Problèmes supplémentaires 2

STT1700

Automne 2008

1. Un célèbre magicien qui prétendait avoir des pouvoirs de perception extrasensorielle a accepté de se livrer à une expérience dans laquelle il se proposait de deviner le résultat du lancer d'un dé. En 12 essais, il a réussi à deviner le résultat 10 fois. Vérifier que la probabilité d'un nombre de succès aussi grand que 10 (c'est-à-dire supérieur ou égal à 10) est excessivement petite pour quelqu'un qui répond au hasard; et expliquer à quelle conclusion ce fait a tendance à mener.

2. Un certain test psychologique consiste à lire un paragraphe, et puis répondre à 20 questions portant sur le texte lu. Un choix de 5 réponses est donné pour chaque question. Un évaluateur, tentant de démontrer que le test ne mesure pas l'aptitude à la lecture, répond aux 20 questions sans avoir lu le texte. Il choisit la bonne réponse à 8 des questions. Calculer la probabilité d'avoir 8 succès ou plus, et discuter les implications sur la qualité du test.

3. Sachant que X suit une loi $\mathfrak{N}(100 ; 256)$, calculer chacune des probabilités suivantes.

   a) $P(100 \leq X \leq 150)$ *b*) $P(X \leq 92)$ *c*) $P(|X\text{-}100| \leq 8)$
   d) $P(52 < X \leq 148)$ *e*) $P(76 \leq X \leq 92)$ *f*) $P(|X\text{-}100| > 8)$

4. Soit $X \sim \mathfrak{N}(32 ; 81)$. Déterminer la valeur de a telle que

   a) $P[X > a] = 0{,}1056$ *b*) $P[32 < X < a] = 0{,}3554$ *c*) $P[X < a] = 0{,}7734$
   d) $P[X < a] = 0{,}0188$ *e*) $P[|X\text{-}32| < a] = 0{,}8836$ *f*) $P[|X\text{-}32| > a] = 0{,}3422$

5. Supposons que la taille des hommes d'une population soit distribuée normalement de moyenne $\mu = 70$ pouces et d'écart-type $\sigma = 10$ pouces Quelle devrait être la longueur d'un matelas pour qu'il puisse accommoder 99 % des hommes de la population?

6. Lorsqu'une machine est réglée pour mettre μ grammes de petits pois dans des boîtes de conserve, elle n'en met pas exactement μ grammes. Le poids réel du contenu varie selon une loi normale de moyenne μ et d'écart-type σ grammes. Si $\sigma = 4$, à quelle valeur doit-on régler μ pour que 1% seulement des boîtes contiennent moins de 300 grammes?

7. Soit X_1 et X_2 deux variables aléatoires indépendantes, $X_1 \sim \mathfrak{N}(25 ; 16)$ et $X_2 \sim \mathfrak{N}(23 ; 9)$. Calculer

   a) $P\{41 \leq X_1 + X_2 \leq 51{,}5\}$ *b*) $P\{-1{,}5 \leq X_1 - X_2 \leq 16\}$, *c*) $P\{X_2 - X_1 > 0\}$.

8. Supposons que le Q. I. moyen de 10 personnes choisies au hasard d'une population normale d'écart-type $\sigma = 10$ est $\overline{X} = 110$. Calculer $P\{\overline{X} \geq 110\}$ et dites si la moyenne observée de 110 est trop grande pour qu'on puisse accepter l'hypothèse que la moyenne de la population vaut 100.

9. Un terrain est découpé en 10 lots identiques. Sans engrais, la production de céréales, en tonnes, pour chaque lot, suit une loi $\mathfrak{N}(6 ; 1)$. En utilisant un certain engrais, la production d'un lot sera de loi $\mathfrak{N}(6{,}3 ; 1)$. Parmi les 10 lots, 6 sont semés sans engrais et 4 reçoivent de l'engrais. Quelle est la probabilité que les 6 lots sans engrais produisent, en moyenne, plus de céréales que les 4 lots avec engrais?

10. Soit $\overline{X}_1$ le Q.I. moyen de 10 personnes choisies au hasard d'une population normale de variance 100; soit $\overline{X}_2$ le Q.I. moyen de 15 personnes choisies au hasard d'une population de variance 121.

    a) Supposons que les deux populations sont de même moyenne. Quelle est la distribution de $\overline{X}_1 - \overline{X}_2$?

    b) Supposons qu'on observe $\overline{X}_1 - \overline{X}_2 = 20$. Cet écart est-il trop grand pour qu'on trouve acceptable l'hypothèse que les deux populations sont de même moyenne?

11. On suppose que le poids (en *kg*) des adultes se distribue avec une moyenne de 64 et un écart-type de 12. Dans un ascenseur, une plaque indique « Capacité maximale: 12 personnes ou 1000 *kg* ». Si 14 personnes s'entassent dans l'ascenseur, quelle est la probabilité que leur poids total dépasse 1000 *kg*?

