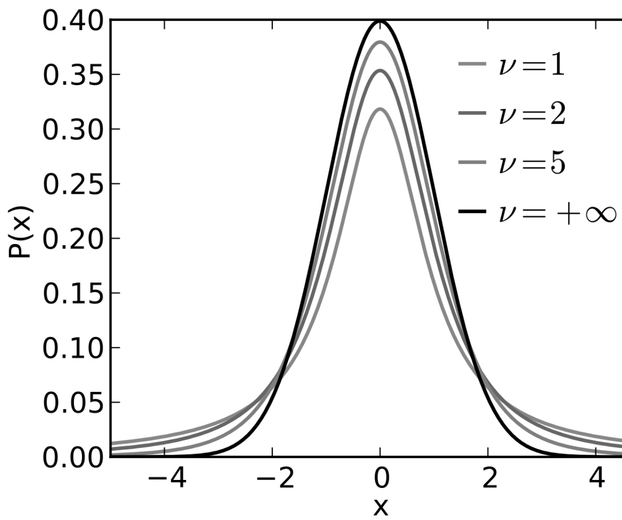
M3.5.1 Statistiques

Corrigé TD3



v0.1

18 juillet 2017

Exercice 1. Dans une chaine de production particulière, chaque boite pèse en moyenne 16 g. Le sous–ou sur-remplissage est un problème sérieux et la chaine de production sera fermée si l'un des deux cas se produit. Grâce aux données passées, on sait que σ est égal à 0.8 g. Un contrôleur de la qualité échantillonne 30 boites toutes les deux heures et décide de fermer ou non la chaine de production pour ajustement.

(a) Quelle est la règle de rejet pour un seuil de signification égal à 0.05 ?

(b) Si la moyenne d'un échantillon $\bar{x}$ est égale à 16.32 g, quelle action recommanderiez-vous ?

(c) Si la moyenne d'échantillon est égale à 15.82 g, quelle action préconiseriez-vous ?

Exercice 2. Une opération particulière dans la chaine de montage automobile nécessite, en moyenne, 2.2 minutes. À cause des effets du temps de montage d'une pièce à la fois sur les opérations d'assemblage précédentes et suivantes, il est important de conserver un temps moyen d'assemblage de 2.2 minutes. Un échantillon aléatoire de 45 pièces assemblées révèle un temps moyen de montage de 2.39 minutes, avec un écart-type d'échantillon de 0.20 minute. Utiliser un seuil de signification égal à 0.02 et tester si l'opération requière en moyenne 2.2 minutes pour être exécutée.

Exercice 3. L'entreprise Simtech fabrique des tubes de verre pour l'entreprise Giscom. L'entreprise cliente exige que les lots expédiés par Simtech contiennent au plus 2% de défectueux. Les lots sont habituellement constitués de 5000 tubes de verre. Avant d'expédier les lots, Simtech effectue un contrôle en prélevant au hasard 200 tubes.

(a) En utilisant un risque de $\alpha = 0.05$ de rejeter à tort un lot dont la proportion de défectueux est de 2% (ou mieux), quelle est la valeur critique de la proportion de défectueux dans un échantillon de taille $n = 200$ qui ne doit pas être dépassée pour considérer un lot comme acceptable ?

(b) On doit expédier un lot de 5000 tubes. Lors du contrôle final, on a observé 4 tubes de verre défectueux dans un échantillon de 200 tubes. Est-ce que ce lot peut être considéré comme acceptable d'après les exigences de Giscom ?

(c) Giscom réceptionne des lots de Simtech en prélevant également 200 tubes et fait usage de la même règle de décision que Simtech. Quelles sont les chances sur 100 d'accepter un lot comportant 6% de tubes défectueux alors que Simtech certifie 2% de défectueux dans 95% des cas ? Comment appelle-t-on ce risque ?

Exercice 4. Un fabricant de fournitures électriques fabrique des résistances dont la valeur nominale doit être de 1000 ohms (ohm-mètre). Pour vérifier le procédé de fabrication, on prélève un échantillon aléatoire de 64 résistances. On mesure ces résistances avec un ohm-mètre de précision et les calculs de la moyenne et l'écart-type conduisent aux valeurs suivantes :

$$\bar{x} = 990 \text{ ohm}, \quad s = 100 \text{ ohm}$$

(a) En supposant que le risque de première espèce est fixé à 0.05, élaborer une règle de décision qui permettrait de tester l'hypothèse nulle selon laquelle le procédé est centré à 1000 ohm.

