

Umabb/FS/Maths

Filière: Master MSS Semestre2

Module: Stat Non Paramétrique

Mars 2022.

SERIE D'EXERCICES N°1

Statistique Non Paramétrique

Tests Non Paramétrique: Test des signes, test **W** de Wilcoxon ,test **U** de Mann-Whitney et généralisation test Kde Kruskal- Wallis

Exercice 1

Dans le tableau qui suit on rapporte le nombre de pièces défectueuses produites par deux machines différentes A et B en 12 semaines consécutives.

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
A	47	56	54	49	36	48	51	38	61	49	56	52
B	71	63	45	64	50	55	42	46	53	57	75	60

En supposant que la production est la même des deux machines durant ce laps de temps.

On veut tester H_0 L'hypothèse qu'il n'y a pas de différence entre les machines c-à-d que les différences des nombres des pièces défectueuses est dû au hasard.

(i) Appliquer le test des **signes** (dit aussi test de la loi Binomiale) au seuil $\alpha = 0.05$ et $\alpha = 0.01$ (test unilatéral et bilatéral)

(ii) Trouver le seuil minimal de rejet de H_0 .

Exercice 2

Tester l'hypothèse que l'âge d'étudiants des deux échantillons A et B se distribue suivant la même loi.

A: 13; 11; 14ans

B: 10; 16; 15 ans

Appliquer le test **W** de Wilcoxon et le test **U** de Mann-Whitney.

Exercice 3

A la fin de l'année on a relevé les journées d'absences de deux ouvriers et on a trouvé pour le lier

I: 0; 1; 2; 6; 4; 19; 8; 7; 13; 15; 17, étalé sur 11 mois de travail.

Pour le second comptabilisé sur 8 mois.

II: 3; 5; 18; 10; 14; 11; 16; 12.

Peut-on dire que la distribution des absences des deux ouvriers est identique?

Appliquer le test **W** de Wilcoxon et le test **U** de Mann-Whitney.

Exercice 4

Deux groupes de 10 lapins chacun, nourris avec un régime enrichi en cholestérol, ont été soumis à deux traitements différents X et Y. On voudrait savoir si les résultats diffèrent (i-e) S'il y'a une différence significative de la distribution des cholestérolémies observés comme suit:

Avec le traitement X: 23; 15; 28; 26; 13; 8; 21; 25; 24; 29 dg/l

Avec le traitement Y: 18; 22; 33; 34; 19; 12; 27; 32; 31; 30 dg/l

(i) Appliquer le test paramétrique (le test t classique de Student) comparaison des moyennes.

(ii) Appliquer les test NP de Mann-Whitney et de Wilcoxon.

(iii) Comparer entre les deux approches paramétrique et non paramétrique. Conclure.

Exercice 5

Dans le tableau qui suit, on rapporte le nombre de pièces défectueuses produite par deux machines différentes (M1 et M2) en 12 semaines consécutives, en supposant que la production soit la même sur les deux machines durant cette période. On veut tester l'hypothèse H_0 qu'il

n'y a pas de différence significative entre les deux machines.

(i-e) On veut répondre à la question: La distribution du nombre des pièces défectueuses des deux machines est elle identique?

Semaine n°	1	2	3	4	5	6	7	8	9	10	11	12
Machine M1	47	56	54	49	36	48	51	38	61	49	56	52
Machine M2	71	63	45	64	50	55	42	46	53	57	75	60

Appliquer le test **U** et **W**.

Exercice 6

Le tableau suivant donne la résistance de deux câbles d'alliages différents

Alliage I	Alliage II
18.3 ; 16.4; 22.7; 17.8	12.6; 14.1; 20.5; 10.7; 15.9
18.9; 25.3; 16.1; 24.2	19.6; 12.9; 15.2; 11.8; 14.7

Réaliser les tests U et W pour décider s'il y a une différence significative entre ces deux échantillons.

Exercice 7

Une entreprise veut acquérir l'une des cinq machines métalliques différentes $M_i : i=1, \dots, 5$. Une expérience est réalisée afin de tester leurs performances. Cinq opérateurs travaillent sur chacune des machines pendant la même durée de temps.

Le tableau suivant montre le nombre de pièces produites par machine.

