Université Grenoble Alpes
L2 - DLST - UGA **Année 2023**

STA401 - Examen terminal - SUJET A (INM) - 1ère session - Correction

---

Exercice :

1.

	Espérance de X	Moyenne empirique	Moyenne estimée de X	Variance de X	Variance empirique	Variance estimée de X
Valeurs	20	19,8	19,8	121	0,96	1,0667

2. $P(20-a<X<20+a) \;=\; P(-\dfrac{a}{4}<\dfrac{X-20}{4}<\dfrac{a}{4})=0,98 \iff P(\dfrac{X-20}{4}<\dfrac{a}{4})=0,99 \iff$ $\dfrac{a}{4}=2,3263 \iff a=9,3052$

3. L'intervalle $]a;b[$ voulu est l'intervalle de confiance de μ au niveau de confiance 0,99 avec σ^2 inconnue : $P(\mu\in I)=1-\alpha \iff I=\left[\overline{x}\pm t^{n-1}_{1-\alpha/2}\dfrac{s}{\sqrt{n-1}}\right]$ avec $t^{n-1}_{1-\alpha/2}=2,7195$. Donc $[a;b]\simeq[18,64025;21,35975]$

4. $qt(0.995,19)$

5. Rejet de $\mathcal{H}_0 \Leftrightarrow T<-t^{n-1}_{1-\alpha/2}\; ou\; T>t^{n-1}_{1-\alpha/2}$ avec $T\rightsquigarrow\mathcal{T}_{29}$
   Or, $T_{calc}=2,21$ donc $p_{valeur}=2*P_{H_0}(T>2.21)\simeq 0,03515$. On accepte $\mathcal{H}_1$ pour tous les risques supérieurs à 3,515%

6. 2*(1-pt(2.21,29)) ou bien 2*pt(2.21,29, lower.tail = FALSE)

Problème :

PARTIE A

1. Calculs : $\hat{\mu}=\overline{x}=3.1063636$ et $\widehat{\sigma^2}=s'^2=0.19494988^2=0.0380$

2. a) Intervalle de confiance de la variance avec $\alpha=0,05$: $\left[\dfrac{ns^2}{z^{(n-1)}_{1-\alpha/2}};\dfrac{ns^2}{z^{(n-1)}_{\alpha/2}}\right]\approx\left[\dfrac{0.38}{20.48};\dfrac{0.38}{3.25}\right]\approx[0.01855;0.11692]$
   L'intervalle de l'écart-type est donc : $[0.13621;0.3419]$
   b) Les paramètres μ et σ^2 du modèle sont inconnues, donc l'intervalle de confiance de μ au niveau de 99% est : $\left]\overline{x}\pm t^{n-1}_{1-\alpha/2}\dfrac{s'}{\sqrt{n}}\right[$. On trouve: $]2.9201;3.2927[$ avec $t=3.1693$
   c) Pour avoir une précision de l'intervalle de $\pm0,1$, il faut que $t^{(n-1)}_{1-\alpha/2}\dfrac{s'}{\sqrt{n}}=0,1\Rightarrow$
   $t^{10}_{1-\alpha/2}=0,1*\sqrt{11}/0,0.19494977=1,70127$. Lecture de table (p presque 95%); à la calculatrice : $p=1-\alpha/2\simeq0,94013$, donc $\alpha\simeq0,11974$. Il faut donc un niveau de confiance de 0,88026.

3. Test : $\begin{cases}\mathcal{H}_0: \sigma^2=0,05\\ \mathcal{H}_1: \sigma^2\neq0,05\end{cases}$ Test paramétrique de la variance bilatéral.
   Sous $\mathcal{H}_0$ la statistique est : $T=\dfrac{nS^2}{0,05}$ suit la loi $\mathcal{X}^2_{\mathrm{n}-1}$; Rejet de $\mathcal{H}_0 \Leftrightarrow T<z^{n-1}_{\alpha/2}\quad ou\; T>z^{n-1}_{1-\alpha/2}$
   $z^{n-1}_{\alpha/2}=2.156\quad z^{n-1}_{1-\alpha/2}=25.19$ et $T_{calc}=7,6$. On ne peut donc pas rejeter $\mathcal{H}_0$ au seuil de 1% (on conclut donc que $\sigma^2=0,05$)

4. Test : $\begin{cases} \mathcal{H}_0 : \mu = 3 \\ \mathcal{H}_1 : \mu > 3 \end{cases}$ On considère que la variance est inconnue.