(b) Est-ce que l'hypothèse de normalité de la population (l'ensemble des valeurs ohmiques de la fabrication) est requise pour l'élaboration de la règle de décision en (a) ?

(c) Selon les résultats de l'échantillon de 64 résistances, doit-on rejeter l'hypothèse selon laquelle le procédé est centré à 1000 ohm ?

(d) Quelle est la probabilité de ne pas rejeter l'hypothèse nulle H_0 : "$m = 1000$ ohm" si le procédé opère en réalité à $m = 1050$ ohm ? Comment appelle-t-on ce risque ?

Exercice 5. La résistance d'un composant électronique doit être, en moyenne, de 400 ohm. Un échantillon de 16 composants, prélevé d'un grand lot, conduit aux résultats suivants :

392	396	386	389
388	387	403	397
401	391	400	402
394	406	406	400

On considère que la distribution de la résistance est celle d'une loi normale.

(a) Peut-on considérer, au seuil de signification $\alpha = 0.05$, que le lot respecte la norme de 400 ohm ?

(b) Avec les résultats de cet échantillon, calculer un intervalle de confiance ayant un niveau de confiance de 95% de contenir la vraie moyenne. Est-ce que cet intervalle contient la norme spécifiée ?

Exercice 6. Une entreprise fabrique des dispositifs électroniques dont la durée de vie moyenne est de 800 heures. La durée de vie des dispositifs est distribuée normalement avec un écart-type $\sigma = 50$ heures. Pour vérifier la qualité des dispositifs, un échantillon de 25 dispositifs est soumis à un essai de fiabilité et on adopte la règle de décision suivante :

Les dispositifs sont de qualité inacceptable si la durée de vie moyenne de 25 dispositifs est inférieure à 783.55 heures ; on les considère de qualité acceptable si la durée de vie moyenne est supérieure ou égale à 783.55 heures.

(a) Quelles sont les hypothèses statistiques que l'on veut tester avec cette règle de décision ?

(b) Quelle est la probabilité de rejeter à tort un lot de dispositifs de qualité acceptable ?

(c) Quel est le risque de deuxième espèce pour chacune des valeurs suivantes de μ : 750, 760, 770, 780, 800.

(d) Le responsable du contrôle a décidé de modifier son plan de contrôle en prélevant 36 dispositifs de la production (au lieu de 25). Quelle règle de décision doit-il alors adopter pour tester les hypothèses spécifiques en (a) ?

Exercice 1. Dans une chaine de production particulière, chaque boite pèse en moyenne 16 g. Le sous–ou sur-remplissage est un problème sérieux et la chaine de production sera fermée si l'un des deux cas se produit. Grâce aux données passées, on sait que σ est égal à 0.8 g. Un contrôleur de la qualité échantillonne 30 boites toutes les deux heures et décide de fermer ou non la chaine de production pour ajustement.

(a) Quelle est la règle de rejet pour un seuil de signification égal à 0.05 ?

Solution.

On note $\mu_0 = 16$. On souhaite tester H_0 : "$\mu = \mu_0$" contre H_1 : "$\mu \neq \mu_0$".
Pour le seuil fixé $\alpha = 0.05$, la région critique du test est

$$W = \{(X_1, \dots, X_n); |\bar{X} - \mu_0| > l_\alpha\}$$

avec l_α réel.
On a la fonction puissance du test

$$\begin{aligned} \beta : \mathbb{R} &\to [0, 1] \\ \mu &\mapsto \beta(\mu) = P((X_1, \dots, X_n) \in W; \mu) \end{aligned}$$

($\beta(\mu)$ est la probabilité de rejeter H_0 quand la vraie valeur de la moyenne est μ.)
Le seuil du test α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned} \alpha &= \sup_{H_0} \beta(\mu) \\ &= \sup_{H_0} P((X_1, \dots, X_n) \in W; \mu) \\ &= \sup_{\mu=\mu_0} P(|\bar{X} - \mu_0| > l_\alpha; \mu) \\ &= P(|\bar{X} - \mu_0| > l_\alpha; \mu_0) \\ &= P(|\sqrt{n}\frac{\bar{X} - \mu_0}{\sigma}| > \sqrt{n}\frac{l_\alpha}{\sigma}; \mu_0) \end{aligned}$$