M_1	68	72	77	42	53
M_2	72	53	63	53	48
M_3	60	82	64	75	72
M_4	48	61	57	64	50
M_5	64	65	70	68	53

(i) Utiliser le test K de Kruskal- Wallis avec et sans correction avec l'hypothèse nulle H_0 selon laquelle il n'y a pas de différence entre les machines au seuil $\alpha = 5\%$ et $\alpha = 0.01$, discuter les différents seuils et déduire le seuil minimal en utilisant l'approche par p.

(ii) Utiliser le test Anova à un facteur (test F d'égalités des moyennes)

(iii) Comparer les résultats des deux approches (Paramétrique et NP)

solution Détaillée

Tests Non Paramétrique: Tests des signes , U et Wet H(Kruskal Wallis)

Exercice 1

Dans le tableau qui suit on rapporte le nombre de pièces défectueuses produites par deux machines différentes A et B en 12 semaines consécutives.

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
A	47	56	54	49	36	48	51	38	61	49	56	52
B	71	63	45	64	50	55	42	46	53	57	75	60

Solution

Le test des signes consiste à évaluer le signe des différences des nombres des pièces défectueuses des deux machines pour les mêmes jours. Ainsi par exemple pour le premier jour 47-71, ce qui négatif, on obtient ainsi la série des signes:

Semaines	1	2	3	4	5	6	7	8	9	10	11	12
Machine A	47	56	54	49	36	48	51	38	61	49	56	52
Machine B	71	63	45	64	50	55	42	46	53	57	75	60
Le signe de la différence	-	-	+	-	-	-	+	-	+	-	-	-

Résultat: 3 signes + et 9 signes -

Si la probabilité était la même, on attendrait 6 de chaque.

Le test avec comme hypothèse nulle H_0 équivaut donc à tester si une pièce est équilibrée ($p=1/2$) alors que 12 lancers donnent 3 faces (+) et 9 piles (-)

Ce test fait appelle à la loi binomiale $B(n=12, p=\frac{1}{2})$

$$X \hookrightarrow B(n, p) : P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Comme H_1 : Il y a une différence entre les deux machines.

Soit $X=k$ le nombre des + (faces) (ou le nombre des - piles)

(i) **Test unilatéral:**

Sous l'hypothèse H_0 : $P(X \leq 3) = 0.00024 + 0.00293 + 0.01611 + 0.05371 = 0.07299 = p_0$

pour un seuil $\alpha = 0.05$ on accepte H_0

accepter H_0 pour $\alpha < p_0$

à fortiori pour un seuil $\alpha = 0.01$ et tous les seuils inférieurs

Sous l'hypothèse H_0 : $P(X \geq 9) = 0.07299 = p_0$

Remarques:

1: $P(X \leq 3) = P(X \geq 9)$ par symétrie de $B(n, 0.5)$

En effet:

$$X \hookrightarrow B(n, p) \Leftrightarrow n - X \hookrightarrow B(n, 1 - p)$$

comme $p=1/2$: X et $n-X$ suivent la même loi.

2: Pour $X=2$ (s'il y a 2 +) ou 10-

$$P(X \leq 2) = P(X=0, 1, \text{ ou } 2) = 0.00024 + 0.00293 + 0.01611 = 0.01928$$

on aurait rejeter H_0 pour le seuil 0.05 car $0.05 > 0.01928$

et accepter pour le seuil 0.01 car $0.01 < 0.01928$

(ii) Test bilatéral

notre cas le nombre de + est 3

$$P(X \leq 3) = 0.07299$$

comme $0.025 < 0.07299$ on accepte H_0

Aussi comme le nombre de - est 9: $P(X \geq 9) = 0.07299$

On accepte H_0

Reprenons avec l'approximation de la loi binomiale avec correction de continuité.

$$X \hookrightarrow B(n, p) \Rightarrow Z = \frac{X - np}{\sqrt{np(1-p)}} \hookrightarrow N(0, 1)$$

Comme X suit une loi binomiale discrète alors que la loi normale est continue, on fait donc **une correction de continuité**

on dit que 3 faces correspondent aux valeurs entre 2.5 et 3.5 de la loi normale

La règle consiste à:
$$\begin{cases} \text{retrancher } 0.5 \text{ à } X & \text{si } X > np \\ \text{ajouter } 0.5 & \text{à } X & \text{si } X < np \end{cases}$$

notre cas $np=6$ et $X=3$ on ajoute 0.5

$$Z = \frac{(3 + 0.5) - 6}{1.73} = -1.4451$$

comme cette valeur est supérieur à -1.96, on accepte H_0

Remarquer: $P(Z \leq -1.445) = 0.0735$ très proche de $P(X \leq 3) = 0.07299$

Exercice 2

Tester l'hypothèse que l'âge d'étudiants des deux échantillons A et B se distribue suivant la même loi où

l'échantillon A: 13; 11; 14ans et B: 10; 16; 15 ans.

