L3-S6 : Statistique Inférentielle

Examen 2ème Session 2024-2025

Durée 1h 30

<u>Recommandations</u> : *Soigner la rédaction. La note prendra largement en compte la qualité des explications. La copie-brouillon et la copie qui ne comporte que des résultats sont mal perçues par le correcteur. Les exercices sont indépendants, des extraits de tables statistiques sont donnés en annexe.* ***Aucun document n'est permis. Les machines à calculer non programmables sont autorisées. Les dictionnaires pour les étudiants étrangers sont autorisés.***

Exercice 1 : **(Barème de notation : a) 1.5 pt b) 1.5 pt c) 1.5 pt d) 2 pts e) 2 pts = 8.5 pts)**

Une institution bancaire envisage de modifier les heures d'ouvertures de ses agences dans les départements du Rhône, de la Savoie et du Cantal de la région Auvergne-Rhône-Alpes. Toutefois, la direction ne veut apporter aucune modification à l'horaire actuel sans avoir sondé au préalable une partie représentative de sa clientèle. Pour cela, des clients de chaque département desservi par l'institution bancaire ont été interrogés. La répartition est présentée dans le tableau suivant :

Changement d'horaire	Rhône	Savoie	Cantal
Favorable	304	237	219
Défavorable	96	63	81
Total	400	300	300

a) Peut-on affirmer avec un risque d'erreur $\alpha = 5\%$ que plus de 75% des clients du département du Rhône sont en faveur du changement horaire ?

b) Peut-on conclure avec un risque d'erreur $\alpha = 5\%$ que les clients du département de la Savoie sont plus favorables au changement horaire que ceux du département du Rhône ?

c) Estimer par intervalle de confiance et ceci pour l'ensemble des trois départements, la proportion de clients favorables au changement d'horaire avec un niveau de confiance de 95%.

d) Peut-on conclure avec un risque d'erreur $\alpha = 5\%$ que les proportions de clients en faveur d'un changement d'horaire sont identiques dans les trois départements.

On s'intéresse aux clients favorables au changement d'horaire dans les trois départements. Ci-dessous les résultats obtenus dans les trois départements :

Changement d'horaire	Rhône	Savoie	Cantal
Favorable	304	237	219

e) Peut-on conclure, au seuil de signification $\alpha = 5\%$, que les clients favorables au changement d'horaire se répartissent de façon uniforme dans les trois départements ?

*************** °°° ***************

Exercice 2 : **(Barème de notation : a) 1.5 pt b) 1.5 pt c) 1.5 pt d) 1.5 pt e) 1.5 pt f) 2 pts g) 2 pts = 11.5 pts)**

Dans 2 exploitations agricoles, on a comparer les rendements journaliers de 2 races bovines : la Montbéliarde et la Normande. La production de lait par jour de la race Normande est supposée normalement distribuée. Les résultats sont consignées dans le tableau suivant :

Races bovines	Monbéliarde	Normande
Nombre de vaches	75	25
Production journalière moyenne observée par vache (litre/jour)	21	25
SCE : Somme des Carrés des Ecarts à la moyenne	1550	900

a) Peut-on considérer avec un risque d'erreur $\alpha = 5\%$, l'hypothèse selon laquelle la production journalière moyenne des vache Normandes est significativement inférieure à 28 litres de lait ?

b) Peut-on conclure avec un risque d'erreur $\alpha = 5\%$ que la variance de la production journalière des vaches Normandes est significativement supérieure à 35 litres^2 ?

c) Peut-on affirmer avec un risque d'erreur $\alpha = 5\%$, l'hypothèse selon laquelle les variances des productions journalières de lait de ces deux races sont différentes ?

d) Doit-on admettre, avec un risque d'erreur $\alpha = 5\%$, que la vache Normande produit plus de lait par jour que la vache Monbéliarde ?

Les services de contrôle des 2 exploitations agricoles ont établi, avec les mêmes données, l'intervalle de confiance bilatéral symétrique $[-6.197\ ;\ -1.803]$ de la différence des productions journalières de lait.

e) Quel est le niveau de confiance que l'on peut associer à cet intervalle de confiance ?

