

EISTI ING1 : Examen de Statistiques Descriptives

documents autorisés, durée 1/2heure

21 janvier 2010

Question 1) On considère un caractère quantitatif. Si on trouve une médiane beaucoup plus grande que la moyenne. Que peut-on en déduire ?

Questions 2) On considère un caractère quantitatif. On étudie ce caractère sur plusieurs populations. Notre étude se borne à l'étude de la moyenne de ce caractère. Dans notre cas on va donc comparer les moyennes de ce caractère sur les différentes populations. On envisage que lors de la saisie des données il y ait quelques valeurs mal saisies et donc que certaines valeurs soient anormalement trop petites ou trop grandes. Ces valeurs risquent donc de fausser les interprétations. Que préconisez-vous pour remédier à cela ? justifier (vous devez utiliser des outils vus en cours).

Questions 3) On s'intéresse aux salaires dans les grandes entreprises. Dans deux grandes entreprises, on calcule les salaires moyens et les salaires médians. Les deux salaires moyens sont égaux et les deux salaires médians sont égaux. Pourquoi n'est-ce pas suffisant pour affirmer que ces deux entreprises ont la même structure salariale ? Quels indicateurs supplémentaires proposez-vous ? (justifier).

Question 4) On étudie un caractère numérique sur deux populations. Dans la première population, on connaît pour chaque individu la valeur exacte du caractère. Dans la deuxième population les valeurs possibles du caractère avait été préalablement découpées en tranche. Donc pour chaque individu, on connaît la tranche dans laquelle se situe l'individu. Entre les deux populations, les deux caractères ne sont donc pas représentés de la même façon. On ne peut donc pas directement faire des études comparatives. Que proposez-vous comme étape préalable pour résoudre ce problème ?

Question 5) En analyse multivariée, les caractères ne sont pas de même nature et donc ne peuvent pas être exprimés dans la même unité (un poids s'exprime en kg et une taille en cm). Pourquoi, en divisant la réponse d'un individu à un caractère par l'écart-type de ce caractère on résout le problème.

Question 6) Quelle est la principale valeur ajoutée de l'analyse bivariée par rapport à l'analyse univariée ?

Question 7) Quand on croise deux caractères quantitatifs, on est amené à calculer le coefficient de corrélation linéaire r . Pourquoi quand $r = 0$ (ou proche de 0) ne peut-on pas conclure que les deux caractères sont indépendants ?

Question 8) On fait un sondage sur la population française à la veille d'une élection présidentielle. Pour chaque personne interrogée, on connaît la tranche d'âge à laquelle il appartient et le candidat pour lequel il désire voter. On veut savoir si le choix du candidat diffère selon la tranche d'âge. Quelle méthode proposez-vous d'utiliser ? (on justifie).

Question 9) Pour faire une analyse bivariée, pourquoi est-il utile de faire auparavant une analyse univariée sur chacun des deux caractères ?

Question 10) On considère une population d'étudiants. On les évalue sur un thème donné. Ils obtiennent chacun une note sur 20. Après cette évaluation, ces étudiants suivent un cours. Après ce cours on évalue une deuxième fois les étudiants sur le même thème. Quelle méthode proposez-vous pour vérifier l'impact du cours sur le changement de niveau des étudiants sur le thème donné ?