Appliquer le test W de Wilcoxon (somme des rangs) et le test U de Mann-Whitney.

Solution:

L'échantillon combiné ordonné(par ordre croissant):

Observations: 10; 11; 13; 14; 15; 16.

Rangs: 1 2 3 4 5 6

Les observations de B sont soulignées, leurs rangs sont: 1,5,6. $W_B = 1 + 5 + 6 = 12$

la somme des rangs de A est: $W_A = 2 + 3 + 4 = 9$

Remarquer que $W_A + W_B = \frac{N(N+1)}{2}$ où $N = n_A + n_B = 3 + 3 = 6$;

$$12 + 9 = \frac{6 \times 7}{2} = 21$$

L'hypothèse à tester est que les échantillons A et B se distribuent suivant la même loi

Soit H_0 l'hypothèse nulle: $F_A = F_B$ i-e: Il n'y-a pas de différence significative entre les distributions A et B (ainsi entre les rang W_A et W_B)

et soit H_1 l'hypothèse alternative: $F_B \leq F_A$ (W_B est grand qui est équivalent à W_A faible car leur somme est constante)

Ainsi H_0 sera rejetée pour W_B grand ou W_A faible

La table peut être utilisée aussi bien pour W_A que pour W_B suivant $\inf(n_A, n_B)$

i) si $n_A < n_B$ la table est relative à W_A et on teste W_A faible

ii) si $n_A > n_B$ la table est relative à W_B et on teste W_B grand

iii) si $n_A = n_B$ la table est relative aussi bien à W_A qu'à W_B

On est dans le iii) cas

1) Relative à W_B

Comparons la valeur de W_B empirique (observée sur l'échantillon) avec la v.a W_B (théorique) sous l'hypothèse H_0

et évaluons la probabilité de l'événement $\{W_B > 12\}$ en regardant la table qui est basée sur la formule

$$P_{H_0}(W_B \leq k) = \frac{\#\{W_B \leq k\}}{\binom{N}{n_b}}$$

La table nous donne les valeurs du numérateur et du dénominateur

Comme la plus petite valeur de W_B est $\frac{n_B(n_B + 1)}{2} = \frac{3 \times 4}{2} = 6$ centrons W_B

soit $W'_B = W_B - 6$

l'événement $\{W_B > 12\} = \{W' > 6\}$

$p_0 = P_{H_0}(W'_B > 6) = 1 - \frac{\#\{W'_B \leq 6\}}{\binom{6}{3}} = 1 - \frac{16}{20} = \frac{4}{20} = 0.2$ on rejette H_0 pour tous les

$\alpha : \alpha \geq p_0$

ainsi pour $\alpha = 0.05$ on accepte H_0

2) Relative à W_A

comme $\{W_B > 12\} = \{21 - W_A > 12\} = \{W_A < 9\}$

et $\{W_A < 9\} = \{W' = W_A - 6 < 9 - 6 = 3\}$

$P_{H_0}(W' < 3) = P_{H_0}(W' \leq 2) = \frac{4}{20} = 0.2$

Même conclusion: Rejet H_0 pour tous les $\alpha : \alpha \geq p_0$

ainsi pour $\alpha = 0.05$ on accepte H_0

Conclusion finale

pour le seuil $\alpha = 0.05$ et tous les seuils ne dépassant pas 0.2 on accepte H_0 : il n'y a pas de différence significative entre les distributions des deux échantillon

Exercice 6

Le tableau suivant donne la résistance de deux câbles d'alliages différents

Alliage I: $A=a_i$	Alliage II: $B=b_i$
18.3 ; 16.4; 22.7; 17.8	12.6; 14.1; 20.5; 10.7; 15.9
18.9; 25.3; 16.1; 24.2	19.6; 12.9; 15.2; 11.8; 14.7

(i) Réaliser les test U et W pour décider s'il y a une différence significative entre ces deux échantillons

(ii) Discuter les différents seuils et déduire le seuil minimal en utilisant l'approche par p

Solution

(i) Les observations: 10.7; 11.8; 12.6; 12.9; 14.1; 14.7; 15.2; 15.9; 16.1; 16.4; 17.8; 18.3; 18.9; 19.6; 20.5; 22.7; 24.2; 25.3.