   Sous $\mathcal{H}_0$ la statistique est : $T = \dfrac{\overline{X} - 3}{S'/\sqrt{n}}$. Elle suit une loi $\mathcal{T}_{10}$; Rejet de $\mathcal{H}_0 \Leftrightarrow T > t_{1-\alpha}^{n-1}$

   Pour $\alpha = 1\%$, on obtient : { Rejet de $\mathcal{H}_0 \iff T > 2.7638$ }. Or, $T_{calc} = 1.8078$. On ne peut donc pas rejeter $\mathcal{H}_0$ au seuil de 1% (on conclura que $\mu = 3$)

5. $p_{valeur} = P_{H_0}(T > 1.8078) \simeq 0.05038$. En conséquence, pour tous les risques $\alpha > p_{valeur} = 5,038$, on accepterait $\mathcal{H}_1$: la durée de vie moyenne des composants est supérieure à 3, au risque αde se tromper.

PARTIE B

1. Test comparaison de 2 échantillons indépendants (2 séries de composants différents), même variable. Condition : Normalité des variables (donné dans l'énoncé). Test sur les moyennes unilatéral (échant 1 > échant 2 - le traitement augmente X) : $\{\mathcal{H}_0 : \mu_1 = \mu_2 \}$ contre $\{\mathcal{H}_1 : \mu_1 > \mu_2 \}$

   Calculs : $\overline{x}_1 = 3.10636$, $s_1^{'2} = 0.0380$; $\overline{x}_2 = 1.67545$, $s_2^{'2} = 0.143247$

2. a) Test : $\{\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2 \}$ contre $\{\mathcal{H}_1 : \sigma_1^2 \neq \sigma_2^2 \}$. Calcul de la statistique (attention au plus grand !) : $T_{calc} = \dfrac{s_2^{'2}}{s_1^{'2}} = 3.76966$; lecture loi Fisher : $f_{(0,995)}^{(10;10)} = 5.85$; Rejet de $\mathcal{H}_1$. On accepte l'égalité des variances au seuil de 1%.

   b) $p_{valeur} = 2 * P_{\mathcal{H}_0}(T > 3.76966) = 0,04773$ (calculatrice avec la loi de Fisher). On conclut qu'il faut prendre des risques α supérieurs à 4,77% pour déclarer que les variances sont différentes. Sinon, à 1% par exemple, on conclura à l'égalité, et on pourra faire le test des moyennes suivant.

3. Test : $\{\mathcal{H}_0 : \mu_1 = \mu_2 \}$ contre $\{\mathcal{H}_1 : \mu_1 > \mu_2 \}$. (l'échantillon traité a une moyenne supérieure à celui non traité ?)

   Ici, variances égales (test avant), n petits (11), donc $T = \sqrt{\dfrac{n_1 + n_2 - 2}{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{n_1 S_1^2 + n_2 S_2^2}} \rightsquigarrow \mathcal{T}_{n_1+n_2-2}$

   Calculs : $T_{calc} = 11.14736$; T suit la loi $\mathcal{T}_{20}$

   $p_{valeur} = P_{\mathcal{H}_0}(T > T_{calc}) = P(T > 11.14736) = 2,46.10^{-10}$

   Pour $\alpha = 1\%$ (donc $\alpha > p_{valeur}$), on accepte $\mathcal{H}_1$. On conclut que la durée de vie des composants avec traitement est supérieure avec un risque de 1%. Le traitement augmente donc la durée de vie moyenne des composants pour un risque de 1% mais aussi quelque soit le risque α supérieur à 2,46.10^{-10} (mais inférieur à 4.77%). Il faudrait prendre des risques $\alpha < p_{valeur}$ pour ne pas accepter $\mathcal{H}_1$. Avec des risques quasiment nul de se tromper, on conclura à l'efficacité du traitement qui augmente la durée de vie significativement.

PARTIE C

a) On dispose de deux échantillons appariés car ce sont les 11 mêmes individus (composants) : X_3 traités, et X_2 non traités. On veut comparer la moyenne des différences, et savoir si le traitement est efficace (la durée de vie augmente avec le traitement) [donc (éch.3)> (éch.2)].

On suppose que $D = X_3 - X_2 \rightsquigarrow \mathcal{N}(\mu_d, \sigma_d^2)$.

b) Le test est : $\begin{cases} \mathcal{H}_0 : \mu_3 = \mu_2 \\ \mathcal{H}_1 : \mu_3 > \mu_2 \end{cases} \iff \begin{cases} \mathcal{H}_0 : \mu_d = 0 \\ \mathcal{H}_1 : \mu_d > 0 \end{cases}$ La statistique est $T = \dfrac{\overline{D}\sqrt{n-1}}{S_d} \rightsquigarrow \mathcal{T}_{n-1}$

Calculs : $\overline{d} = 1,353636$, $s_d = 0.266706$, $T_{calc} = 16,04978$.

Calcul de la $p_{valeur} = P_{\mathcal{H}_0}(T > 16,04978) = 9,11.10^{-9}$ [table stat donne la majoration : $p_{valeur} < 0,0005$]

Pour tous les risques $\alpha > p_{valeur}$ (par exemple 0,1%, 1% ou 5%), on accepterait $\mathcal{H}_1$. On conclut donc que le traitement est efficace et augmente la durée de vie des composants, avec des risques faibles de se tromper.