UNIVERSITE DE LOME **ECOLE POLYTECHNIQUE DE LOME**

TD STATISTIQUE INFERENTIELLE

Parcours : LF/S4 Année : 2023-2024

Exercice 1 (pour devoir)

Une entreprise fabrique des sacs en plastique pour les enseignes de distribution. Elle s'intéresse au poids maximal que ces sacs peuvent supporter sans se déchirer.

On suppose ici que le poids maximal que ces sacs peuvent supporter suit une loi normale d'espérance mathématique 58 Kg et d'écart-type 3 Kg.

1. Sur 200 sacs reçus, une grande enseigne de distribution constate un poids moyen de 57,7 Kg.

1.1. Donner un intervalle de confiance bilatéral de la moyenne des poids sur un échantillon de taille 200, au seuil de risque 1%.

1.2. Quelle est votre conclusion sur le poids moyen constaté ?

2. Donner le poids moyen dépassé dans 97% des cas, sur un échantillon de taille 200

Exercice 2

Les résultats d'une enquête, effectuée sur une population de 1500 salariés d'une entreprise, a montré que dans 65% des cas les individus avaient au moins un crédit en cours.

Trouver la probabilité pour que deux échantillons de 200 personnes chacun, indiquent plus de 10 points d'écart entre les proportions de personnes ayant au moins un crédit en cours.

Exercice 3

Afin de mieux gérer les demandes de crédits de ses clients, un directeur d'agence bancaire réalise une étude relative à la durée de traitement des dossiers, supposée suivre une distribution normale. Un échantillon non exhaustif de 30 dossiers a donné :

Durée mn	0 – 10	10 –20	20 – 30	30 – 40	40 – 50	50 – 60
Effectif	3	6	10	7	3	1

1. Calculer la moyenne et l'écart type des durées de traitement des dossiers de cet échantillon.
2. En déduire les estimations ponctuelles de la moyenne m et de l'écart type σ de la population des dossiers.
3. Donner une estimation de m par intervalle de confiance au seuil de risque 5%

Exercice 4

La société G@E a mis au point un logiciel de gestion destiné essentiellement aux PME. Après une enquête, dans la région Aquitaine, auprès de 100 entreprises déjà équipées d'un matériel informatique (micro-ordinateur) apte à recevoir ce logiciel, la société G@E décide de fixer le prix de vente à 200 €.

Elle espère diffuser son produit auprès de 68% des PME de la région (cette valeur constituera la proportion de ventes sur l'échantillon).

On peut admettre que les 100 PME interrogées constituent un échantillon représentatif des 4 500 PME formant le marché potentiel.

1. Déterminer l'intervalle de confiance de la proportion p des entreprises intéressées par le logiciel, au seuil de risque 1%.
2. Quelle aurait dû être la taille de l'échantillon pour que l'amplitude de l'intervalle de confiance soit de 20 points (erreur de 0,1).

Exercice 5

On veut savoir si la résistance moyenne de composants produits dans une usine est 400Ω. On considère que la distribution des résistances est normale, et on mesure pour 16 composants les valeurs 392, 396, 386, 389, 388, 387, 403, 397, 401, 391, 400, 402, 394, 406, 406, 400.

(a) Donner les estimations ponctuelles des moyenne et variance.

(b) Peut-on considérer, au seuil de signification $\alpha = 5\%$, que le lot respecte la norme de 400Ω ? Même question avec un seuil de $\alpha = 1\%$.

Exercice 6

Un fabricant se vante de proposer des tubes à essai d'une durée de vie supérieure à 2000h de chauffage. A l'aide d'un échantillon de 100 tubes testés, on estime la durée de vie moyenne à 1975h, avec un écart-type de 130h. Peut-on affirmer, au risque 5%, que le fabriquant ment ?

Exercice 7 (pour devoir)

On relève chaque jour pendant 200 jours le nombre d'atterrissages entre 14h et 15h dans un aéroport :

Nb d'atterrissages	0	1	2	3	4	5	6	7	8	9	10
Nb de jours	11	28	43	47	32	28	7	0	2	1	1

(a) Soit X la variable « nombre d'atterrissages par jour entre 14h et 15h ». Donner les estimations ponctuelles de E(X) et Var(X) et estimer E(X) par un intervalle de confiance 95%. Ces résultats sont-ils compatibles avec une loi de poisson ? Quel serait son paramètre ?

(b) Tester la validité de ce modèle (test du $\chi 2$ au risque 5%).

(c) Calculez la probabilité d'avoir dans cet aéroport, toujours entre 14h et 15h : 0 atterrissage un jour donné, 1 ou 2 atterrissages un jour donné, 2 atterrissages en tout sur 3 jours quelconques.

<u>Exercice 8</u>

Soient $X_1,\dots X_n$ n-variables aléatoires indépendantes, telles que $E\{X_k\}=\mu$ et $Var(X_k)=\sigma^2$ pour $k = 1,\dots n$.

1. Montrer que $\overline{X}_n$ est un estimateur sans biais pour μ.
2. Montrer que $\overline{X}_n^2$ n'est pas un estimateur sans biais de μ^2. Déterminer son biais.
3. Déterminer m tel que $\overline{X}^2_n - m\tilde{S}^2_n$ soit un estimateur sans biais de μ^2.

