

Dispositif INRA de formation permanente à la statistique.
Niveau 2, module 1 :

Introduction à la décision statistique

Questionnaire de positionnement des connaissances

Les *connaissances mathématiques requises* pour suivre ce module (situé à l'« entrée » du niveau 2 du dispositif) ont été volontairement réduites au strict minimum (notations des variables avec indices et des sommes, calculs de probabilités les plus élémentaires), à peu près incontournable pour la compréhension et donc l'utilisation correcte des résultats statistiques.

Ce questionnaire devrait permettre, d'une part à *chacun des stagiaires* de vérifier si son niveau mathématique est suffisant pour suivre cette formation avec profit, et d'autre part *au formateur* d'évaluer l'ensemble de ces niveaux pour ajuster sa formation au groupe, former (éventuellement) plusieurs groupes plus homogènes, et le cas échéant conseiller à certains une préparation préliminaire susceptible de l'amener au niveau requis (lectures personnelles, stages de niveau plus élémentaire).

Il a donc pour unique objectif d'assurer à la formation un maximum d'efficacité, et en aucun cas d'élaborer un quelconque « classement définitif » des stagiaires en « bons » et « mauvais » (qui aurait d'autant moins de sens que l'objectif de la formation ne peut être que « tous bons »)... C'est pourquoi l'on remet aux stagiaires le « corrigé » *en même temps que le questionnaire*, en espérant qu'ils sauront comprendre que le test perd tout intérêt (pour le formateur comme pour eux-mêmes) s'ils ont la faiblesse d'utiliser ce corrigé pour leurs réponses !

Il ne sera pas délivré de « notes » à proprement parler : le formateur fera, à la suite du test, un commentaire du corrigé au cours d'une discussion avec l'ensemble des stagiaires qui s'y sont soumis, et restera disponible pour tout entretien individuel jugé nécessaire ; à l'issue de la réunion, il rassemblera les « copies » (il récupérera aussi les corrigés, pour qu'ils ne tombent pas accidentellement sous les yeux de collègues susceptibles de se soumettre au test une autre fois...) et les restituera dans un délai d'une semaine.

Nom, prénom :	Station :	Téléphone :
		Mél :

Veillez répondre (sans consulter le corrigé...) aux questions **dans l'ordre** (car l'énoncé de certaines questions apporte des indications sur les réponses à des questions précédentes...) et noter vos réponses **à droite** (ou *à la suite*, dans l'espace laissé entre les énoncés des questions, si une *justification* ou un *commentaire* vous paraît nécessaire); pour les choix multiples, le point d'interrogation « ? » correspond à la réponse « je ne sais pas ». Enfin, nous ne voyons aucun inconvénient (bien au contraire) à ce que vous utilisiez les marges pour les brouillons de vos calculs (sauf si vous estimez la place insuffisante).

Première partie : indices et sommes

On note, dans l'ordre, X_1, X_2, \dots, X_8 les huit valeurs de la liste suivante :

10,1 20,4 28,7 39,2 320,8 331,3 339,6 349,9

1. Calculer les valeurs des différences : $X_5 - X_1 =$
 et $X_4 - X_8 =$

2. Les deux différences de la 1^{re} question ont la même valeur absolue : est-ce une coïncidence? oui / non / ?
 Autre formulation (pour une réponse inverse) : serait-ce encore forcément vrai avec 8 autres valeurs quelconques? (non / oui / ?)

3. Comment *noter* (en utilisant seulement la lettre X et les indices de 1 à 8, car il ne s'agit *pas* ici de *calculer*) la différence entre d'une part la 6^e valeur de la liste, et d'autre part la 4^e *en partant de la fin de la liste*?

4. Calculer les deux sommes : $\sum_{i=1}^4 X_i =$
et $\sum_{i=5}^8 X_i =$
et leur différence :

5. Comment *noter* (en utilisant seulement la lettre X et les indices de 1 à 8, il ne s'agit *pas* encore de la *calculer*) la moyenne (arithmétique) des 8 valeurs de la liste?

6. Que vaut la moyenne des 8 valeurs de la liste?

7. Calculer la somme: $\sum_{i=1}^4 (X_{9-i} - X_i) =$

8. Pour calculer la moyenne à la 6^e question, avez-vous utilisé les valeurs des sommes calculées à la 4^e question? oui / non

9. La différence de la 4^e question, $\sum_{i=5}^8 X_i - \sum_{i=1}^4 X_i$, et la somme de la 7^e question, $\sum_{i=1}^4 (X_{9-i} - X_i)$, ont la même valeur : est-ce une coïncidence? oui / non / ?