On pose

$$U = \sqrt{n}\frac{\bar{X} - \mu_0}{\sigma}$$

On suppose que les X_i sont indépendantes et identiquement distribuées, et on a $n = 30$. Donc avec le TCL, U suit approximativement la loi normale centrée réduite.
On a donc

$$\begin{aligned} \alpha &= P(|U| > \sqrt{n}\frac{l_\alpha}{\sigma}; \mu_0) \\ \iff 1 - \frac{\alpha}{2} &= P(U < \sqrt{n}\frac{l_\alpha}{\sigma}; \mu_0) \\ \iff 1 - \frac{\alpha}{2} &= \Phi(\sqrt{n}\frac{l_\alpha}{\sigma}) \end{aligned}$$

À partir de la table, on trouve la valeur $u_\alpha = 1.96$ correspondant à $1 - \frac{\alpha}{2} = 0.975$.

On a donc $\sqrt{n}\frac{l_\alpha}{\sigma} = u_\alpha$, soit $l_\alpha = \frac{\sigma}{\sqrt{n}} u_\alpha$.
A.N. : $l_\alpha = \frac{0.8}{\sqrt{30}} \times 1.96 = 0.286$.
Donc la règle de rejet pour un seuil de signification de 0.05 est :

$$\boxed{(X_1, \ldots, X_n) \in W \iff |\bar{X} - \mu_0| > 0.286}$$

(b) Si la moyenne d'un échantillon $\bar{x}$ est égale à 16.32 g, quelle action recommanderiez-vous ?

Solution.

On a
$$|\bar{x} - \mu_0| = |16.32 - 16| = 0.32$$
$|\bar{x} - \mu_0| > 0.286$, donc $(x_1, \ldots, x_n) \in W$, et alors on rejette H_0.
La chaine de production doit donc être fermée pour ajustement.

(c) Si la moyenne d'échantillon est égale à 15.82 g, quelle action préconiseriez-vous ?

Solution.

On a
$$|\bar{x} - \mu_0| = |15.82 - 16| = 0.18$$
$|\bar{x} - \mu_0| \leq 0.286$, donc $(x_1, \ldots, x_n) \notin W$, et alors on ne rejette pas H_0.
La chaine de production n'a donc pas besoin d'être fermée pour ajustement.

Exercice 2. Une opération particulière dans la chaine de montage automobile nécessite, en moyenne, 2.2 minutes. À cause des effets du temps de montage d'une pièce à la fois sur les opérations d'assemblage précédentes et suivantes, il est important de conserver un temps moyen d'assemblage de 2.2 minutes. Un échantillon aléatoire de 45 pièces assemblées révèle un temps moyen de montage de 2.39 minutes, avec un écart-type d'échantillon de 0.20 minute. Utiliser un seuil de signification égal à 0.02 et tester si l'opération requière en moyenne 2.2 minutes pour être exécutée.

Solution.

On note $\mu_0 = 2.2$. On souhaite tester H_0 : "$\mu = \mu_0$" contre H_1 : "$\mu \neq \mu_0$".
Pour le seuil fixé $\alpha = 0.02$, la région critique du test est

$$W = \{(X_1, \ldots, X_n); |\bar{X} - \mu_0| > l_\alpha\}$$

avec l_α réel.

