Examen

Corrigé

Utilisez exactement ce qui est autorisé et ce que vous allez utiliser, lors de votre examen : documents, calculatrices, stylos, boissons….

Mettez-vous dans des conditions de l'examen.

Question 1

1) Un sondage probabiliste est un sondage aléatoire

☒Vrai ☐faux

2) L'homogénéité d'un échantillon est proportionnelle à sa non-dispersion

☒Vrai ☐faux

3) Un bon estimateur est un estimateur qui

a) Un biais nul
b) Une petite variance
c) Une grande variance
d) Une petite erreur quadratique
e) Une grande erreur quadratique

4) Pour doubler la précision il faut :

a) Doubler la taille de l'échantillon
b) Multiplier par 4 la taille de l'échantillon
c) Diviser par 2 la taille de l'échantillon
d) Diviser par 4 la taille de l'échantillon

5) Un sondage effectué auprès des étudiant-e-s de la 3ème année de Bachelor sur leurs choix d'orientation pour leur master l'année suivante :

S'agit-il d'un :

☒SAS

☐Sondage stratifié

☐ Sondage en grappe

5) A mesure que la taille de l'échantillon augmente,

a) la longueur de l'IC diminue
b) la longueur de l'IC augmente
c) la longueur de l'IC reste la même

Question 2

La croix rouge vous demande d'estimer le pourcentage d' étudiant-e-s qui seraient prêt-e-s à donner 1,5 litres de leur sang. Un de ses agents planifie sa tournée des prochaines semaines pour récolter le sang. Il souhaiterait que vs lui présentiez une estimation avec une précision de 3%, un degré de confiance de 95%.

1. Quelle devrait être la taille de votre échantillon ?

a) Si vous n'avez aucune idée du véritable pourcentage.

Réponse

Quand on n' a aucune idée, on prend le pire cas soit 50%

On cherche à déterminer la taille de l'échantillon qu'il faut pour estimer la proportion π d' étudiant-e-s qui seraient prêt-e-s à donner 1,5 litres de leur sang, il suffit d'appliquer la formule.

$$\boldsymbol{n} = \left(\frac{\boldsymbol{Z}_{\alpha/2}}{\boldsymbol{d}}\right)^2 \boldsymbol{\pi}(1-\boldsymbol{\pi}) = \left(\frac{1{,}96}{0{,}03}\right)^2 0{,}5(1-0{,}5) = 1067{,}1 \approx 1067$$

Donc il faut un échantillon de 1067 étudiants

b) 70% en se basant sur les autres fois.

Réponse

$$\boldsymbol{n} = \left(\frac{\boldsymbol{Z}_{\alpha/2}}{\boldsymbol{d}}\right)^2 \boldsymbol{\pi}(1-\boldsymbol{\pi}) = \left(\frac{1{,}96}{0{,}03}\right)^2 0{,}7(1-0{,}7) = 896{,}37 \approx 896$$

Donc il faut un échantillon de 896 étudiants

2. Supposons que sur les 10000 étudiants de l'université, un échantillon de 800 a été trié au hasard, 710 ont affirmé qu'ils seraient prêt-e-s à donner 1,5 litres de leur sang. Donner un intervalle de confiance à 95% pour la proportion de ces étudiant-e-s.

Réponse

D'abord, il faut vérifier les conditions nécessaires :

$$\boldsymbol{n\bar{p} \geq 5 \quad et\; n(1-\bar{p}) \geq 5}$$

$$\bar{p} = \frac{710}{800} \approx 0{,}89$$

800.0,89 =710 et 800.(1-0,89) =90. C'est ok

L'intervalle de confiance pour estimer la proportion d'étudiant-e-s qui seraient prêt-e-s à donner 1,5 litres de leur sang est donné par :

$$\boldsymbol{IC = \bar{p} \mp Z_{\frac{\alpha}{2}}\sqrt{\widehat{var}(\bar{p})}}$$

On a déjà calculé $\bar{p}$ =0,89. Calculons $\widehat{var}(\bar{p})$ **:**

$$\widehat{var}(\bar{p}) = (1-f)\frac{\bar{p}(1-\bar{p})}{n-1}$$

$$\widehat{var}(\bar{p}) = \left(1-\frac{800}{10000}\right)\frac{0{,}89.(1-0{,}89)}{800-1} = 0{,}00011$$

IC=0,89∓ 0,02=[0,87 ;0,91]

Ainsi, avec un niveau de confiance de 95%, le pourcentage des étudiant-e-s qui seraient prêt-e-s à donner 1,5 litres de leur sang se trouve entre 87% et 91% avec une marge d'erreur d'env. 2%.

Question 3

Un promoteur immobilier veut construire des villas de luxe dans une nouvelle région.

Il se base sur une enquête effectuée dans une région semblable qui a conclu que l'écart-type des prix pour ce genre d'habitation est de 15 000 CHF. Ce promoteur vous engage de faire cette étude en vous précisant qu'il veut être sûr à 95% que les résultats que vous alliez fournir donneront une estimation à 2 500 CHF près du véritable coût moyen de construction, quelle est la taille de l'échantillon qu'il vous faudra ?

