

Sont admis les tables statistiques, le formulaire manuscrit (5 feuille A4 en RV) et la calculette personnels

NOM : Prénom :

Exercice 1 :

Dans une ferme, à une date déterminée, on a pesé les œufs qui ont été produits. La masse d'un œuf étant exprimée en grammes (g), le tableau statistique ci-dessous contient les **effectifs** des différentes **classes** observées sur l'échantillon prélevé.

Tableau 1

Masse de l'œuf (g)	28-37	38-47	48-52	53-57	58-62	63-72	73-82
Effectifs n_i	3	51	74	112	92	62	6
Effectifs cumulés N_i							
Fréquences f_i							
Fréquences cumulées F_i							
Amplitudes des classes a_i							
Centres des classes c_i							
Catégories	Petits		Moyens		Grands		

Statistique descriptive uni-dimensionnelle

1. Considérons que la variable aléatoire *masse d'un œuf* est une variable aléatoire quantitative **continue**. Compléter le tableau statistique avec les **effectifs cumulés** N_i , les **fréquences** f_i , les **fréquences cumulées** F_i , les **amplitudes** a_i et les **centres** c_i des classes.
2. Tracer l'**histogramme** et la **fonction de répartition** (courbe cumulative) à partir des données statistiques du tableau 1.
3. Calculer les **quartiles à 25 %**, à **50 %** et à **75 %**.
4. Caractériser la **tendance centrale** de la v.a. *masse d'un œuf* par le **mode** et par la **médiane**. Caractériser la **dispersion** de la v.a. *masse d'un œuf* par l'**intervalle inter-quartile**.
5. Calculer les valeurs de la **moyenne** et de la **variance empiriques** (sur les données statistiques). **Comparer** ces valeurs aux **indicateurs statistiques** précédents et conclure.
6. Que pensez-vous quant à la **distribution** des masses des œufs de cette ferme **autour de la valeur moyenne** ?

Échantillon et statistiques sur l'échantillon

Pour éviter de trier les œufs en fonction de leur masse en « **petits, moyens ou grands** », le fermier propose aux particuliers d'acheter des œufs par douzaine, au prix des œufs « **moyens** », en les prenant directement dans le conteneur des œufs non-triés.

Sont admis les tables statistiques, le formulaire manuscrit (5 feuille A4 en RV) et la calculette personnels

NOM : Prénom :

Exercice 1 :

Dans une ferme, à une date déterminée, on a pesé les œufs qui ont été produits. La masse d'un œuf étant exprimée en grammes (g), le tableau statistique ci-dessous contient les **effectifs** des différentes **classes** observées sur l'échantillon prélevé.

Tableau 1

Masse de l'œuf (g)	28-37	38-47	48-52	53-57	58-62	63-72	73-82
Effectifs n_i	3	51	74	112	92	62	6
Effectifs cumulés N_i							
Fréquences f_i							
Fréquences cumulées F_i							
Amplitudes des classes a_i							
Centres des classes c_i							
Catégories	Petits		Moyens		Grands		

Statistique descriptive uni-dimensionnelle

1. Considérons que la variable aléatoire *masse d'un œuf* est une variable aléatoire quantitative **continue**. Compléter le tableau statistique avec les **effectifs cumulés N_i** , les **fréquences f_i** , les **fréquences cumulées F_i** , les **amplitudes a_i** et les **centres c_i** des classes.
2. Tracer l'**histogramme** et la **fonction de répartition** (courbe cumulative) à partir des données statistiques du tableau 1.
3. Calculer les **quartiles à 25 %**, à **50 %** et à **75 %**.
4. Caractériser la **tendance centrale** de la v.a. *masse d'un œuf* par le **mode** et par la **médiane**. Caractériser la **dispersion** de la v.a. *masse d'un œuf* par l'**intervalle inter-quartile**.
5. Calculer les valeurs de la **moyenne** et de la **variance empiriques** (sur les données statistiques). **Comparer** ces valeurs aux **indicateurs statistiques** précédents et conclure.
6. Que pensez-vous quant à la **distribution** des masses des œufs de cette ferme **autour de la valeur moyenne** ?

Échantillon et statistiques sur l'échantillon

Pour éviter de trier les œufs en fonction de leur masse en « **petits, moyens ou grands** », le fermier propose aux particuliers d'acheter des œufs par douzaine, au prix des œufs « **moyens** », en les prenant directement dans le conteneur des œufs non-triés.

7. Supposons que la *masse d'un œuf* suit une loi de distribution de probabilité gaussienne, dont les paramètres, **espérance** et **variance**, inconnus doivent être estimés. Donner les **valeurs estimées de ces paramètres** sur l'échantillon i.i.d. donné dans le tableau 1.

8. Calculer la **probabilité** que la *masse moyenne d'un œuf* (dans une boîte de 12 œufs) soit **supérieure à 58 g** (la masse de l'œuf le moins lourd pouvant être considéré dans la catégorie « **Grands** »). Conclure.

Statistique descriptive bi-dimensionnelle

Dans la ferme on cherche à déterminer la liaison entre le **TYPE DE CÉRÉALES** consommés et la *catégorie des œufs* produits.

Tableau 2

CÉRÉALES / Catégorie	Petits	Moyens	Grands	Effectifs marginaux
CÉRÉALES TYPE 1	54	186	160	
CÉRÉALES TYPE 2	71	152	193	
Effectifs marginaux				

9. Vous trouverez dans le Tableau 2 *la table des contingence* avec les *effectifs conjoints*. Compléter la table de contingence avec les valeurs des **effectifs marginaux** et de **l'effectif total**. Calculer le **coefficient de liaison khi-deux**, puis le coefficient de liaison **phi-deux** et enfin celui de **Cramer**. Existe-t-il une **dépendance** entre la catégorie des œufs et le type de céréales ? Motiver votre réponse.

Statistique inférentielle

Soit X une v. a. continue de densité de probabilité *f* définie pour tout *x* réel par : $f(x) = pf_a(x) + (1-p)f_b(x)$ où *p* est un paramètre inconnu à déterminer et $f_a(x)$ et $f_b(x)$ les densités de deux lois uniformes respectivement sur $[0, a]$ et $[0, b]$, *a* et *b* étant deux paramètres connus tels que $a < b$. On dispose d'un échantillon i. i. d de la v. a. X. Sachant que la *fonction de vraisemblance* du paramètre *p* est égale à :

$$L(x_1, \dots, x_n; p) = \left(\frac{p}{a} + \frac{(1-p)}{b}\right)^y \left(\frac{(1-p)}{b}\right)^{(n-y)} \quad \text{où } y \text{ est le nombre d'observations } x_i \leq a$$

10. Déterminer l'expression de l'**espérance** de la v. a. X. En déduire l'**estimateur** \hat{p}_m de *p* par la **méthode des moments**. Sachant que $a=5$ et $b=13$, calculer l'estimation de *p* sur l'échantillon i.i.d. suivant :

1	2	2	5	13	4	3	8	11	5	6	3	1
---	---	---	---	----	---	---	---	----	---	---	---	---

11. Comparer, **sans la calculer**, la variance de l'**estimateur** \hat{p}_m avec celle d'un **estimateur max de vraisemblance** \hat{p}_{MV} .

12. Montrer que l'**estimateur** \hat{p}_m du paramètre *p* est **sans biais et convergent**. Est-ce qu'il est **exhaustif** ?

NOTA : Le médian est noté sur 20 points.