

Contrôle 1 de statistique

Ce contrôle est un questionnaire à choix multiples : certaines questions seulement comportent plus d'une réponse exacte.

1),1) Dans le cas d'une variable statistique continue, l'hypothèse d'uniformité dans chaque classe permet de :

- a) calculer les caractéristiques de position centrale et de dispersion de la variable
 b) remplacer chaque classe par son centre c) tracer le diagramme intégral
 d) calculer les effectifs et les fréquences cumulés e) tracer l'histogramme

2) On a réalisé l'étude d'un paramètre quantitatif et on a obtenu la distribution non groupée suivante :

x_i	n_i	f_i
100	1	0,10
150	3	0,3
...

Quelles sont les propositions exactes ?

- a) on ne peut pas connaître l'effectif total de l'échantillon. b) l'effectif total de l'échantillon est de 4 individus.
 c) l'effectif total de l'échantillon est de 10 individus. d) il y a 100 individus qui ont la valeur 1.
 e) dans l'échantillon, aucun individu n'a eu des valeurs inférieures à 100.
 3) Le coefficient de corrélation linéaire sert à
 a) étudier l'existence d'une relation linéaire entre les variables b) calculer la covariance c) montrer qu'il n'y a pas de relation entre les variables d) déterminer la droite d'ajustement e) créer une relation linéaire entre les variables

4) Quand le coefficient de corrélation est proche de 1 nous comprenons que

- a) la covariance est proche de 1 b) les points sont proches de la droite c) la relation entre les 2 variables est non linéaire d) la pente de la droite est négative e) les points se trouvent exactement sur la droite

5) indiquer les affirmations exactes :

a) l'interpolation linéaire est une technique pour calculer la médiane dans le cas continu

b) l'écart type est toujours inférieur à l'étendue

c) on n'ajuste une droite à un nuage de points que lorsque le coefficient de corrélation linéaire est proche de 1, en valeur absolue

d) la variance donne une moyenne de la distance des observations par rapport à n'importe quelle caractéristique de position centrale

e) la droite d'ajustement d'une variable Y par rapport à une variable X est la meilleure courbe au sens où elle est celle qui est la plus proche des points

6) indiquer les affirmations incorrectes

a) une variance est toujours positive b) la variance est positive si la variable est continue

c) l'histogramme est le diagramme différentiel d'une variable discrète

d) l'histogramme est le diagramme différentiel d'une variable continue

e) l'histogramme est le diagramme intégral d'une variable continue

7) Le tableau suivant représente la distribution des accidents par jour dans une ville

Accidents	0	1	2	3
Fréquences relatives	0.53	0.20	0.1	0.15

- Indiquer quelles sont les affirmations correctes

- a) le mode et la médiane sont égaux b) l'effectif total est 100 c) la moyenne et la médiane sont égales d) 15 est l'effectif de la modalité 3 e) le troisième quartile est calculé par interpolation linéaire dans l'intervalle [1,2]

8) Nous avons vu en cours qu'une liaison linéaire entre 2 variables X et Y peut être formulée sous la forme de l'équation d'une droite $Y = aX + b$, seulement dans beaucoup de cas le nuage de points a une forme de parabole qu'on peut exprimer par l'équation

$$Y = a_0 + a_1X + a_2X^2$$

Il s'agit alors, de déterminer les coefficients a_0 , a_1 et a_2 pour établir définitivement la relation entre X et Y.

- chercher la parabole des moindres carrés ajustant le nuage de points $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$ revient à chercher les coefficients a_0 , a_1 et a_2 qui minimisent la quantité β tel que :

a) $\beta = \frac{1}{N} \sum_{i=1}^N (Y_i - a_1X_i - a_0)^2$ b) $\beta = \frac{1}{N} \sum_{i=1}^N (Y_i - a_0 - a_1X_i - a_2X_i^2)^2$

c) $\beta = \frac{1}{N} \sum_{i=1}^N (Y_i - a_0 - a_1X_i - a_2X_i^2)^2$ d) $\beta = \sum_{i=1}^N (Y_i - a_0X_i)^2$

e) toutes les réponses précédentes sont fausses.