Réponse

On cherche à déterminer la taille de l'échantillon qu'il faut pour estimer le coût moyen de construction avec une précision (une marge d'erreur) de 2 500 CHF

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{d}\right)^2 . \sigma^2 = \left(\frac{1{,}96}{2\,500}\right)^2 . 15\,000^2 = 138{,}3$$

Donc, il faut prendre un échantillon de 138 villas pour estimer la moyenne de leurs coûts de construction avec une marge d'erreur (précision absolue) de 2 500 CHF.

Question 4

Le canton de Genève souhaite connaître le prix d'un litre de lait frais. La liste complète de tous les distributeurs compte 1200 magasins (des grandes surfaces aux superettes). Un échantillon de 85 magasins est déterminé. Sur celui-ci, le sondage réalisé montre que le prix moyen pour un litre de lait frais est de 2,75 CHF avec une variance de 0,35 CHF^2.

1) Quel est le taux de sondage ?

Réponse

Le taux de sondage correspond à :

f=n/N=85/1200=0,071 soit env. 7%

2) Estimer la variance de l'estimateur de la moyenne

Réponse

La variance de l'estimateur de la moyenne est donnée par :

On peut négliger f

$$\widehat{var(\overline{x})} = (1-f)\frac{\hat{\sigma}^2}{n} = (1-0{,}071).\frac{0{,}35}{85} = 0{,}0038 \approx 0{,}004$$

3) Donner un intervalle de confiance à 95% pour le prix d'un litre de de lait frais dans le canton de Genève

Réponse

On cherche à estimer l'IC pour la moyenne à 95%, ce qui revient à calculer :

$$IC = \bar{x} \mp z_{\frac{\alpha}{2}}.\sqrt{\widehat{var(\bar{x})}} = 2{,}75 \pm 1{,}96.\sqrt{0{,}004}$$

IC=[2,63 ; 2,87]

Ainsi, avec un niveau de confiance de 95%, le prix moyen d'un litre de lait se trouve entre 2,63 et 2,87 avec une marge d'erreur d'env. 0,13.

4) Est-il plausible que le prix moyen sur les 1200 distributeurs excède 2,85 FCH ? Justifiez votre réponse.

Réponse

Oui, c'est possible que le prix moyen sur les 1200 distributeurs excède 2,85 CHF puisque cette valeur est incluse dans l'IC (la borne sup. De l'IC est de 2,87 CHF)

Question 5

Une entreprise voudrait investir dans une supérette pour faire gagner du temps à ses employés. On demande aux employés de fournir leurs dépenses mensuelles des produits non périssables (Produits ménagers, d'hygiènes, jus, lait, conserves…)

Le tableau ci –dessous donne la réparation.

Strate h	**Catégorie**	**Effectifs N_h**	**Nb d'observations. n_h**	**Moyenne observée en CHF $\bar{x}_h$**	**Ecart-type en CHF s_h**
1	Ouvriers	500	50	1300	200
2	Employés de bur.	300	30	1400	250
3	Cadres	120	12	2200	400
4	Cadres supérieurs	20	2	3100	500
Total		N=940	n=94		

1) Avec un degré de confiance de 95%, estimer la moyenne des dépenses mensuelles des produits non périssables ;

Réponse

On cherche à estimer l'IC pour la moyenne à 95%, et comme on a un sondage stratifié, l'IC est :

$$IC = \bar{x}_{str} \mp Z_{\frac{\alpha}{2}}\sqrt{Var(\bar{x}_{str})}$$

On calcule la moyenne strate :

$$\hat{\mu} = \bar{x}_{str} = \frac{1}{N}\sum_{h=1}^{H} N_h \bar{x}_h = \frac{1}{940}(500.1300 + \cdots. +20.1300)$$
$$= 1485{,}1 \text{ CHF}$$

On calcule la variance de l'estimateur :

$$\widehat{Var}(\bar{x}_{str}) = \sum_{h=1}^{H} W_h^2 \, \widehat{Var}(\bar{x}_h) = \sum_{h=1}^{H} W_h^2 . (1 - f_h) . \frac{s_h^2}{n_h}$$
$$= \left(\frac{500}{940}\right)^2 . \left(1 - \frac{50}{500}\right) . \frac{200^2}{50} + \cdots + \left(\frac{20}{940}\right)^2 . \left(1 - \frac{2}{20}\right) . \frac{500^2}{2}$$
$$= 641{,}2 \; CHF^2$$

IC=1484,1∓1,96.√641,2=[1435,5 ; 1534,7]

Donc, la moyenne des dépenses mensuelles des produits non périssables se trouve entre 1435,5 et 1535 CHF avec une marge d'erreur d'env. 49,63 CHF

2) Avec un degré de confiance de 95% et en gardant la même taille de l'échantillon (n), estimer l'IC pour la moyenne des dépenses mensuelles des produits non périssables avec une allocation proportionnelle

Réponse

Définissons d'abord la réparation des effectifs pour une allocation proportionnelle.

n_h= $W_h n$

n_1= (500/940).94=50

n_2=(300/940).94=30

n_3=(120/940).94=12

n_4=(20/940).94=2

C’est la même répartition que précédemment. Donc, c’est le même IC.

