12$

5 Comparaison de moyennes

5.1 Temps de sommeil des femmes et des hommes

L'INSEE mène régulièrement des enquêtes sur les temps sociaux journaliers. Parmi ceux-ci le temps libre est compté. L'INSEE a interrogé au hasard des femmes et des hommes actifs.

- 14 femmes ont noté leur temps de sommeil quotidien (il s'agit en fait de la moyenne calculée sur 2 semaines)

- 12 hommes ont également noté leur temps de sommeil. Les résultats exprimés en minutes sont reportés dans le tableau suivant:

femmes	498	435	460	511	487	510	502	460	470	494	510	570	422	478
hommes	414	379	404	446	514	366	431	438	442	422	454	455		

Y-a-t-il une différence entre temps de sommeil des hommes et des femmes ? Construisez votre raisonnement en suivant les consignes suivantes:

- on considère que la distribution de la variable aléatoire temps de sommeil est gaussienne
- le niveau de l'intervalle que vous devez construire est fixé à 98%
- vous devez rédiger précisément votre réponse.

5.2 Comparaison de ventes*

Le directeur RH d'une importante chaine de magasins de meubles s'interroge sur l'efficacité d'une couteuse formation complémentaire à la vente. Avant d'imposer cette formation à tous les vendeurs de la chaine, le DRH décide de commander une étude statistique. Les montants des ventes de 7 vendeurs pour le mois de mai de l'année précédente sont relevés (échantillon A). Ces 7 vendeurs suivent une formation et leurs montants des ventes du mois de mai de l'année suivante sont reportés (échantillon B). Les résultats sont représentés dans le tableau suivant:

mai 2015	1024	345	647	544	978	241	450
mai 2016	874	451	450	875	1201	346	374

En supposant qu'une étude préalable a montré que le montant des ventes ainsi que la différence des montants des ventes entre les 2 années suivent une distribution gaussienne, construisez un intervalle de confiance de niveau 95% pour la comparaison des montants moyens des ventes. La formation est-elle efficace ?

6 Sujet antérieur

L'entreprise MANGER PAS CHER qui exploite un service de restauration réalise un sondage auprès de n clients qui fréquentent le restaurant sans prendre de repas. A la question "le prix du repas vous paraît-il trop élevé ?" k personnes ont répondu "oui".

Soit p la proportion véritable de réponses "oui" dans l'ensemble des clients qui fréquentent le restaurant sans prendre de repas et F_n la variable aléatoire proportion des individus qui ont répondu "oui" parmi n clients interrogés.

1) Donner la loi exacte suivie par la variable Fn

2) Justifier que F_n est un estimateur sans biais de p

3) Donner un intervalle de confiance à 95% de la proportion p pour $n = 400$ et $k = 250$

4) Si l'on veut, à 95%, que l'intervalle de confiance pour p soit [0.50,0.75]