10. Pour calculer la somme à la 7^e question, avez-vous utilisé les valeurs des sommes calculées à la 4^e question? oui / non

Seconde partie : pourcentages et probabilités

Tous les dés que nous allons utiliser sont « ordinaires », à six faces carrées régulières numérotées de 1 à 6 sans omission ; ils ne sont « truqués » d'aucune manière, de même que les pièces de monnaie à deux côtés (*pile* et *face*) que nous jetterons avec prodigalité...

11. Dans les réserves du monastère augustin de Brno, on a trouvé 1 200 graines de pois ridées et 6 800 graines lisses. Quel est le pourcentage de graines ridées?

12. 80 % des graines ridées de la question précédente sont à cotylédon jaune : quel est leur nombre?

13. Dans le Grand-Duché de G erolstein, une personne sur 200 est militaire, et 8 personnes sur 10 000 sont officiers. Quel est le pourcentage d'officiers parmi les militaires? .. .

14. On lance un d e. Quelle est la probabilit e d'obtenir la face marqu ee 5? .. .

15. On lance le m eme d e deux fois. Quelle est la probabilit e d'obtenir aux deux fois la face marqu ee 5? .. .

16.  a chaque emplacement vide de la phrase suivante, on peut ins erer un mot choisi parmi une liste de deux ou trois possibilit es (les listes suivent la phrase) :

« Si deux  ev enements sont [...(A)...], la probabilit e pour que [...(B)...] se produise(nt) s'obtient en faisant [...(C)...] de leurs probabilit es respectives. »

A1 : incompatibles .. . A2 : ind ependants.

B1 : tous les deux .. . B2 : au moins l'un des deux.

C1 : la somme .. . C2 : la diff erence .. . C3 : le produit.

Combien de phrases (m eme fausses) peut-on construire ainsi? .. .

17. Parmi les phrases construites  a la question pr ec edente, laquelle (ou lesquelles) exprime(nt) une « v erit e » math ematique? (coder la (les) r eponse(s) sous la forme: $A_i.B_j.C_k$, en rempla cant chacune des lettres i , j et k par 1 ou 2 ou 3). .. .

18. Un matin, la poubelle  a l'arr et du bus contient 9 tickets verts et 1 ticket jaune. Le soir, elle contient 72 verts et 8 jaunes. Les agents de surface charg es du ramassage biquotidien jouent  a parier sur la couleur d'un ticket tir e au hasard, sans tricher, dans le fond de la poubelle suppos ee opaque, le ticket jaune  etant gagnant.

 a leur place, choisiriez-vous de jouer : (A) plut ot le matin, ou (B) plut ot le soir, ou (C) indiff eremment le matin ou le soir? (*N. B.* : un refus de jouer  a ce jeu – certes d ego tant – ne pourrait que vous faire perdre la face dans l' equipe) .. A / B / C / ?

19. Vous jetez une pièce de monnaie soit (*i*) 2 fois, soit (*ii*) 100 fois, et vous gagnez (par exemple l'estime du tricheur professionnel...) si vous obtenez exactement autant de « pile » que de « face ».

Préférez-vous : (A) la partie de type (*i*), ou (B) la partie de type (*ii*), ou (C) pensez-vous que les chances de gagner soient les mêmes?

A / B / C / ?

20. Vous pouvez jeter une pièce de monnaie dans deux types de parties :

(*i*) vous effectuez 100 jets, et vous gagnez si vous obtenez au moins 60 « face »;

(*ii*) vous effectuez 1 000 jets, et vous gagnez (la même chose...) si vous obtenez au moins 600 « pile ».

Préférez-vous : (A) la partie de type (*i*), ou (B) la partie de type (*ii*), ou (C) pensez-vous que les chances de gagner soient les mêmes?

A / B / C / ?

Corrigé

Première partie : indices et sommes

1. $X_5 - X_1 = 310,7$ et $X_4 - X_8 = -310,7$.
2. Réponse: *oui* à la première formulation, *non* à la seconde.
3. $X_6 - X_5$, ou $X_5 - X_6$.
4. $\sum_{i=1}^4 X_i = 98,4$, $\sum_{i=5}^8 X_i = 1\,341,6$ et leur différence vaut $1\,243,2$.
5. $(\sum_{i=1}^8 x_i) / 8$.
6. Réponse 1 (de très loin la plus attendue) : $1440 / 8 = 180$. D'autres réponses sont possibles : 0 (dans un cas assez particulier qu'il est à peu près impossible d'imaginer sans indications complémentaires...), et une « non-réponse » consciemment choisie, assumée et justifiée par une phrase du genre « il ne faut pas calculer sans précautions n'importe quoi (même une moyenne) sur n'importe quelles données » ; cette dernière réponse, un peu du domaine du « rêve », ne serait pas impossible mais plutôt inespérée (soyons réalistes)...
7. $\sum_{i=1}^4 (X_{9-i} - X_i) = 1\,243,2$.
9. Réponse: *non*, car on peut changer l'ordre des termes des sommes, et/ou associer/dissocier les termes des sommes sans changer le résultat.