rangs:		1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18				

Les observations de l'alliage I notées par: $A=a_i$ sont soulignées

$$W_I = 9 + 10 + 11 + 12 + 13 + 16 + 17 + 18 = 106.0$$

$$W_{II} = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 14 + 15 = 65.0$$

$$\text{Soit } N_I = \inf(N_I, N_{II}) = 8, N_{II} = 10$$

$$\text{Evaluons } U = N_I N_{II} + \frac{N_I(N_I + 1)}{2} - W_I = 8 \times 10 + \frac{8 \times 9}{2} - 106 = 10.0$$

$$\text{Cette statistique } U \text{ à pour } \mu_U = \frac{N_I N_{II}}{2} = \frac{8 \times 10}{2} = 40.0$$

$$\text{et pour variance } \sigma_U^2 = \frac{N_I N_{II} (N_I + N_{II} + 1)}{12} = \frac{8 \times 10 \times 19}{12} = 126.67$$

$$\text{d'où } \sigma_U = \sqrt{126.67} = 11.25$$

$$\text{Ainsi la statistique } Z = \frac{U - \mu_U}{\sigma_U} = \frac{10 - 40}{11.25} = -2.6667$$

L'hypothèse H_0 étant de tester qu'il n'y a pas de différence entre les deux alliages au risque

On va réaliser un test bilatérale

La règle de décision est donc : Au risque de $\alpha = 5\%$ Accepter H_0 si

$$-1.96 \leq Z \leq 1.96$$

$$Z \text{ suit la loi } N(0,1) \Rightarrow P(-1.96 \leq Z \leq 1.96) = 0.95$$

sinon rejeter H_0

Comme $Z = -2.6667$ on rejette H_0

On peut utiliser le même raisonnement de l'exercice 1 (l'approche par p)

$$\text{Sous l'hypothèse de } H_0 \text{ évaluons } P(Z \geq -2.667) = 1 - P(Z < -2.667) = 1 - \text{NormalDist}(2.67; 0, 1) = 1 - 0.99621 = 0.00379$$

$$p_0 = 0.00379$$

La regle de décision est de rejeter H_0 pour tous les

$$\text{Rejet de } H_0 : \alpha : \alpha \geq p_0$$

et d'accepter pour tous les

$$\text{Accepter } H_0 \quad \alpha : \alpha < p_0$$

d'où pour le seuil $\alpha = 0.05$ et $\alpha = 0.01$ on rejette H_0

Conclusion finale

les deux alliage ont une différence significative

Remarques: $U = U_{AB}$ peut être calculer par cette méthode

$$\begin{aligned} U_{AB} &= \#\{(a_i, b_i) : a_i < b_i\} \\ &= 5 + 5 = 10 \end{aligned}$$

Autre méthode:

Méthode W: comme $n_I < n_{II}$

Table relative à W_I

On va tester W_I : petit

$$P(W_I \leq 106)$$

$$\begin{aligned} \text{on utilise: } \mu_W &= \frac{N_I(N_I + N_{II} + 1)}{2} = \frac{8 \times 19}{2} = 76.0 \\ \sigma_U^2 &= \frac{N_I N_{II} (N_I + N_{II} + 1)}{12} = \frac{8 \times 10 \times 19}{12} = 126.67 \\ \sigma_W &= 11.25 \end{aligned}$$

$$P(W_I \leq 106) = P\left(\frac{W_I - 76}{11.25} \leq \frac{106 - 76}{11.25}\right) = \Phi\left(\frac{106 - 76}{11.25}\right)$$

$$Z = \frac{106 - 76}{11.25} = 2.6667$$

Au risque de $\alpha = 5\%$ Accepter H_0 si

$$-1.96 \leq Z \leq 1.96$$

comme $z = 2.6667$ on rejette H_0

Remarques:

$$\text{NormalDist}(2.67; 0, 1) = 0.99621$$

$$\text{NormalDist}(1.96; 0, 1) - \text{NormalDist}(-1.96; 0, 1) = 0.95$$

$$\text{rejet de } H_0 \text{ car } p_0 = 1 - 0.99621 = 0.00379 < \alpha = 1 - 0.95 = 0.05$$

Exercice 7

Une entreprise veut acquérir l'une des cinq machines métalliques différentes $M_i : i=1, \dots, 5$. Une expérience est réalisée afin de tester leurs performances. Cinq opérateurs travaillent sur chacune des machines pendant la même laps de temps.

