Tests d'hypothèses

Inférence à un échantillon

Exemple

Après l'obtention de votre diplôme, vous travaillez comme acheteur d'épicerie. Le gérant vous demande d'acheter du maïs pour tester vos compétences. Le problème n'est pas le prix, mais la quantité de maïs par boîte, prétendument 10 grammes selon le vendeur. Vous considérez quatre possibilités :

- 1. Il est honnête : $\mu = 10$ g.
- 2. Conservateur : $\mu > 10$ g.
- 3. Trompeur : μ < 10 g.
- 4. Novice : $\mu \neq 10$ g.

Cas I. Vérifier que le vendeur est conservateur

Supposons que le vendeur, plutôt timide et peu confiant, semble conservateur dans sa revendication de 10 g/boîte. Deux prédictions émergent de cette observation : l'hypothèse nulle $(H_0:\mu\leq 10\mathrm{g})$ représente la déclaration du vendeur que nous testons, tandis que l'hypothèse alternative $(H_1:\mu>10\mathrm{g})$ exprime votre croyance.

$$H_0: \mu \leq 10g$$

$$H_1: \mu > 10g$$

On prend n=25 cases et supposons que $\sigma=1$ g, on obtient $\bar{x}=10.36$ g. Puisqu'on ne peut pas calculer $\bar{x}\geq 10.36$ directement, on calcule la transformation Z comme suit :

$$Z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z = \frac{10.36 - 10.00}{\frac{1}{\sqrt{15}}} = \frac{0.36}{0.20} = 1.8$$

On obtient :

$$\begin{split} P(\bar{X} \geq 10.36 \mathrm{g}) &= P(Z \geq 1.8) \\ &= 1 - P(Z < 1.8) \\ &= F(1.8) \\ &= 1 - 0.9641 \\ &= 0.0359 \end{split}$$

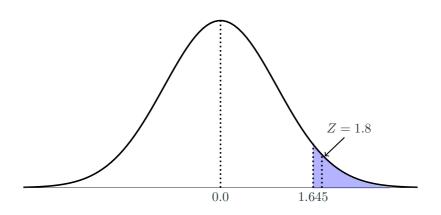


Figure 1: Distribution des valeurs de z

Cas II. Vérifier que le vendeur est un tricheur

Dans le premier cas, nous avons formé une position et l'avons testée objectivement à l'aide d'un échantillon et d'une compréhension de la distribution d'échantillonnage de la moyenne. Maintenant, avec un point de vue différent sur le vendeur — peut-être en le considérant comme malhonnête — vous soupçonnez des pratiques contraires à l'éthique. Votre hypothèse nulle est $H_0: \mu \geq 10$ g, et l'hypothèse alternative est $H_1: \mu \leq 10$ g. Notons l'exclusivité des hypothèses et le signe égal toujours présent dans l'hypothèse nulle, qui représente l'affirmation à tester : l'honnêteté du vendeur.

 $H_0: \mu \geq 10g$

 $H_1: \mu < 10g$

$$Z = \frac{10.36 - 10.00}{\frac{1}{\sqrt{25}}} = \frac{0.36}{0.20} = 1.8$$

$$P(\bar{X} < 10.36g) = P(Z < 1.8)$$

$$= F(1.8)$$

$$= 0.9641$$

 $Z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}}$

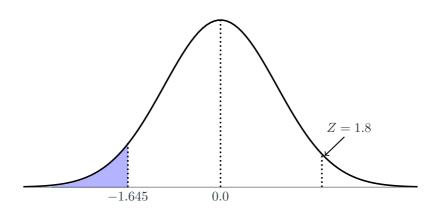


Figure 2: Distribution des valeurs de \boldsymbol{z}

Cas III. Vérifier que le vendeur n'a aucune idée

Dans ce cas, nous ne savons pas si la moyenne de l'échantillon sera supérieure ou inférieure à ce que prétend le vendeur. Le vendeur, étant nouveau dans son métier et peu familier avec le produit, s'appuie sur l'affirmation donnée de 10 g par boîte. Contrairement au cas I (conservateur) ou au cas II (malhonnête), nous ne le percevons ni comme trop conservateur ni malhonnête. L'hypothèse alternative est moins spécifique, suggérant que la moyenne est différente de 10 g. Les prédictions sont alors:

$$H_0: \mu = 10 \text{ g}$$

 $H_1: \mu \neq 10 \text{ g}$

$$H_1: \mu \neq 10 \text{ g}$$

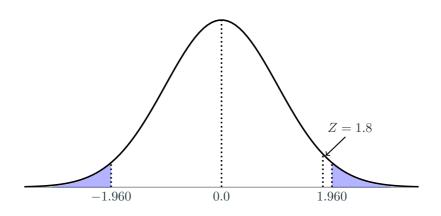


Figure 3: Distribution des valeurs de \boldsymbol{z}

Étapes d'un test d'hypothèse

Nous énumérons ici les étapes générales suivies dans un test d'hypothèse :

- 1. Indiquer le problème. Dois-je acheter des maïs chez ce vendeur?
- 2. Formuler les hypothèses nulles et alternatives :

$$H_0: \mu = 10 \text{ g}$$

 $H_1: \mu \neq 10 \text{ g}$

- Choisissez le niveau de signification. Cela signifie choisir la probabilité de rejeter une vraie hypothèse nulle. Nous avons choisi 5% ou 0.05.
- 4. Déterminez le test statistique approprié.
- 5. Calculez le test statistique approprié :

$$Z = \frac{10.36 - 10.00}{\frac{1}{\sqrt{25}}} = \frac{0.36}{0.20} = 1.8$$

Étapes d'un test d'hypothèse

6. (a) Déterminez les valeurs critiques pour la distribution d'échantillonnage et le niveau de signification approprié.



- (b) Déterminer la valeur p du test statistique.
- 7. (a) Comparer le test statistique a la valeur critique.
 - (b) Comparer la valeur p au niveau de signification choisi.
- 8. Sur la base de la comparaison de l'étape 7, accepter ou rejeter H_0 .
 - (a) Si Z se situe entre les valeurs critiques, elle n'est pas suffisamment extrême pour rejeter H_0 .
 - (b) Si P(Z) est supérieure a 0.05, Z n'est pas suffisamment rare pour rejeter H_0 .
- Énoncez votre conclusion et répondez à la question posée à l'étape
 1.

Quand H_0 est vraie, alors H_1 est fausse

- Si on accepte H_0 , on a fait le bon choix.
- Si on rejette H_0 , on a commit une erreur.

Ce type d'erreur est appelé erreur de type I.

Definition

P(Erreur de type I) = la probabilité de rejeter une hypothèse nulle vraie.

Quand H_0 est fausse, alors H_1 est vraie

- Si on accepte H_0 , on a commit une erreur.
- Si on rejette H_0 , on a fait le bon choix.

Ce deuxième type d'erreur est appelé erreur de type II.

Definition

P(Erreur de type II) = la probabilité d'accepter une hypothèse nulle fausse.

Résultat du test	Situation réelle inconnue	
	$ m H_0$ est vraie	H ₀ est fausse
Accepter H_0	Décision correcte	Erreur type II
	$1-\alpha$	β
Rejeter H_0	Erreur type I	Décision correcte
	α	$1-\beta$

Table 1: Les résultats possibles du test H_0

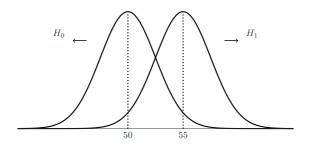
Exemple

Dans cet exemple particulier, supposons qu'une population est soit normalement distribuée avec une moyenne μ de 50 et un écart type σ de 10, soit qu'elle est normalement distribuée avec une moyenne de μ 55 et un écart type σ de 10. Nous ne pouvons pas savoir laquelle de ces populations nous échantillonnons, mais c'est l'une ou l'autre.

$$H_0: \mu = 50, \quad \sigma = 10$$

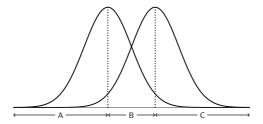
$$H_1: \mu = 55, \quad \sigma = 10$$

Si nous prenons un échantillon de taille n=25 et calculons \bar{X} pour cet échantillon, alors \bar{X} provient de l'une des deux distributions d'échantillonnage possibles.