9) Si X et Y sont des variables positives et $a_0 = b^2$, $a_1 = 2ab$ et $a_2 = a^2$, indiquer les réponses correctes dans ce qui suit :

a) $\begin{cases} a = \frac{\text{Cov}(X, W)}{\text{Var}(X)} \text{ où } W = \sqrt{Y} \\ b = \bar{W} - a\bar{X} \end{cases}$ b) $\begin{cases} a = \frac{\text{Cov}(X, Z)}{\text{Var}(X)} \text{ où } Z = \ln Y \\ b = \bar{Z} - a\bar{X} \end{cases}$

c) $\begin{cases} a = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \\ b = \bar{Y} - a\bar{X} \end{cases}$ d) $\begin{cases} a = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} \\ b = \bar{Y} - a\bar{X}^2 \end{cases}$ e) $a = b = 0$

10) Afin d'évaluer les réactions allergisantes d'un produit, on l'injecte à des rats à trois périodes suffisamment espacées pour que l'effet de la précédente injection soit annulé. En supposant que les chances d'avoir une réaction allergisante soit les mêmes pour chaque animal, chaque rat recevant trois injections peut développer 0, 1, 2, ou 3 réactions allergisantes.

Nombre de réactions allergisantes	0	1	2	3
Nombre de rats	95	27	11	2

- Quelles sont les propositions exactes ?

a) le nombre total d'injections est 405.

b) la population dans ce cas est : les rats.

- c) le nombre total de réactions allergisantes est 55. d) le nombre total de réactions non-allergisantes est 300.
 e) la proportion de rats ayant présenté une réaction allergisante est 23,5%.

11) Nous avons vu dans le cours que la quantité

$$\frac{1}{N} \sum_{i=1}^n [R(X_i - \bar{X}) - (Y_i - \bar{Y})]^2$$

nous a permis de conclure que la covariance était toujours, en valeur absolue, inférieure ou égale au produit des écarts types. Nous avons aussi vu que (M) était équivalente à un trinôme du second degré $P(R) = mR^2 - nR + u$.

Indiquer parmi les propositions suivantes celles qui sont correctes

a) R est un paramètre statistique bien connu. b) R est un paramètre arbitraire (ou quelconque)

c) le déterminant Δ de P(R) peut être strictement supérieur à 0

d) le déterminant Δ de P(R) est, pour tout R, inférieur ou égal à 0

e) $m = \text{Var}(X)$, $n = 2\text{Cov}(X, Y)$ et $u = \text{Var}(Y)$

12) Lorsque n, m et u de la question 11) sont tel que $|n|/2 = (m \cdot u)^{1/2}$, choisir les bonnes réponses dans ce qui suit

a) Le coefficient de corrélation ρ , en valeur absolue, vaut 1

b) Le coefficient de corrélation ρ , en valeur absolue, est proche de 1

c) Tous les points du nuage sont sur une même droite

d) certains points du nuage sont sur une même droite

e) Aucune des réponses précédente n'est vraie

13) Dans quel cas pensez vous que c'est la variable Y qui doit être ajustée par la variable X

a) X : jours - Y : nombre de nouveaux cas de malades par jour dans une épidémie

b) X : le nombre de boîtes vendues par mois - Y : mois

c) X : âge - Y : Temps de sommeil

d) X : nombre de bactéries dans un milieu approprié après leur culture sous différentes températures - Y : Température

e) aucune des variables Y dans a), b), c) et d) ne peut être ajustée par X

14) Indiquer les fonctions pouvant être ramenées au cas linéaire (comme ça a été le cas pour les fonctions exponentielles et puissance dans le cours de l'ajustement linéaire)

a) $Y = a + bx^2 - cx^4$ b) $Y = AX^B + 1$ c) $Y = b + a/X$ d) $Y = (aX + b)^2$, Y et X étant des variables positives e) $Y = aX + bX^{1/2}$

11) Soit la répartition des salaires journaliers des employés d'une usine « A » :

Salaires (DA)	[500 ; 600[[600 ; 700[[700 ; x[[x ; 900[[900 ; 1000[
Nombre d'employés	120	190	240	100	50

Un employé a un salaire de 720 DA et il y a 342 employés qui ont un salaire plus bas que lui.

15) Indiquer les bonnes réponses dans ce qui suit

a) la population dans ce cas est « les salaires » des employés

b) la population dans ce cas est « les employés » de l'usine A

c) le caractère dans ce cas est « les salaires » des employés

d) le caractère dans ce cas est « les employés » de l'usine

e) le caractère étudié est un caractère quantitatif discret

16) la valeur de x est

a) 750 b) 850 c) 800 d) 740 e) 780

17) La médiane est égale à

a) 725 b) 750 c) 780 d) 800 e) 350

18) Le nombre d'employés qui perçoivent un salaire compris entre 600 et 800 est

a) 430 b) 350 c) 400 d) 450 e) 550

19) Le nombre d'employés ayant un salaire supérieur ou égal à 600 DA est

a) 120 b) 190 c) 580 d) 290 e) 700

20) Les représentations graphiques pouvant représenter graphiquement ces données sont

a) histogramme b) le diagramme en bâtons c) le diagramme intégral

d) le diagramme en escaliers e) le nuage de points

2 est calculé par interpolation linéaire :
~~342~~
 $N(720) = 342$