On souhaite comparer les productions journalières de lait de 10 vaches de chaque race bovine. Les résultats sont donnés dans le tableau suivant :

Vache	1	2	3	4	5	6	7	8	9	10
Monbéliarde	26	23	21	24	25	22	24	21	26	21
Normande	28	25	24	25	23	26	27	25	28	28

f) Peut-on conclure, au seuil de signification $\alpha = 5\%$, que les deux distributions des productions journalières de lait de ces 2 races sont différentes ?

On a relevé les productions journalières de lait de 8 vaches Normandes en Hiver et en Eté.

Vache	1	2	3	4	5	6	7	8
Eté	26	25	21	21	22	21	24	21
Hiver	28	25	24	25	23	26	27	25

g) Peut-on conclure avec un risque d'erreur $\alpha = 5\%$, que les distributions des productions journalières en été et en hiver sont différentes ?

*************** °°° ***************

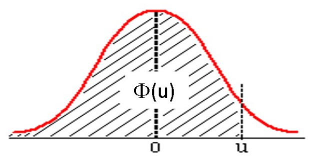

Table de la loi Normale Centrée Réduite : $U \to N(0;1)$

Fonction de répartition : Φ

$$\Phi(u) = P(U \leq u) \quad ; \quad \Phi(-u) = P(U \leq -u) = 1 - \Phi(u)$$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900

Exemples : $\Phi(1.26) = P(U \leq 1.26) = 0.89617$; $\Phi(u) = P(U \leq u) = 97.50\% \Rightarrow u = 1.96$

*************** °°° ***************

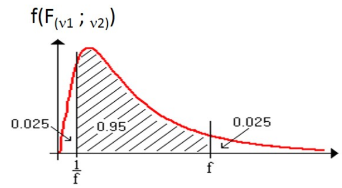

Table de la loi de Fisher-Snedecor

Valeur f de la variable de Fisher-Snedecor $F(\nu_1;\nu_2)$ ayant la probabilité 2.5% d'être dépassée.

ν_1 : degrés de liberté du numérateur ν_2 : degrés de liberté du dénominateur

$$F(f) = P(F(\nu_1,\nu_2) \leq f) = 97.50\%$$

ν_2 ν_1	24	25	45	50	74	75
24	2,269	2,257	2,124	2,107	2,054	2,052
25	2,242	2,230	2,096	2,079	2,025	2,024
45	1,965	1,952	1,807	1,788	1,728	1,726
50	1,931	1,919	1,772	1,752	1,691	1,689
74	1,835	1,822	1,669	1,648	1,583	1,581
75	1,833	1,819	1,666	1,645	1,580	1,578

Exemples : $\nu_1 = 45$ d.d.l. et $\nu_2 = 50$ d.d.l. $P(F_{97,5\%\,;45\,;\,50} \leq f) = 0,975 \Rightarrow f = 1,772$

$$P(F_{2,5\%\,;\,45\,;\,50} \leq f') = 0,025$$

$$P(F_{97,5\%\,;\,50\,;\,45} \leq f) = 0,975 \Rightarrow f = 1,788 \Rightarrow f' = \frac{1}{f} = \frac{1}{1,788} = 0.5592$$

*************** °°° ***************

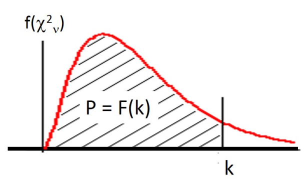

Table de la loi du $\chi^2_{\nu\ d.d.l.}$ Fonction de répartition : $F(k) = P(\chi^2_\nu \leq k)$

Exemples : $\nu = 3$ d.d.l. $F(k) = P(\chi^2_3 \leq k) = 0.95 \Rightarrow k = 7.815$

$k = 9.348 \Rightarrow F(9.348) = P(\chi^2_3 \leq 9.348) = 0.975$

ν P	0.010	0.020	0.025	0.050	0.100	0.150	0.200	0.800	0.900	0.950	0.975	0.980	0.990
1	0.000	0.001	0.001	0.004	0.016	0.036	0.064	1.642	2.706	3.841	5.024	5.412	6.64
2	0.020	0.040	0.051	0.103	0.211	0.325	0.446	3.219	4.605	5.991	7.378	7.824	9.21
3	0.115	0.185	0.216	0.352	0.584	0.798	1.005	4.642	6.251	7.815	9.348	9.837	11.35
24	10,856	11,992	12,401	13,848	15,659	16,969	18,062	29,553	33,196	36,415	39,364	40,270	42,980
25	11,524	12,697	13,120	14,611	16,473	17,818	18,940	30,675	34,382	37,652	40,646	41,566	44,314