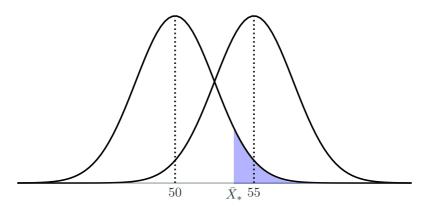
- Si H_0 est vraie, \bar{X} suit une loi $\mathcal{N}(\mu=50,\frac{\sigma}{\sqrt{n}}=2)$.
- Si H_1 est vraie, \bar{X} suit une loi $\mathcal{N}(\mu=55,\frac{\sigma}{\sqrt{n}}=2).$

Notez que ces distributions d'échantillonnage ont la même forme (normale) et la même répartition $(\frac{\sigma}{\sqrt{n}})$, mais sont centrées à différentes positions sur la droite, c'est-à-dire qu'elles ont des moyennes différentes.



- A Les moyennes de l'échantillon dans cette partie de la courbe indiquent évidemment $H0: \mu=50$ comme hypothèse correcte.
- B Les moyennes de l'échantillon dans l'intersection sur la droite pourraient indiquer l'une ou l'autre population comme distribution d'échantillonnage.
- C Les moyennes de l'échantillon dans cette partie de la droite numérique indiquent évidemment $H_1: \mu=55$ comme hypothèse correcte.

On suppose n=25 and $\alpha=0.05$. Soit \bar{X}_* le point de coupure pour accepter H_0 . Nous pouvons trouver cette valeur en utilisant la loi normale.



Nous savons que $0.05 = P(Z \ge 1.645)$. Mais

$$Z = \frac{\bar{X}_* - \mu}{\frac{\sigma}{\sqrt{n}}}$$

alors

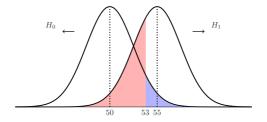
$$1.645 = \frac{\bar{X}_* - 50}{\frac{10}{\sqrt{25}}}$$

Où $\mu = 50$ vient de H_0 . Donc

$$\bar{X}_* = 50 + 1.645 \frac{10}{\sqrt{25}} = 53.29$$

En utilisant $\alpha=0.05$ nous rejetterions $H_0:\mu=50$ chaque fois que l'échantillon $\bar{X}\geq 53.29$ et ferions une erreur de type I.

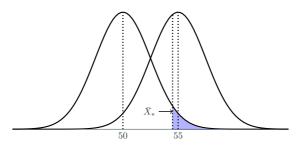
Et si H_0 était faux et H_1 était vrai ? Désormais, la distribution d'échantillonnage est H_1 et non H_0 et toute valeur inférieure à 53.29 serait considérée comme confirmant H_0 .



Puisque H_0 est faux, tout \bar{X} inférieur à 53.29 entraînerait une erreur de type II. Pour ce cas particulier, nous pouvons maintenant calculer la probabilité d'une erreur de type II qui correspond à la zone rouge de la courbe.

$$P(\bar{X} < 53.29) = P\left(Z < \frac{53.29 - 55}{\frac{10}{\sqrt{25}}}\right) = P(Z < -0.855) = 0.1963$$

Supposons maintenant que nous ayons exactement la même situation mais réduisons α à 0,01. Que se passe-t-il et quelle est la puissance du test?