12. D'une population normale de moyenne $\mu = 200\,\$$ et d'écart-type $\sigma = 30\,\$$, un statisticien prélève un échantillon de $n = 25$ unités afin d'estimer la moyenne (μ) de la population par la moyenne échantillonnale $\overline{X}$.

    a) Quelle est la probabilité qu'il surestime de plus de 4 \$ la vraie moyenne?

    b) Quelle est la probabilité qu'il commette une erreur (de sous-estimation ou de surestimation) de plus de 10\$ dans son estimation?

    c) Quelle est la probabilité qu'il commette une erreur (de sous-estimation ou de surestimation) de plus de 1% par rapport à la vraie moyenne?

    d) Quelle devrait être la taille de son échantillon pour que la probabilité d'une erreur de plus de 1% par rapport à la vraie moyenne ne soit pas supérieure 0,05?

13. On prélève un échantillon de taille $n = 30$ d'une population d'écart-type $\sigma = 18\$$ pour estimer la moyenne de la population.

    a) Quelle est la probabilité de surestimer la vraie moyenne de plus de 6\$?

    b) Quelle est la probabilité de commettre une erreur (de sous-estimation ou de surestimation) de plus de 6\$?

    c) Quelle doit être la taille de l'échantillon s'il faut que la probabilité d'une erreur de plus de 6\$ ne dépasse pas 0,05?

14. Un acheteur de pièces de manufacture et son fournisseur sont d'accord pour déclarer «acceptable» un lot dont la proportion réelle de pièces défectueuses est de 5% ou moins. Les deux parties s'entendent pour que, à chaque livraison, on effectue une inspection selon le plan suivant: l'acheteur prélève 120 pièces au hasard et rejette le lot si 12 pièces ou plus dans l'échantillon sont défectueuses.

    a) Avec le plan d'inspection proposé, quelle est la probabilité que le lot soit rejeté alors qu'il est juste acceptable (c'est-à-dire, si le pourcentage réel de pièces défectueuses est de 5 %)?

    b) Avec le plan d'inspection proposé, quelle est la probabilité que le lot soit rejeté alors qu'il est mieux qu'acceptable, c'est-à-dire, si le pourcentage réel de pièces défectueuses est 4 %?

    c) Supposons qu'on trouve 14 pièces défectueuses parmi les 120 inspectées, et que le fournisseur insiste quand même que le lot est acceptable, attribuant au hasard le fait que l'échantillon contient un grand nombre de pièces défectueuses. Expliquer clairement pourquoi on n'accepte pas cette argumentation.

    d) Selon le plan proposé, on rejette le lot si le nombre de pièces défectueuses dans l'échantillon est supérieur ou égal à 12. Modifier ce critère de façon que la probabilité de rejeter le lot alors qu'il est juste acceptable (proportion est 5 %) soit inférieure à 0,5 %.

15. Pour contrôler la qualité de certains sachets pharmaceutiques, on décide de tirer de chaque lot un échantillon de 16 sachets et prendre note des poids (en *mg*) $X_1, X_2, \ldots, X_{16}$. Supposons que $X_i : \mathfrak{N}(\mu\,;\,144)$, μ inconnue. Le poids moyen dans la population ne doit pas être supérieur à $\mu = 50$. Aussi, on décide de rejeter le lot si $\overline{X} > C$, où $\overline{X}$ est la moyenne échantillonnale.

    a) Quelle doit être la valeur de C s'il faut que la probabilité de rejeter un lot pour lequel $\mu = 50$ soit de 5 %?

    b) Avec la valeur de C trouvée en a), déterminer la probabilité de rejeter un lot pour lequel $\mu = 51$, 55, 60, 70.

16. Sachant que 20 % de la population étudiante sont gauchers, une administration universitaire meuble son nouvel amphithéâtre de 60 pupitres pour gauchers et 240 pupitres pour droitiers. Quelle est la probabilité que lorsque 250 étudiants entrent dans l'amphithéâtre il n'y ait pas assez de places pour tous les gauchers?

17. D'une population d'employés, on compte prélever un échantillon de 100 personnes afin d'estimer le pourcentage p de fumeurs parmi eux. On estimera p, bien sûr, par le pourcentage échantillonnal équivalent, soit $X/100$, où X est le nombre de fumeurs dans l'échantillon. Quelle est la probabilité de se tromper d'au moins 5 points de pourcentage si la valeur *réelle* de p est *a*) 10 %, *b*) 20 %, *c*) 30 %?

18. Soit X_1, X_2 et X_3 trois variables aléatoires indépendantes, chacune de loi $\mathfrak{N}(\mu ; \sigma^2)$. Soit $X = X_1 + 2 X_2 - 2 X_3$. Quelle est la distribution de $(X-\mu)^2/(9\sigma^2)$?