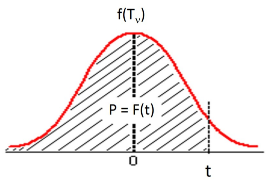

Table de la loi de Student - Fractiles de la loi de Student à ν degrés de liberté

Valeur du fractile t - Fonction de répartition : $P = F(t) = P(T_\nu \leq t)$.

ν \ P	0,9350	0,9400	0,9450	0,9500	0,9550	0,9600	0,9650	0,9700	0,9750	0,980	0,985	0,9900	0,9950
24	1,568	1,612	1,660	1,711	1,767	1,828	1,896	1,974	2,064	2,172	2,307	2,492	2,797
25	1,566	1,610	1,657	1,708	1,764	1,825	1,893	1,970	2,060	2,167	2,301	2,485	2,787
40	1,546	1,589	1,635	1,684	1,737	1,796	1,862	1,936	2,021	2,123	2,250	2,423	2,704
75	1,531	1,573	1,617	1,665	1,718	1,775	1,838	1,910	1,992	2,090	2,212	2,377	2,643
98	1,527	1,568	1,613	1,661	1,712	1,769	1,832	1,903	1,984	2,081	2,202	2,365	2,627
100	1,527	1,568	1,613	1,660	1,712	1,769	1,832	1,902	1,984	2,081	2,201	2,364	2,626

Exemple : $\nu = 40$ d.d.l. $\quad P(T_{40} \leq t) = 0.975 \Rightarrow t = +2.021$ et $P(T_{40} \leq t) = 0.025 \Rightarrow t = -2.021$

Valeurs critiques du T de Wilcoxon - Echantillons appariés.

α	5%	2.5%	1.0%	0.5%
α^*	10%	5%	2.0%	1.0%
n				
6	2	0		
7	2	2	0	
8	5	3	2	0
9	8	5	3	1
1	10	8	5	3
11	13	10	7	5
12	17	13	10	97
13	21	17	13	90
14	25	21	16	12

n : nombre de différences non nulles

α : Niveau de signification (test unilatéral)

α^* : Niveau de signification (test bilatéral)

Hypothèses statistiques : Test bilatéral symétrique

H_0 : les échantillons ont des distributions identiques

H_1 : les échantillons ont des distributions différentes.

Exemple : n = 12 , $\alpha^* = 5\%$ Test bilatéral

Rejet de H_0 si : $T = min(T^+, T^-) \leq T_{\alpha^*} = 13$

Tables de Wilcoxon - Echantillons indépendants

Test bilatéral | **Tests Unilatéraux**

m	$\alpha = 5\%$	$\alpha = 1\%$		m	$\alpha = 5\%$	$\alpha = 1\%$
6	0	*	$\widetilde{m}$: la médiane de $X - Y$.	6	2	*
7	2	*	W_X la somme des rangs des différences positives.	7	2	*
8	3	0		8	5	*
9	5	1	Valeur critique $w*$ tel que : $P(W_X \leq w*)$.	9	8	2
10	8	3	m : taille de l'échantillon de X, le plus petit échantillon.	10	10	4
11	10	5	n : taille de l'échantillon de Y, $(m \leq n)$.	11	13	7
12	13	9	$*$ indique que le test ne peut être significatif.	12	17	9

Test unilatéral "risque à gauche"

$\begin{cases} H_0 : \widetilde{m} = 0 \\ H_1 : \widetilde{m} < 0. \end{cases}$

Rejet de H_0 si $W_x \leq w*$

Test bilatéral

$\begin{cases} H_0 : \widetilde{m} = 0 \\ H_1 : \widetilde{m} \neq 0. \end{cases}$

Rejet de H_0 si $W_x \leq w*$

Test unilatéral "risque à droite"

$\begin{cases} H_0 : \widetilde{m} = 0 \\ H_1 : \widetilde{m} > 0. \end{cases}$

Rejet de H_0 si $W_x > w*$

*************** °°° ***************