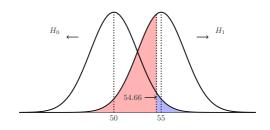


Avec un α plus petit, le point de coupure pour rejeter H_0 est décalé vers la droite car dans la distribution normale standard $0.01=P(Z\geq 2.33)$. Alors cette fois

$$\bar{X}_* = 50 + 2.33 \frac{10}{\sqrt{25}} = 54.66$$

Et si H_0 était faux et H_1 était vrai? La distribution d'échantillonnage est illustrée dessous, et toute valeur inférieure à 54.66 serait considérée comme confirmant H_0 . Encore une fois, comme H_0 est désormais faux, tout \bar{X} inférieur à 54.66 entraînerait une erreur de type II. Cette probabilité est

$$P(\bar{X} < 54.66) = P\left(Z < \frac{54.66 - 55}{\frac{10}{\sqrt{25}}}\right) = P(Z < -0.17) = 0.4325$$



Exemple

Un écologiste forestier étudie la régénération des communautés de la forêt tropicale suite à la chute d'arbres pendant des tempêtes. Elle a lu que les semis de l'arbre à aiguillons, *Dendrocnide excelsa*, croîtraient de 1.5 m par an en plein soleil dans de telles trouées. Dans sa parcelle d'étude, elle a identifié et mesuré 9 spécimens de cette espèce en 2013 et un an plus tard. Les changements de hauteur de ces 9 spécimens sont indiqués ci-dessous. Ses données confirment-elles l'affirmation selon laquelle les semis de cette espèce grandiront en moyenne de 1.5 m par an à la lumière directe du soleil ?

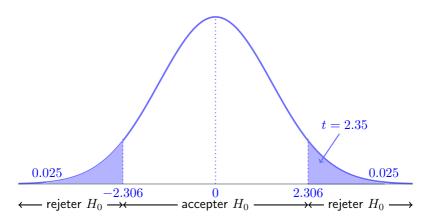
1.9 2.5 1.6 2.0 1.5 2.7 1.9 1.0 2.0

L'écologiste recherche des écarts de 1.5 m dans les deux sens, nous avons donc ici un test bilatéral :

$$H_0: \mu_d = 1.5$$
 m/an $H_1: \mu_d
eq 1.5$ m/an

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{1.90 - 1.50}{\frac{0.51}{\sqrt{9}}} = \frac{0.40}{\frac{0.51}{3}} = 2.35$$

Nous avons une distribution d'échantillonnage pour t avec v=n-1-=9-1=8, et $\alpha=0.05$ pour un test bilatéral



Exemple

Les documents de l'industrie baleinière des îles du Vent indiquent que dans les années 1920, les globicéphales à bec court pesaient en moyenne 360 kg. Une étude récente dans les mêmes eaux a montré qu'un échantillon aléatoire de 25 individus avait un poids moyen de 336 kg, avec un écart-type de 30 kg. Suggère-t-on une diminution significative du poids moyen des globicéphales depuis les années 1920 ?

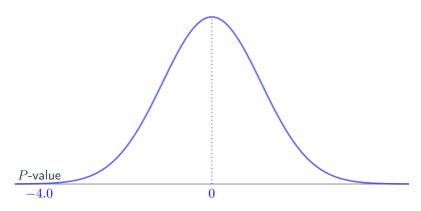
Nous prévoyons un écart par rapport à la valeur revendiquée dans une direction particulière, c'est-à-dire une diminution du poids. Il s'agit d'un test unilatéral à gauche

$$H_0: \mu \ge 360 \text{ kg}$$

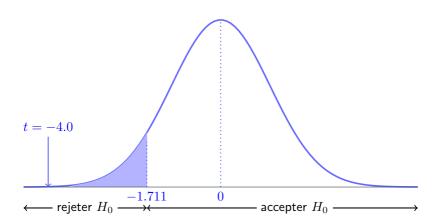
 $H_1: \mu < 360 \text{ kg}$

$$t = \frac{X - \mu}{\frac{s}{\sqrt{n}}} = \frac{336 - 360}{\frac{30}{25}} = \frac{-24}{\frac{30}{\sqrt{25}}} = \frac{-24}{6} = -4.0$$

Nous avons v=n-1=25-1=24, et $P\left(t<-3.091\right)=0.0025$, alors $P\left(t\leq-4.0\right)<0.0025$.