On a la fonction puissance du test :

$$\begin{aligned}\beta : \mathbb{R} &\to [0,1]\\ \mu &\mapsto \beta(\mu) = P((X_1,\dots,X_n)\in W;\mu)\end{aligned}$$

($\beta(\mu)$ est la probabilité de rejeter H_0 quand la vraie valeur de la moyenne est μ.)
Le seuil du test α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned}\alpha &= \sup_{H_0}\beta(\mu)\\ &= \sup_{H_0} P((X_1,\dots,X_n)\in W;\mu)\\ &= \sup_{\mu=\mu_0} P(|\bar{X}-\mu_0|>l_\alpha;\mu)\\ &= P(|\bar{X}-\mu_0|>l_\alpha;\mu_0)\\ &= P(\left|\sqrt{n-1}\frac{\bar{X}-\mu_0}{S}\right| > \sqrt{n-1}\frac{l_\alpha}{S};\mu_0)\end{aligned}$$

On pose

$$T = \sqrt{n-1}\frac{\bar{X}-\mu_0}{S}$$

On suppose que les X_i sont indépendantes et identiquement distribuées, et on a $n = 45 \geq 30$. Donc avec le TCL et le théorème de Fisher, T suit approximativement la loi de Student à $n-1$ degrés de liberté.
On a donc

$$\alpha = P(|T| > \sqrt{n-1}\frac{l_\alpha}{S};\mu_0)$$
$$\iff 1-\alpha = P(|T| \leq \sqrt{n-1}\frac{l_\alpha}{S};\mu_0)$$

À partir de la table, pour trouver la valeur $t_{n-1,\alpha}$ correspondant à $n-1 = 44$ et $1-\alpha = 0.98$, on fait une interpolation linéaire : $\frac{t_{n-1,\alpha}-2.4233}{44-40} = \frac{2.4033-2.4233}{50-40}$ qui donne $t_{n-1,\alpha} = 2.4153$.
On a donc $\sqrt{n-1}\frac{l_\alpha}{s} = t_{n-1,\alpha}$, soit $l_\alpha = \frac{s}{\sqrt{n-1}}t_{n-1,\alpha}$.
A.N. : $l_\alpha = \frac{0.2}{\sqrt{44}} \times 2.4153 = 0.0728$.
Donc la règle de rejet du test au seuil de 0.02 est :

$$(X_1,\dots,X_n)\in W \iff |\bar{X}-\mu_0| > 0.0728$$

On a

$$|\bar{x}-\mu_0| = |2.39-2.2| = 0.19$$

$|\bar{x}-\mu_0| > 0.0728$, donc $(x_1,\dots,x_n)\in W$, et alors on rejette H_0.
On conclut que l'opération ne requière pas en moyenne 2.2 minutes pour être exécutée.

Exercice 3. L'entreprise Simtech fabrique des tubes de verre pour l'entreprise Giscom. L'entreprise cliente exige que les lots expédiés par Simtech contiennent au plus 2% de défectueux. Les lots sont habituellement constitués de 5000 tubes de verre. Avant d'expédier les lots, Simtech effectue un contrôle en prélevant au hasard 200 tubes.

(a) En utilisant un risque de $\alpha = 0.05$ de rejeter à tort un lot dont la proportion de défectueux est de 2% (ou mieux), quelle est la valeur critique de la proportion de défectueux dans un échantillon de taille $n = 200$ qui ne doit pas être dépassée pour considérer un lot comme acceptable ?

Solution.

On note $p_0 = 0.02$. On souhaite tester H_0 : "$p \leq p_0$" contre H_1 : "$p > p_0$".
Pour le seuil fixé $\alpha = 0.05$, la région critique du test est

$$W = \{(X_1, \dots, X_n); \hat{p} > l_\alpha\}$$

avec $l_\alpha \geq p_0$, et $\hat{p} = \bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$.
On a la fonction puissance du test :

$$\begin{aligned} \beta : \mathbb{R} &\to [0, 1] \\ p &\mapsto \beta(p) = P((X_1, \dots, X_n) \in W; p) \end{aligned}$$

($\beta(p)$ est la probabilité de rejeter H_0 quand la vraie valeur de la proportion est p.)
Le seuil du test α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned} \alpha &= \sup_{H_0} \beta(p) \\ &= \sup_{H_0} P((X_1, \dots, X_n) \in W; p) \\ &= \sup_{p \leq p_0} P(\hat{p} > l_\alpha; p) \\ &= \sup_{p \leq p_0} P(\sqrt{n}\frac{\hat{p} - p}{\sqrt{p(1-p)}} > \sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p) \end{aligned}$$

On pose

$$U = \sqrt{n}\frac{\hat{p} - p}{\sqrt{p(1-p)}}$$

On suppose que les X_i sont indépendantes et identiquement distribuées, et on a $n = 200 \geq 30$. Donc avec le TCL, U suit approximativement la loi normale centrée réduite.