Le tableau suivant montre le nombre de pièces produites par machine.

M_1	68	72	77	42	53
M_2	72	53	63	53	48
M_3	60	82	64	75	72
M_4	48	61	57	64	50
M_5	64	65	70	68	53

(i) Utiliser le test de Kruskal- Wallis avec et sans correction avec l'hypothèse nulle H_0 selon laquelle il n'y a pas de différence entre les machines au seuil $\alpha = 5\%$ et $\alpha = 0.01$, discuter les différents seuils et déduire le seuil minimal en utilisant l'approche par p.

(ii) Utiliser le test Anova à un facteur (test F d'égalités des moyennes)

(iii) Comparer les résultats des deux approches(Paramétrique et NP)

Solution

(i)

Le tableau résumant les rangs

M_1	17.5	21	24	1	6.5	S=70
M_2	21	6.5	12	6.5	2.5	S=48.5
M_3	10	25	14	23	21	93
M_4	2.5	11	9	14	4	40.5
M_5	14	16	19	17.5	6.5	73

k=5; N=25

- Sans corrections

$$\text{La statistique de Kruskal- Wallis: } H = \frac{12}{N(N+1)} \sum_j \frac{W_j^2}{N_j} - 3(N+1)$$

$$H = \frac{12}{25 \times 26} \left(\frac{(70)^2}{5} + \dots + \frac{(73)^2}{5} \right) - 3(25+1) = 6.44$$

pour $k-1=4$, au risque $\alpha = 0.05$

$$\chi^2_{0.95}(4) = \text{ChiSquareInv}(0.95; 4) = 9.4877$$

Comme la valeur $6.44 < 9.4877$ on accepte H_0 au risque 5% et a fortiori à 1%

$$\chi^2_{0.99}(4) = \text{ChiSquareInv}(0.99; 4) = 13.277$$

On peut le seuil minimal de rejet (l'approche par p)

$$\text{Evaluons } p = 1 - \text{ChiSquareDist}(6.44; 4) = 0.16861$$

Donc on accepte H_0 (ou on réserve son jugement) pour tous les risques inférieurs à 0.16861
(a fortiori 0.05 et 0.01)

- Avec correction $1 - \frac{\sum(T^3 - T)}{N^3 - N}$

Observations	48	53	64	68	72	
Nombre d'ex aequo(T)	2	4	3	2	3	
$T^3 - T$	6	60	24	6	24	$\sum(T^3 - T) = 120$

$$1 - \frac{\sum(T^3 - T)}{N^3 - N} = 1 - \frac{120}{(25)^3 - 25} = 0.99231$$

La statistique de Kruskal- Wallis corrigée H_c

$$H_c = \frac{H}{1 - \frac{\sum(T^3 - T)}{N^3 - N}}$$

$$H_c = \frac{6.44}{0.99231} = 6.4899$$

Remarque: Cette correction ne modifie pas la décision

(ii) Anova

Afin d'alléger les calculs, on va soustraire 60 de toutes les observations (les écarts sont invariants par translation)

M ₁	68	72	77	42	53
M ₂	72	53	63	53	48
M ₃	60	82	64	75	72
M ₄	48	61	57	64	50
M ₅	64	65	70	68	53

j/k	1	2	3	4	5	T _j	T _j ²
M ₁	8	12	17	-18	-7	12	144
M ₂	12	-7	3	-7	-12	-11	121
M ₃	0	22	4	15	12	53	2809
M ₄	-12	1	-3	4	-10	-20	400
M ₅	4	5	10	8	-7	20	400
Les sommes:		$\sum X_{jk}^2 = 2658$				54	3874

Rappels

	Traitement j=1	X ₁₁				X _{1b}
Pour un tableau	..					
	...					
	Traitement j=a	X _{a1}				X _{ab}

On définit:

- $\bar{X}_j = \frac{1}{b} \sum_{k=1}^b X_{jk}$ la moyenne des mesures de la j-ième ligne (Traitement j fixé j=1..,a)

Les valeurs \bar{X}_j sont appelées moyennes de groupe, moyennes des traitements ou moyennes des lignes.