Si nous avions utilisé une approche de valeur critique et un niveau $\alpha=0.05$, nous aurions trouvé que le seuil de rejet de H_0 est de -1.711 pour v=24



Exemple

Pour évaluer l'efficacité d'un nouveau produit de pulvérisation contre les acariens de la rouille dans les vergers, un chercheur a comparé le rendement moyen de 16 vergers traités, pulvérisés selon un calendrier recommandé (moyenne de 814 boîtes de fruits commercialisables, écart-type de 40 boîtes), avec le rendement moyen de vergers non traités dans la même région au cours des 10 dernières années (moyenne de 800 boîtes). Les données suggèrent-elles un rendement moyen significativement supérieur dans les vergers traités par rapport aux vergers non traités ?

Nous anticipons un changement particulier, une augmentation du rendement, nous disposons donc d'un test unilatéral droit :

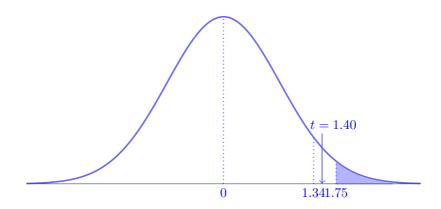
$$H_0: \mu \leq 800$$
 boites

$$H_1: \mu > 800 \text{ boites}$$

$$t = \frac{X - \mu}{\frac{s}{\sqrt{n}}} = \frac{814 - 800}{\frac{40}{16}} = \frac{14}{10} = 1.40$$

$$P\left(t > 1.341\right) = 0.10 \text{ et } P\left(t > 1.753\right) = 0.05 \text{, alors}$$

$$0.05 < P\left(t \ge 1.40\right) < 0.10$$



Hypothèses concernant la variance σ^2

Exemple

Un obtenteur a sélectionné des pommiers Crispin pendant de nombreuses années pour réduire la variabilité de la taille des fruits. Lorsqu'il a commencé, l'écart-type était de 20 g. Un échantillon aléatoire de 30 pommes de sa dernière récolte a montré un écart-type de 15 g. Y a-t-il des preuves que l'obtenteur a réussi à réduire la variabilité des pommes Crispin? Quelle est la valeur attendue de la statistique de test si H_0 est vraie?

Hypothèses concernant la variance σ^2

Écrivons H_0 et H_1 en termes de variance car nous disposons d'une distribution d'échantillonnage pour les variances mais pas pour les écarts types :

$$H_0: \sigma^2 \ge 400 \text{ g}^2$$

 $H_1: \sigma^2 < 400 \text{ g}^2$

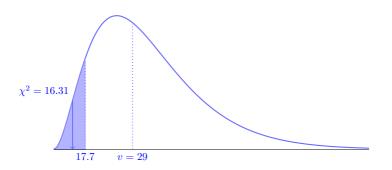
lci n=30 et $s^2=225~{\rm g}^2$, le test est le suivant :

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2}$$

Hypothèses concernant la variance σ^2

avec v=n-1=30-1=29 et en choisissant $\alpha=0.05$, la valeur critique est 17.7. Pour notre échantillon nous avons :

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2} = \frac{(30-1)\cdot 15}{400} = 16.31$$



Exemple

Les *dugongs*, grands mammifères marins, vivent normalement en troupeaux d'une cinquantaine d'individus. Les activités humaines le long de la côte sud-est du Queensland ont limité les populations, conduisant à des groupes dispersés et probablement plus consanguins. Cette consanguinité peut augmenter la variabilité de la taille des animaux entre les troupeaux. Les populations de *dugongs* sains ont un poids moyen de 350 kg avec une variance de 900 kg². Un échantillon de 25 *dugongs* de la baie de Moreton a montré une variance de 1600 kg². La population de la baie de Moreton est-elle significativement plus variable que les populations standard?