On a donc

$$\begin{aligned}
&\alpha = \sup_{p\le p_0} P(U > \sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p) \\
\iff &\alpha = \sup_{p\le p_0} \left(1 - P(U \le \sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p)\right) \\
\iff &\alpha = \sup_{p\le p_0} \left(1 - \Phi(\sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p)\right) \\
\iff &\alpha = 1 - \inf_{p\le p_0} \Phi(\sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p)
\end{aligned}$$

La fonction de répartition de la loi normale centrée réduite $\Phi(u)$ est bien sûr croissante en u, donc $\Phi(\sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}})$ est décroissante en p, et alors

$$\inf_{p\le p_0} \Phi(\sqrt{n}\frac{l_\alpha - p}{\sqrt{p(1-p)}}; p) = \Phi(\sqrt{n}\frac{l_\alpha - p_0}{\sqrt{p_0(1-p_0)}}; p_0)$$

On a donc

$$1 - \alpha = \Phi(\sqrt{n}\frac{l_\alpha - p_0}{\sqrt{p_0(1-p_0)}}; p_0)$$

À partir de la table, on trouve la valeur $u_\alpha = 1.645$ correspondant à $1 - \alpha = 0.95$.
On a donc $\sqrt{n}\frac{l_\alpha - p_0}{\sqrt{p_0(1-p_0)}} = u_\alpha$, soit $l_\alpha = \sqrt{\frac{p_0(1-p_0)}{n}}u_\alpha + p_0$.
A.N. : $l_\alpha = \sqrt{\frac{0.02\times 0.98}{200}} \times 1.645 + 0.02 = 0.0363$.
Donc la valeur critique de la proportion de défectueux qui ne doit pas être dépassée pour considérer un lot comme acceptable (c'est-à-dire ne pas rejeter H_0) est $l_\alpha = 3.63\%$.

(b) On doit expédier un lot de 5000 tubes. Lors du contrôle final, on a observé 4 tubes de verre défectueux dans un échantillon de 200 tubes. Est-ce que ce lot peut être considéré comme acceptable d'après les exigences de Giscom ?

Solution.

La proportion de défectueux dans cet échantillon est $\frac{4}{200} = 0.02$.
Puisque $0.02 < 0.0363$, ce lot peut être considéré comme acceptable.

(c) Giscom réceptionne des lots de Simtech en prélevant également 200 tubes et fait usage de la même règle de décision que Simtech. Quelles sont les chances sur 100 d'accepter un lot comportant 6% de tubes défectueux alors que Simtech certifie 2% de défectueux dans 95% des cas ? Comment appelle-t-on ce risque ?

Solution.

On cherche $1-\beta(p_1)$ avec $p_1 = 0.06$, c'est-à-dire la probabilité de décider que $p \leq p_0$ (ne pas rejeter H_0) alors qu'en réalité $p = p_1$.

$$\begin{aligned}
1-\beta(p_1) &= 1 - P((X_1,\dots,X_n)\in W; p_1)\\
&= P((X_1,\dots,X_n)\notin W; p_1)\\
&= P(p \leq l_\alpha; p_1)\\
&= P(\sqrt{n}\frac{p-p_1}{\sqrt{p_1(1-p_1)}} \leq \sqrt{n}\frac{l_\alpha - p_1}{\sqrt{p_1(1-p_1)}}; p_1)\\
&= P(Z \leq \sqrt{n}\frac{l_\alpha - p_1}{\sqrt{p_1(1-p_1)}}; p_1)\\
&= 1 - P(Z \leq \sqrt{n}\frac{p_1 - l_\alpha}{\sqrt{p_1(1-p_1)}}; p_1)
\end{aligned}$$

en posant

$$Z = \sqrt{n}\frac{p-p_1}{\sqrt{p_1(1-p_1)}}$$

On a $n = 200 \geq 30$ et les X_i sont supposées indépendantes et identiquement distribuées, donc avec le TCL, Z suit approximativement la loi normale centrée réduite. Alors

$$1-\beta(p_1) = 1 - \Phi(\sqrt{n}\frac{p_1 - l_\alpha}{\sqrt{p_1(1-p_1)}})$$

A.N. :

