

Corrigé de l'examen de décembre 2020

Exercice 1 de l'examen de 2020

Question 1 : loi normale

Question 2 : loi du chi-deux

Question 3 : test de normalité

Question 4 : test d'égalité de deux variances

Question 5 : la moyenne

Question 6 : diminuer la zone de rejet

Question 7 : détermine si on rejette une hypothèse

Question 8 : HP.

Exercice 2 de l'examen de 2020

Question 1 :

53

1

continue

note

3

groupe, alphabet, genre.

2

10,64

13,00

30

10,97

3,36

23

10,22

3,96.

Question 2 :

La fonction `shapiro.test()` réalise un test statistique d'adéquation avec une loi normale dont les hypothèses suivantes sont :

H_0 : la variable étudiée suit une loi normale

H_1 : la variable étudiée ne suit pas une loi normale.

Question 3 :

p-value = 0.466 (début)

Comme $0,466 > 0,05$, le test de Shapiro-Wilk n'est pas significatif au seuil $\alpha = 0,05$.

Donc nous décidons de ne pas rejeter H_0 au seuil $\alpha = 0,05$ et par conséquent nous décidons d'accepter H_0 au seuil $\alpha = 0,05$.

Nous commençons alors un risque d'erreur de deuxième espèce noté β qu'il faudrait évaluer.

$n=30 \geq 30$ vous n'avez pas à calculer β .

p-value = 0.05409

Comme $0,05409 > 0,05$ le test de Shapiro-Wilk n'est pas significatif au seuil $\alpha = 0,05$.

Donc nous décidons de ne pas rejeter H_0 au seuil $\alpha = 0,05$ et par conséquent nous décidons d'accepter H_0 au seuil $\alpha = 0,05$.

Nous commençons alors un risque d'erreur de deuxième espèce noté β qu'il faudrait évaluer.

N=23...un peu limite

Question 4 :

(a) Quel test proposez-vous de faire ?

On va réaliser un test de dispersion avec la fonction `var.test()`.

(b) Quelles sont les conditions d'applications et les hypothèses de test ?

⇒ Normalité de la variable « notes » au sein de chacun des deux groupes ? Oui d'après la question 3.

H_0 : variance_fin = variance_début

H_1 : variance_fin != variance_début

(c) Donner la statistique de test, la p-valeur et conclure.

F = 0.72146,

Comme la p-valeur=0,4059>0,05, le test de variance n'est pas significatif au seuil alpha=0,05. Nous décidons de ne pas rejeter H_0 au seuil alpha et par conséquent d'accepter H_0 au seuil alpha=0,05. Nous commettons un risque d'erreur de deuxième espèce beta qu'il faudrait évaluer.

⇒ H_0 : les deux variances sont égales.

(d)(bonus) Quelle quantité faudrait-il calculer pour préciser le risque d'erreur de la conclusion du test?

beta=Probabilité (H_0 | H_1).

Question 5. Comparons maintenant les moyennes de ces deux groupes. En suivant le même procédé qu'à la question précédente, répondre à la problématique.

H_0 les deux moyennes (début et fin) sont égales

H_1 les deux moyennes (début et fin) ne sont pas égales.

T=0,74319

Comme la p-valeur=0,4608>0,05, le test de moyennes n'est pas significatif au seuil alpha=0,05.

Nous décidons de ne pas rejeter H_0 au seuil alpha =0,05...

⇒ H_0 : les deux moyennes sont égales.

Exercice 3 de l'examen de 2020

Question 1. Définir les variables étudiées. Donner leurs types.

Les deux variables étudiées sont : genre et salaire. Les deux variables sont qualitatives : une est nominale (genre) l'autre ordinale (salaire)

Question 2. Donner le nom du tableau précédent.

Tableau d'effectifs observés.

Question 3. Donner le nom du test à effectuer.

Test du chi-deux d'indépendance.

Question 4. Donner les deux hypothèses H0 et H1 associées au test.

H_0 les deux variables genre et salaire sont indépendantes

H_1 les deux variables genre et salaire ne sont pas indépendantes.

Question 5. Vérifier les conditions d'applications.