Nous prévoyons des écarts plus importants que la valeur déclarée ce qui en fait un test unilatéral :

$$H_0: \sigma^2 \le 900 \text{ kg}^2$$

 $H_1: \sigma^2 > 900 \text{ kg}^2$

ici n=25 et $s^2=1600~{\rm kg}^2$. Sous H_0 on s'attend à ce que χ^2 soit n-1(24), et sous H_1 on s'attend à ce que χ^2 soit supérieure à n-1

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2} = \frac{(25-1)\,1600}{900} = 42.67$$

On voit que
$$P\left(\chi^2<39.4\right)=0.975$$
, donc $P\left(\chi^2>39.4\right)$, et $P\left(\chi^2<43.0\right)=0.990$, donc $P\left(\chi^2>43.0\right)=0.01$, alors
$$0.025>P\left(\chi^2\leq42.67\right)>0.01$$

Exemple

Un généticien étudie les modèles de croissance des hommes américains depuis 1900. Une monographie de 1902 indique que la taille moyenne des hommes adultes était de 67.0 pouces avec un écart type de 3.5 pouces. Pour vérifier les changements au cours du vingtième siècle, le généticien a mesuré un échantillon de 28 hommes adultes, obtenant X=69.4 pouces et s=4.0 pouces. Ces valeurs diffèrent-elles significativement de celles de 1902 ?

$$H_0: \mu = 170 \; \mathrm{cm}$$
 $H_1: \mu \neq 170 \; \mathrm{cm}$

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{69.4 - 67.0}{\frac{4.0}{\sqrt{28}}} = \frac{2.4}{0.76} = 3.16$$

for v=n-1=27 la valeur critique est ± 2.771 . Puisque 3.16>2.771, on rejette H_0 .

$$H_0: \sigma^2 = 12.25 \text{ cm}^2$$

 $H_1: \sigma^2 \neq 12.25 \text{ cm}^2$

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2} = \frac{(28-1)\,16}{12.25} = 35.3$$

avec $\alpha=0.01$ pour v=27, nous avons les valeurs critiques 11.8 et 49.6, et puisque 11.8 < 35.3 < 49.6, on ne rejette pas H_0 .

Exercice 1

Les boîtes de cornflakes de taille économique normales pèsent en moyenne environ 1134 grammes avec un écart-type d'environ 114 grammes. Un nouveau processus d'emballage est développé pour améliorer le contrôle qualité, visant à réduire significativement la variabilité. Trente boîtes emballées selon le nouveau procédé ont un écart-type d'environ 85.0 grammes. Existe-t-il des preuves d'une significative amélioration de l'uniformité des boîtes de cornflakes avec le nouveau procédé d'emballage ? Quelle est la valeur attendue de la statistique du test si H_0 est vraie ? Trouvez la valeur P de ce test.

Exercice 2

Dans le cadre d'une étude de la communauté benthique de l'île Lady Elliot, 16 étoiles de mer, *Linckia laevigata*, ont été collectées et leur bras le plus long a été mesuré au dixième de centimètre le plus proche. de centimètre le plus proche. Supposez la normalité de cette population.

10.3	11.0	10.5	10.0	11.3	14.5	13.0	12.1
12.1	9.4	11.3	12.0	11.5	9.3	10.1	7.6

- 1. Calculez la moyenne et la variance de l'échantillon.
- 2. Calculez l'intervalle de confiance à 95% pour la moyenne de la population. Expliquer ce que représente cet intervalle de confiance.
- 3. Calculez un test t avec $\alpha=0.05$ pour déterminer si la moyenne est significativement différent de 12 cm. Discuter la relation entre l'intervalle de confiance dans (2) et ce test t bilatéral.
- 4. Calculez l'intervalle de confiance à 95% de la variance de la population. Expliquez à nouveau ce que signifie cet intervalle de confiance.
- 5. Calculez un test χ^2 avec $\alpha=0.05$ pour déterminer si la variance est significativement différente de $5.0~{\rm cm}^2$. Discuter la relation entre l'intervalle de confiance dans (4) et ce test χ^2 .