3) Avec un degré de confiance de 95% et en gardant la même taille de l’échantillon (n), estimer l’IC pour la moyenne des dépenses mensuelles des produits non périssables avec une allocation optimale

Réponse

Définissons d’abord la réparation des effectifs pour une allocation optimale.

$$n_h = \frac{N_h s_h}{\sum_{k=1}^{4} N_k \hat{\sigma}_k}.n$$

$$n_1 = \frac{500.200}{500.200 + 300.30 + 120.12 + 20.2}.94 \approx 40$$

$$n_2 = \frac{300.30}{500.200 + 300.30 + 120.12 + 20.2}.94 \approx 30$$

$$n_3 = \frac{120.12}{500.200 + 300.30 + 120.12 + 20.2}.94 \approx 20$$

$$n_4 = \frac{20.2}{500.200 + 300.30 + 120.12 + 20.2}.94 \approx 4$$

La moyenne stratifiée reste la même, recalculons la variance :

$$Var(\bar{x}_{opt}) = \frac{1}{n}\left(\sum_{h=1}^{H} W_h\sigma_h\right)^2 - \frac{1}{N}\sum_{h=1}^{H} W_h{\sigma_h}^2$$
$$= \frac{1}{94}\left(\frac{\mathbf{500}}{\mathbf{940}}.\mathbf{200} + \cdots + \frac{\mathbf{20}}{\mathbf{940}}.\mathbf{500}\right)^2$$
$$- \frac{1}{940}\left(\frac{\mathbf{500}}{\mathbf{940}}.\mathbf{200^2} + \cdots + \frac{\mathbf{20}}{\mathbf{940}}.\mathbf{500^2}\right)$$
$$= \frac{1}{94}(247{,}87)^2 - \frac{1}{940}.66968{,}08 \approx 582{,}4$$

IC=1484,1∓1.96.√582,4= [1437,81 ;1532,41]

Donc, la moyenne des dépenses mensuelles des produits non périssables se trouve entre 1438 et 1532 CHF avec une marge d’erreur d’env. 47,3 CHF.

Comme on s’y attendait la variance a diminué et donc un IC moins long, donc plus précis par rapport à une répartition avec allocation proportionnelle.

$$1 - \frac{Var(\bar{x}_{opt})}{Var(\bar{x}_{prop})} = 1 - \frac{582{,}4}{641{,}2} \cong 0{,}1$$

La variance d'échantillonnage a diminué de 10%.

Question 6

Une boutique vend en ligne 5000 articles par mois et souhaite estimer le pourcentage des retours, elle prend un échantillon de 140 ventes et trouve 16 articles.

1) Estimer à 95 % l'IC pour le pourcentage des retours.

Réponse

X : v.a qui représente la proportion des articles retournés. D'abord, il faut vérifier les conditions nécessaires :

$$\boldsymbol{n\bar{p} \geq 5 \quad et\; n(1-\bar{p}) \geq 5}$$

$$\bar{p} = \frac{16}{140} \cong 0{,}11$$

140.0,11 =16 et 140.(1-0,11) =124. C'est ok

L'intervalle de confiance pour estimer la proportion de la population des articles retournés est donné par :

$$\boldsymbol{IC = \bar{p} \mp Z_{\frac{\alpha}{2}}\sqrt{\widehat{var}(\bar{p})}}$$

On a déjà calculé $\bar{p}$ =0,11. Calculons $\widehat{var}(\bar{p})$ **:**

$$\widehat{var}(\bar{p}) = (1-f)\frac{\bar{p}(1-\bar{p})}{n-1}$$

$$\widehat{var}(\bar{p}) = \left(1-\frac{140}{5000}\right)\frac{0{,}11.(1-0{,}11)}{140-1} = 0{,}00071$$

IC=0,11∓ 0,052=[0,06 ;0,17]

Ainsi, avec un niveau de confiance de 95%, le pourcentage des retours se trouve entre 6% et 17% avec une marge d'erreur d'env. 5%

2) Estimer à 95 % l'IC pour le nombre d'articles retournés mensuel dans la population des ventes.

Réponse

Posons N_C : le nombre dans la population des articles retournés par mois.

On cherche à estimer à 95% l'IC qui est donné par :

IC= $N_C \mp 1,96\sqrt{var\ (N_C)}$

D'après l'échantillon on sait qu'on a env.11% de retour ce qui fait

$$N_c = N.\bar{p}$$

N_C=5000.0,11= 571,4≈571

Donc, cette boutique enregistre env. 571 retours

Maintenant, on calcule Var(N_C):

$$Var(N_c) = Var(N.\bar{p}) = N^2 Var(\bar{p})$$

*Rappel: var(**a**X)= **a²** Var(x)*

Var(N_C)=5000^2.0,00071= 17696,1

IC= 571,4∓260,7

IC=[310,7 ;832,2] ≈[311 ;832]

Donc, le nombre dans la population des articles retournés par mois se trouve entre 311 et 832 avec une marge d'erreur d'env. 261 articles.