(a) $270 > +50$

(b) $\$expected = \text{effectifs théoriques} \geq 5$.

Faible/Homme = $17,77778 > 5$

Faible/Femme = $22,222 > 5$

Moyen/Homme = $P(\text{Homme}) * P(\text{Moyen}) * 270$

$P(\text{Homme}) = 120/270$

$P(\text{Moyen}) = 130/270$

Moyen/Homme = $120/270 * 130/270 * 270 = 120 * 130 / 270 = 12 * 130 / 27 = 57,77778 > 5$

Moyen/Femme = $72.22222 > 5$
Elevé/Homme = $44.444 > 5$
Elevé/Femme = $100 * 150 / 270 = 1500 / 27 = 55.55 > 5$
 $P(\text{Femme}) = 150 / 270$
 $P(\text{Elevé}) = 100 / 270$

Donc toutes les conditions du test du chi-deux d'indépendance sont vérifiées.

Question 6. Donner les effectifs marginaux.

Homme 120

Femme 150

Faible 40

Moyen 130

Elevé 100

Question 7. Donner les effectifs sous hypothèse d'indépendance.

`>chisq.test(Salaire)$expected`

	Faible	Moyen	Elevé
Homme	17.77778	57.77778	44.44444
Femme	22.22222	72.22222	55.555

Question 8. Donner la valeur de la statistique de ce test et la conclusion.

X-squared = 11.579 p-value = 0.00306

Comme la p-valeur $0,00306 < \alpha = 0,05$ le test du chi-deux est significatif au seuil $\alpha = 0,05$.

Nous décidons de ne pas rejeter H_0 au seuil $\alpha = 0,05$. Et par conséquent nous décidons d'accepter H_1 au seuil $\alpha = 0,05$. Nous commençons un risque d'erreur de première espèce égal à $\alpha = 0,05$.

⇒ Les deux variables genre et salaire sont dépendantes.

Question 9. Donner le risque d'erreur avec lequel vous pouvez armer votre décision.

Nous commençons un risque d'erreur de première espèce égal à $\alpha = 0,05$.

Exercice 4 de l'examen 2020

Question 1. Quel est le premier coefficient à estimer pour étudier la liaison de ces variables ? Commenter cette valeur.

Coefficient de corrélation linéaire de Pearson.

`>cor(hom_min_Y,quantite_X)[1]` 0.9939843

Elle est très grande. Donc cela signifie que la corrélation linéaire est très forte.

Question 2. On envisage un modèle linéaire. Donner l'équation de la droite de régression estimée.

$Y(\text{nombre d'hommes-minutes}) = 15,0511 + 4,0166 * \text{quantite}$

Question 3. D'après l'équation de régression, si le nombre d'unités à fabriquer augmente de 10, quelle sera l'augmentation correspondante du nombre moyen d'hommes-minutes requis ?

$10 * 4,0166 = 40,166$.

Question 4. Donner la variation qui est expliquée par la droite de régression et la variation qui est inexpliquée par la droite.

Expliquée 80441

Inexpliquée 977

Question 5. Déterminer le pourcentage de variation qui est expliqué par la droite de régression.

Multiple R-squared: 0.988

98,8%

Question 6. Donner un intervalle de confiance à 95% pour la pente de la droite.

3.7514656 4.281828

Question 7.(bonus) Tester si l'ordonnée à l'origine est nulle.

H_0 ordonnée nulle

H_1 ordonnée non nulle

Comme la p-valeur 0,0612 (lue dans le tableau de la sortie `anova()`) $> 0,050$, alors le test de l'ordonnée à l'origine n'est pas significatif au seuil $\alpha=0,05$. Nous décidons de ne pas rejeter H_0

Question de Guillaume.

Pouvez-vous rappeler les conditions d'application des tests ?

Test de Student : normalité

Test d'égalité de deux variances : normalité càd les mêmes conditions que le test de Student

Test de corrélation linéaire

Test du chi-deux d'indépendance : $n > 50$, $n_{th} \geq 5$.

Test de Shapiro-Wilk : $n > 30$

Estimation de la proportion : HP.