Exercice 3

Chez les hommes adultes en bonne santé, les valeurs de la vitesse de conduction du nerf sciatique sont normalement distribuées avec une moyenne de 65 cm/msec et un écart type de 5 cm/msec. Les vitesses de conduction de 16 sujets admis au centre de contrôle des intoxications d'un hôpital métropolitain avec un diagnostic d'intoxication au méthylmercure avaient une vitesse de conduction moyenne de 55 cm/msec et une variance de 49 cm/msec².

- 1. Ces données fournissent-elles suffisamment de preuves pour indiquer que les vitesses de conduction sont significativement plus lentes chez les individus empoisonnés?
- 2. Ces données fournissent-elles suffisamment de preuves pour indiquer que les individus empoisonnés sont plus variables dans leurs vitesses de conduction du nerf sciatique que les individus normaux?

Exercice 4

Un documentaire télévisé sur la suralimentation prétend que les hommes américains adultes sont en moyenne en surpoids de 6.80 kg. Pour tester cette affirmation, 25 hommes adultes sélectionnés au hasard ont été examinés, et le poids excessif moyen s'est avéré être de 8.16 kg avec un écart type de 4.54 kg. Ces données nous donnent-elles des raisons de croire que l'affirmation de 6.80 kg est incorrecte ? Quel type d'erreur auriez-vous pu commettre ici ? Expliquez.

Exercice 5

Avant d'acheter une grande exploitation de maïs, vous souhaitez déterminer sa rentabilité en analysant sa productivité. Une affirmation faite par le propriétaire actuel est que les champs sont très uniformes dans leur rendement par acre. En fait, on affirme que l'écart-type en boisseaux par acre du rendement en maïs est de 5. Un échantillon aléatoire de 25 parcelles d'un acre donne une moyenne de 140 boisseaux/acre et un écart-type de 6.5 boisseaux/acre. Les champs sont-ils significativement plus variables que ce qui est affirmé ? Quel type d'erreur auriez-vous pu commettre ici ? Expliquez.

Exercice 6

Des préoccupations sur les effets des phtalates sur le développement du système reproducteur masculin ont été soulevées. Dans une étude pilote, un chercheur a administré quotidiennement à des rates enceintes 750 mg/kg de DEHP (phtalate de di-2-éthylhexyle) pendant le développement des organes sexuels de leurs ratons mâles, sacrifiés et pesés ensuite. Ci-dessous, les poids pour les huit mâles (en mg) :

1710 1630 1580 1670 1350 1650 1600 1650

Si les mâles nouveau-nés non traités ont une moyenne de 1700 mg, peut-on dire que les rats exposés au DEHP in utero ont un poids significativement inférieur?

Exercice 7

Le python tapis est commun en forêt tropicale, forêt d'eucalyptus, zones pastorales et urbaines. Utilisant des fosses thermiques le long de la mâchoire, il repère ses proies à sang chaud, détectant des changements de température inférieurs à un trentième de degré Celsius. Sa longueur médiane est d'environ 250 cm. En raison d'une sécheresse prolongée de 1994 à début 1996 à Lamington National Park, la longueur médiane était supposée être inférieure à 250 cm. Une enquête a été menée, enregistrant les longueurs (en cm) ci-dessous.

269 221 204 258 211 198 276 223 207 242

En supposant que les longueurs des serpents sont symétriquement distribuées autour de la moyenne, effectuez un test approprié de l'hypothèse au niveau de 0.05. Interprétez le résultat.

Merci. Des questions?