<u>Exercice 9</u>

On veut estimer le rendement d'un engrais pour la culture du blé. Sur douze parcelles expérimentales, on a trouvé les rendements suivants en tonnes par hectare :

7,7; 8,4; 7,8; 8,2; 7,9; 8,5; 8,4; 8,2; 7,6; 7,8; 8,4; 8,3.

Donner un intervalle de confiance à 95% pour le rendement moyen de l'engrais (on supposera que le rendement à l'hectare est une variable aléatoire gaussienne).

<u>Exercice 10 (pour devoir)</u>

On s'intéresse, sur un site donné, à la concentration d'un polluant m´métallique que l'on note X(µg/g) mesurée dans des sédiments. On suppose que la variable $X - N(\mu_0, \sigma^2_0)$ où le paramètre populationnel $\sigma^2_0 = 100$ est connu. On dispose d'un échantillon $\{X1, \cdots Xn\}$ de taille n = 100.

1. On s'intéresse à $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,l'estimateur de la moyenne de la population. Avec les hypothèses précédentes, calculer son espérance et sa variance en fonction de μ_0, σ_0 et n. Qu'en d´déduisez-vous sur les qualités de cet estimateur ?

2. Quelle loi suit la variable $\overline{X}$?

3. On suppose maintenant que $\mu_0 = 50$ µg/g. Calculer les probabilités suivantes :

(a) $P(\overline{X} \leq 52)$

(b) $P(\overline{X} \leq 48)$

(c) $P(48 \leq \overline{X} \leq 52)$

(d) $P(48 \leq \overline{X} \leq 52/X \geq 50)$

(e) $P(\overline{X} = 48.25)$

4. En fait on a des doutes sur la valeur de μ_0. Sur un échantillon réalisé de taille n = 100, on a mesuré une concentration moyenne de x = 51,2 µg/g. Vérifier, au risque de 5%, si la valeur supposée de μ_0 est la bonne.

5. Construire un intervalle de confiance de la moyenne. Qu'en déduisez-vous ?

<u>Exercice 11</u>

Soit $(X_1, ..., X_{25})$ un échantillon de loi gaussienne d'espérance µ inconnue et de variance v = 100 connue.

1. Construire un test de niveau α = 0,10 de l'hypothèse nulle "µ = 0" contre l'hypothèse alternative "µ = 1,5", fondé sur la moyenne empirique, estimateur du paramètre µ. On observe x = 1. Quelle est la décision du test ?

Quelle est l'erreur de seconde espèce du test ?

2. Répondre à la question précédente si v = 9.

3. Comment modifier le test si l'alternative est "µ = −1,5" ?

4. On souhaite tester (H_0) : µ = 2 contre (H_1) : µ < 2. Définir la région de rejet.

Calculer la puissance du test et étudier ses variations en fonction de µ, n et σ.

<u>Exercice 12</u>

On considère un échantillon de n = 10 variables aléatoires suivant une même loi normale N (μ, σ^2).

1. On connait la variance $\sigma^2 = 2,5$. On a observé la valeur 1,15 de la moyenne empirique $\bar{X}_n$. Peut-on accepter, au seuil de 95%, l'hypothèse H_0 : µ = 0,1 ?

2. En fait, on ne connait pas la variance σ^2, mais on a observé la valeur $\tilde{S}^2_n = 2,7$ de la variance empirique corrigée $\tilde{S}^2_n = \frac{1}{n-1}\sum_{i=1}^{n}(Xi - \bar{X})^2$. Peut-on accepter, l'hypothèse H_0 ?

<u>Exercice 13</u>

Un échantillon de 10 000 personnes sur une population étant donné, on sait que le taux moyen de personnes à soigner pour un problème de cholestérol élevé est de 7,5%. Donner un intervalle dans lequel on soit « sûr » à 95%, de trouver le nombre exact de personnes à soigner sur les 10 000.

<u>Exercice 14 (pour devoir)</u>

Le staff médical d'une grande entreprise fait ses petites statistiques sur le taux de cholestérol de ses employés. Les observations sur 100 employés tirés au sort sont les suivantes.

Taux de cholestérol en cg:(centre classe)	120	160	200	240	280	320
Effectif d'employés :	9	22	25	21	16	7

1. Calculer la moyenne m_e et l'écart-type σ_e sur l'échantillon.

2. Estimer la moyenne et l'écart-type pour le taux de cholestérol dans toute l'entreprise.

3. Déterminer un intervalle de confiance pour la moyenne.

Exercice 15

Une compagnie aérienne a demandé des statistiques afin d'améliorer la sûreté au décollage et définir un poids limite de bagages. Pour l'estimation du poids des voyageurs et du poids des bagages, un échantillon est constitué de 300 passagers qui ont accepté d'être pesés : on a obtenu une moyenne me de 68kg, avec un écart-type σ_e de 7 kg.

1. Définir un intervalle de confiance pour la moyenne des passagers. (On admet que le poids des passagers suit une loi normale de moyenne m, d'écart-type σ)

2. Montrer que l'on peut considérer que le poids des passagers est une variable aléatoire X de moyenne 70kg, d'écart-type 8 kg.

3. En procédant de même pour le poids des bagages, on admet les résultats :

• Si le poids maximum autorisé est de 20 kg, le poids des bagages peut être considéré comme une variable aléatoire Y de moyenne 15 kg, d'écart-type 5 kg.

• La capacité de l'avion est de 300 passagers ; l'avion pèse, à vide, 250 tonnes. Le décollage est interdit si le poids total dépasse 276.2 tonnes. Quelle est la probabilité pour que le décollage soit interdit ?