19. On doit prélever deux échantillons de deux populations de moyennes μ_1 et μ_2 et variances connues σ_1^2 et σ_2^2. On a l'intention d'estimer $\mu_1 - \mu_2$ par $\overline{X}_1 - \overline{X}_2$. On dispose d'un budget d'échantillonnage limité, de sorte que les tailles n_1 et n_2 des deux échantillons doivent satisfaire $n_1 + n_2 = 100$. Comment doit-on choisir n_1 et n_2?

20. Soit $\overline{X}_1, \ldots, \overline{X}_k$ les moyennes de k échantillons indépendants de tailles $n_1, \ldots, n_k$, tirés tous de la même population de moyenne μ et de variance σ^2.

a) Pour estimer μ on considère l'estimateur linéaire $\hat{\mu} = a_1 \overline{X}_1 + a_2 \overline{X}_2 + \ldots + a_k \overline{X}_k$,

où $a_1, \ldots, a_k$ sont des constantes. Quelle(s) condition(s) les a_i doivent-elles satisfaire pour que $\hat{\mu}$ soit sans biais?

b) Parmi tous les estimateurs linéaires sans biais, montrer que celui qui a la plus petite variance est celui où $a_i = n_i/n$ où $n = \sum_{i=1}^{k} n_i$. [*Indice*: montrer que $\sum_{i=1}^{n} \frac{a_i^2}{n_i} \geq \frac{1}{n}$ à partir de l'énoncé

$\sum_{i=1}^{n} n_i \left(\frac{a_i}{n_i} - \frac{1}{n} \right)^2 \geq 0$.]

21. Si $X \sim \mathcal{B}(n ; p)$ et $\hat{p} = \frac{X}{n}$, montrez que $Var(\hat{p}) \leq 1/4n$.

22. Soit $T \sim t_{20}$. En se servant de la table déterminer la valeur de b telle que

a) $P\{T > b\} = 0{,}025$ *b*) $P\{T < b\} = 0{,}10$ *c*) $P\{T < b\} = 0{,}99$
d) $P\{|T| > b\} = 0{,}10$ *e*) $P\{|T| > b\} = 0{,}05$ *f*) $P\{|T| < b\} = 0{,}99$.

23. D'une population de comptes de dépenses on tire un échantillon de 120 comptes pour estimer la proportion p des comptes qui comportent un voyage en dehors de la ville. Soit X le nombre de comptes dans l'échantillon qui comportent un voyage en dehors de la ville. Si $X = 28$, déterminer un intervalle de confiance à 95% pour p.

24. Montrez qu'un intervalle construit par la formule $\hat{p} - 1{,}88\,\hat{\sigma}_{\hat{p}} \leq p \leq \hat{p} + 2{,}05\,\hat{\sigma}_{\hat{p}}$ est un intervalle de confiance à 95% (approximativement) [vous supposerez que $(\hat{p} - p)/\hat{\sigma}_{\hat{p}} \sim \mathfrak{N}(0 ; 1)$].

25. On veut estimer une proportion p avec une probabilité d'au moins 95 % de ne pas se tromper de plus de 0,01. On ne connaît pas la valeur de p mais on peut être sûr qu'elle n'est pas supérieure à 15%. Quelle est, au maximum, la taille de l'échantillon qu'il faudrait prélever? [Servez-vous du fait que la fonction $p(1-p)$ est croissante sur l'intervalle (0 ; 1/2).]

26. Considérons le problème qui consiste à déterminer la taille de l'échantillon qu'on devrait prélever pour qu'avec une probabilité de 95% l'erreur *relative* dans l'estimation de p ne dépasse pas 20% [L'erreur relative est l'erreur exprimée comme proportion de p: erreur relative=$|\hat{p} - p|/p$]. Supposer que la meilleure estimation disponible de p est $p = 0{,}10$ et estimer la taille de l'échantillon qu'il faut prélever.

27. Soit $X_1, \ldots, X_n$ un échantillon d'une population $\mathcal{N}(\mu\ ;\ \sigma^2)$. Dans cet exercice, on suppose μ connu, et on compare l'estimateur $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)^2$ à l'estimateur $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$.

a) Montrez que $n\hat{\sigma}^2/\sigma^2 = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_n^2$. En déduire que $\hat{\sigma}^2$ est un estimateur sans biais de σ^2 .

b) Montrer que $Var[n\ \hat{\sigma}^2/\sigma^2] = 2n$. En déduire que $Var[\hat{\sigma}^2] = 2\sigma^4/n$.

c) De manière similaire, montrer que $Var[S^2] = 2\sigma^4/(n-1)$.

d) Quel est le meilleur estimateur? *Justifier!*