

Examen : Session Normale

Nom et Prénom* :

REMARQUES IMPORTANTES

- Les téléphones portables doivent être éteints.
 - Aucun document n'est autorisé.
 - Seules les calculatrices non programmables sont autorisées.
 - Les exercices sont indépendants. Ils ne sont pas classés par ordre de difficulté.
-

Questions de cours: (7pts)

Cocher la (les) bonne(s) réponse(s):

1. Un recensement est une enquête incomplète, c'est-à-dire une enquête au cours de laquelle quelques unités de base de la population sont observées.
 Oui Non
2. L'échantillonnage est l'ensemble des opérations qui permettent de sélectionner de façon organisée les éléments de l'échantillon.
 Oui Non
3. Un échantillon aléatoire est une suite (X_1, \dots, X_n) de variables aléatoires indépendantes et de même loi qu'un caractère de la population.
 Oui Non
4. Taux de sondage est le rapport entre la taille de l'échantillon et la taille de la population.
 Oui Non
5. Soit T un estimateur d'un paramètre $\theta > 0$ inconnu, tel que $\mathbb{E}(T) = \frac{1}{\theta}$.
Cet estimateur est-il sans biais ?
 Oui Non
6. Soient T_1 et T_2 deux estimateurs d'un paramètre θ inconnu, tels que

$$\mathbb{E}(T_1) = \theta \quad , \quad \mathbb{E}(T_2) = \frac{2n\theta}{2n+1},$$

où n désigne la taille de l'échantillon.

- T_1 est sans biais. T_2 est sans biais.
 T_2 est asymptotiquement sans biais.

*Feuille à rendre avec la copie.

7. Soient T_1 et T_2 deux estimateurs sans biais d'un paramètre θ , tels que $\text{Var}(T_2) = 2\text{Var}(T_1)$.

T_2 est plus efficace que T_1 . T_1 est plus efficace que T_2 .

8. Soient \tilde{S}_X^2 la variance empirique définie par

$$\tilde{S}_X^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n},$$

où n et \bar{X} désignent respectivement la taille et la moyenne de l'échantillon.

\tilde{S}_X^2 est un estimateur la moyenne de la population. \tilde{S}_X^2 est un estimateur sans biais.
 \tilde{S}_X^2 est un estimateur asymptotiquement sans biais.

9. Pour estimer une proportion inconnue p d'une population, on utilise

La moyenne de l'échantillon \bar{X} . La proportion de l'échantillon.
 La médiane de l'échantillon.

Exercice 1: (4pts)

Le tableau suivant présente la distribution des prix (en US \$) du pétrole brut dit BRENT de 01-2020 à 11-2021.

Mois	Prix p_t en US \$	Prévision par M.M.3	Mois	Prix p_t en US \$	Prévision par M.M.3
01-2020	63.6	--	01-2021	54.8	
02-2020	55.7	--	02-2021	62.3	
03-2020	31.9	--	03-2021	65.4	
04-2020	18.5		04-2021	64.8	
05-2020	29.4		05-2021	68.5	
06-2020	40.3		06-2021	73.2	
07-2020	43.2		07-2021	75.1	
08-2020	44.8		08-2021	70.9	
09-2020	40.9		09-2021	74.5	
10-2020	40.2		10-2021	83.5	
11-2020	42.7		11-2021	81.0	
12-2020	50.0		12-2021	--	

Source: *INSEE*. Notons que "M.M.3" désigne les moyennes mobiles d'ordre 3.

- Remplir le tableau ci-dessus.
- On partage la série p_t $t = 1, \dots, 24$ en deux groupes

$E_1 := \{ \text{Les prix du BRENT dans les mois de l'année 2020 (en US \$)} \}$

$E_2 := \{ \text{Les prix du BRENT dans les mois de l'année 2021 (en US \$)} \}$

Déterminer les points moyens G_1 et G_2 des groupes E_1 et E_2 .

- En déduire l'équation de la droite d'ajustement de Mayer.
- Quels prix peut-on prévoir pour le mois 12 de l'année 2021 par les deux méthodes suivantes:
 - Méthode des moyennes mobiles d'ordre 3.
 - Méthode de Mayer.
- Proposer un critère pour comparer la performance des deux méthodes en question.

Exercice 2: (6pts)

Le tableau suivant donne la moyenne Y des maxima de tension artérielle en fonction de l'âge X d'une population donnée :

Âge : X	36	42	48	54	60	66
Tension : Y	12	13.5	12.5	14.25	15.5	15

1. Préciser la nature des variables étudiées.
2. a) Représenter graphiquement la série d'observations (x_i, y_i) de la variable couple (X, Y) , en précisant la droite D des points extrêmes.
b) En déduire une prévision y_7 de tension artérielle pour une personne âgée de 78 ans.
3. Calculer les moyennes \bar{x} et \bar{y} des séries d'observations x_i, y_i en question.
4. Calculer le coefficient de corrélation linéaire entre X et Y , et interpréter le résultat.
5. a) Déterminer l'équation de la droite Δ de régression linéaire de Y en fonction de X .
b) En déduire une nouvelle prévision y'_7 de tension artérielle pour une personne de 78 ans.
6. Représenter sur le même graphique la droite Δ de régression linéaire de Y en fonction de X .
7. Lequel des deux résultats y_7 et y'_7 paraît le plus vraisemblable ? (Justifier votre réponse).

Exercice 3: (3pts)

Un sondage effectué sur 300 votants d'une population de 30 000 personnes a montré que 165 personnes avaient l'intention de voter pour l'acceptation du projet soumis au vote. La proportion p (dans la population) des personnes qui ont l'intention de voter pour l'acceptation du projet est inconnue.

1. Proposer un estimateur adéquat, et donner une estimation ponctuelle de cette proportion.
2. Déterminer l'intervalle de confiance de la proportion p , au niveau de confiance de 95%.

Hint: Le seuil $\alpha = 0.05$ et $z_{1-\alpha/2} = 1.96$.

3. Préciser la taille n de l'échantillon pour assurer un intervalle de confiance de p , au niveau de confiance de 95% avec une marge d'erreur inférieure à 10^{-3} .