- $\bar{X} = \frac{1}{ab} \sum_{j=1}^a \sum_{k=1}^b X_{jk}$ Moyenne totale ou moyenne globale.
- $V = \sum_{j,k} (X_{jk} - \bar{X})^2$ l'écart total noté V_T ou V comme la somme des carrés des écarts de chaque mesure à la moyenne totale \bar{X} .
- En écrivant: $X_{jk} - \bar{X} = (X_{jk} - \bar{X}_{j.}) + (\bar{X}_{j.} - \bar{X})$

$$\begin{aligned} \sum_{j,k} (X_{jk} - \bar{X})^2 &= \sum_{j,k} (X_{jk} - \bar{X}_{j.})^2 + \sum_{j,k} (\bar{X}_{j.} - \bar{X})^2 \\ \sum_{j,k} (X_{jk} - \bar{X})^2 &= \sum_{j,k} (X_{jk} - \bar{X}_{j.})^2 + b \sum_j (\bar{X}_{j.} - \bar{X})^2 \end{aligned}$$

$V_W = \sum_{j,k} (X_{jk} - \bar{X}_{j.})^2$ écart intra traitement noté parfois V_{intra}

$V_B = b \sum_j (\bar{X}_{j.} - \bar{X})^2$ écart entre-inter- traitement (V_{extra} ou V_{inter})

$$V = V_W + V_B$$

Méthodes rapide de calcul des écarts:

$$V_T = \sum X_{jk}^2 - \frac{T^2}{ab}$$

$$V_B = \frac{1}{b} \sum_j T_j^2 - \frac{T^2}{ab}$$

$$V_W = V - V_B$$

avec $T = \sum_j T_j$ et $T_j = \sum_k X_{jk}$

on a $\bar{X}_j = \frac{1}{b} \sum_{k=1}^b X_{jk} = \frac{T_j}{b}$

Modèle Mathématique pour l'analyse de la variance: On peut considérer que chaque ligne du tableau est une v.a de taille b tirée de la population pour le traitement donné. Pour le j-ième traitement, les X_{jk} différeront de la moyenne μ_j de la population d'une erreur aléatoire que l'on note ϵ_{jk} ainsi

$$X_{jk} = \mu_j + \epsilon_{jk} \quad *$$

On suppose que ces erreurs sont de loi $N(0, \sigma^2)$

Remarquer: égalité des variances pour les erreurs

si μ est la moyenne de la population pour tous les traitements et si on pose $\alpha_j = \mu - \mu_j \Leftrightarrow \mu_j = \alpha_j + \mu$ L'équation * devient

$$X_{jk} = \alpha_j + \mu + \epsilon_{jk} \quad **$$

où $\sum_j \alpha_j = 0$

Sous l'hypothèse que les

Variance	degré de liberté	Ecart	F
$V_B = b \sum_j (\bar{X}_j - \bar{X})^2 (inter)$	a-1	$\hat{S}_B^2 = \frac{V_B}{a-1}$	$\frac{\hat{S}_B^2}{\hat{S}_W^2}$
$V_W = V - V_B$ (intra)	a(b-1)	$\hat{S}_W^2 = \frac{V_W}{a(b-1)}$	avec a-1 et b-1 d° de liberté
$V = V_W + V_B$	ab-1		

Application

$$V_T = \sum X_{jk}^2 - \frac{T^2}{ab}$$

$$V = 2658 - \frac{(54)^2}{5 \times 5}$$

$$V_B = \frac{3874}{5} - \frac{(54)^2}{5 \times 5}$$

Variance	degré de liberté	Ecart	F
$V_B = b \sum_j (\bar{X}_j - \bar{X})^2 = 658.2$	a-1=4	$\hat{S}_B^2 = \frac{V_B}{a-1} = \frac{658.2}{4} = 164.5$	$\frac{\hat{S}_B^2}{\hat{S}_W^2} = \frac{164.55}{94.16} = 1.75$
$V_W = V - V_B$ (intra)=1883.2	a(b-1)=5×4=20	$\hat{S}_W^2 = \frac{V_W}{a(b-1)} = \frac{1883.2}{20} = 94.16$	avec a-1=4 et a
$V = V_W + V_B = 2514.4$	ab-1=5×5 - 1 = 24		

$$F_{0.95}(4, 20) = \text{FInv}(0.95; 4, 20) = 2.8661$$

Anisi on accepte H_0 car $1.75 < 2.8661$

d'autre part:

$$1 - \text{FDist}(1.75; 4, 20) = 0.17868$$

La valeur de $p = 0.17868$ est la plus petite valeur du seuil de signification à laquelle H_0 serait rejetée

donc H_0 ne serait pas rejetée au seuil 0.01 ou 0.05.

(iii) Les deux approches ont les mêmes résultats