Seconde partie : pourcentages et probabilités

11. 15 %.
12. 960.
13. 16 %.
14. $1/6$, ou 16,667 %.
15. $1/36$, ou 2,778 %.
16. 12.
17. A1.B2.C1 et A2.B1.C3.
18. C.
19. A.
20. A.

Éléments d'évaluation des réponses

Première partie : indices et sommes

1. Maîtrise de la notation indicielle : sauf accident isolé, il serait *gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question.

2. Question sans grand intérêt si ce n'est de tester le « bon sens », et de préparer à bien comprendre la formulation de la question 9.

3. Maîtrise de la notation indicielle : sauf accident isolé, il serait *gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question.

4. Maîtrise de la notation \sum : sauf accident isolé, il serait *gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question.

5. Maîtrise de la notation \sum : sauf accident isolé, il serait *un peu gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question un peu moins simple que 1, 3, 4, et 6, mais sans doute plus facile que 7.

6. Maîtrise de la notation \sum : sauf accident isolé, il serait *gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question.

Réponse 1 : il faut s'attendre à ce que ce soit la réponse d'une majorité écrasante... Mais on peut toujours rêver de trouver :

Réponse 2 : « Il ne faut pas calculer sans précautions n'importe quoi (même une moyenne) sur n'importe quelles données ». Justification : dans notre cas, on a manifestement 2 groupes de valeurs (les « petites » de 10 à 40 et les « grandes » de 320 à 350 environ) : avec la réponse 1, on court un risque certain de calculer ainsi une « moyenne de choux et de navets » de signification hasardeuse... erreur que n'évite bien sûr pas toujours même le statisticien même confirmé, pour peu qu'il soit un peu trop pressé, et/ou que les deux groupes soient mieux mélangés que dans cet énoncé !

Réponse 2 *bis* : la même que ci-dessus, en remplaçant carrément dans la phrase l'expression « sans précautions » par « n'importe comment ». Justification (peut-être un peu plus tirée par les cheveux) : dans certains cas, la « bonne » réponse *devrait* être 0 et non 180 (prendre quelques minutes pour essayer d'imaginer une telle situation, avant de lire la suite car après il sera trop tard)... Par exemple si les 8 valeurs étaient des mesures d'angles en degrés (direction du vent, écart entre axe du crâne et direction de visée des yeux, etc.), il faudrait faire les calculs *modulo* 360 et donc la moyenne

serait $1440/8 \bmod 360 = 0/8 = 0$ (si cette explication ne suffit pas, représenter les données par des points sur un cercle trigonométrique ou sur une « boussole »). C'est un piège très difficile à éviter si l'on n'est pas prévenu! (se méfier tout de même lorsque l'on doit calculer une moyenne de quantités dont on ignore tout, jusqu'à l'unité de mesure)...

7. Maîtrise de la notation \sum et de la notation indicielle : sauf accident isolé, il serait *peut-être gênant* pour la capacité à suivre la formation avec profit de ne pas savoir répondre à cette question (un peu moins simple que 1, 3, 4, et 6, et 5).

8. Une réponse négative pourrait indiquer une *maîtrise opératoire de la notation \sum un peu légère*, pas forcément réhibitoire pour suivre la formation, mais peut-être pas dans un groupe de niveau initial « excellent » ; voir aussi questions 9 et 10.

9. Une réponse insuffisante (« ? » ou « oui, c'est une coïncidence ») pourrait indiquer une *maîtrise opératoire de la notation \sum un peu légère*, pas forcément réhibitoire pour suivre la formation, mais peut-être pas dans un groupe de niveau initial « excellent » ; voir aussi questions 8 et 10.

10. Une réponse positive pourrait indiquer une *très bonne maîtrise opératoire de la notation \sum* , permettant de participer à un groupe de bon niveau initial ; voir aussi questions 8 et 9.

Seconde partie : pourcentages et probabilités

11, 12 et 13 (question peut-être un peu moins simple : on ne fait pas le rapport de quantités, « concrètes », mais de proportions). Définition des pourcentages. Des réponses insuffisantes seraient gênantes.

14. Probabilités très élémentaires. Réponse insuffisante : gênante.

15. Probabilités élémentaires. Réponse insuffisante : assez gênante.

16. Dénombrement élémentaire. Réponse insuffisante : gênante.

17. Probabilités élémentaires. Réponse insuffisante : assez gênante.

18. Probabilités assez élémentaires. Réponse insuffisante : peut-être gênante.

19. Probabilités un peu moins élémentaires. Une réponse correcte serait assez encourageante.

20. Une réponse correcte pourrait montrer une bonne connaissance des probabilités élémentaires (loi binomiale et son approximation gaussienne), pour un candidat à placer dans un groupe de bons.