$$\sqrt{n}\frac{p_1 - l_\alpha}{\sqrt{p_1(1-p_1)}} = \sqrt{200}\frac{0.06-0.363}{\sqrt{0.06\times 0.94}} = 1.41$$

$$\Phi(\sqrt{n}\frac{p_1 - l_\alpha}{\sqrt{p_1(1-p_1)}}) = \Phi(1.41) = 0.9207$$

donc

$$1-\beta(p_1) = 1 - 0.9207$$
$$\boxed{1-\beta(\mu_1) = 0.0793}$$

Si le lot comporte en réalité 6% de défectueux, alors la probabilité de ne pas rejeter H_0 (c'est-à-dire décider que le lot est acceptable alors qu'il ne l'est pas) est de 7.93%. Ce risque est appelé risque de deuxième espèce.

Exercice 4. Un fabricant de fournitures électriques fabrique des résistances dont la valeur nominale doit être de 1000 ohms (ohm-mètre). Pour vérifier le procédé de fabrication, on prélève un échantillon aléatoire de 64 résistances. On mesure ces résistances avec un ohm-mètre de précision et les calculs de la moyenne et l'écart-type conduisent aux valeurs suivantes :

$$\bar{x} = 990 \text{ ohm}, \quad s = 100 \text{ ohm}$$

(a) En supposant que le risque de première espèce est fixé à 0.05, élaborer une règle de décision qui permettrait de tester l'hypothèse nulle selon laquelle le procédé est centré à 1000 ohm.

Solution.

On note $m_0 = 1000$. On souhaite tester H_0 : "$m = m_0$" contre H_1 : "$m \neq m_0$".
Pour le seuil fixé $\alpha = 0.05$, la région critique du test est

$$W = \{(X_1, \ldots, X_n); |\bar{X} - m_0| > l_\alpha\}$$

avec l_α réel.
On a la fonction puissance du test :

$$\begin{aligned} \beta : \mathbb{R} &\to [0, 1] \\ \mu &\mapsto \beta(m) = P((X_1, \ldots, X_n) \in W; m) \end{aligned}$$

($\beta(m)$ est la probabilité de rejeter H_0 quand la vraie valeur de la moyenne est m.)
Le seuil du test α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned} \alpha &= \sup_{H_0} \beta(m) \\ &= \sup_{H_0} P((X_1, \ldots, X_n) \in W; m) \\ &= \sup_{m=m_0} P(|\bar{X} - m_0| > l_\alpha; m) \\ &= P(|\bar{X} - m_0| > l_\alpha; m_0) \\ &= P(\left|\sqrt{n-1}\frac{\bar{X} - m_0}{S}\right| > \sqrt{n-1}\frac{l_\alpha}{S}; m_0) \end{aligned}$$

On pose

$$T = \sqrt{n-1}\frac{\bar{X} - m_0}{S}$$

On suppose que les X_i sont indépendantes et identiquement distribuées, et on a $n = 64 \geq 30$. Donc avec le TCL et le théorème de Fisher, T suit approximativement la loi de Student à $n - 1$ degrés de liberté.
On a donc

$$\begin{aligned} \alpha &= P(|T| > \sqrt{n-1}\frac{l_\alpha}{S}; m_0) \\ \iff 1 - \alpha &= P(|T| \leq \sqrt{n-1}\frac{l_\alpha}{S}; m_0) \end{aligned}$$

À partir de la table, pour trouver la valeur $t_{n-1,\alpha}$ correspondant à $n - 1 = 63$ et $1 - \alpha = 0.95$, on fait une interpolation linéaire : $\frac{t_{n-1,\alpha} - 2.0003}{63-60} = \frac{1.9944 - 2.0003}{70-60}$ qui donne $t_{n-1,\alpha} = 1.9985$.

On a donc $\sqrt{n-1}\frac{l_\alpha}{s} = t_{n-1,\alpha}$, soit $l_\alpha = \frac{s}{\sqrt{n-1}} t_{n-1,\alpha}$.
A.N. : $l_\alpha = \frac{100}{\sqrt{63}} \times 1.9985 = 25.18$.
Donc la règle de rejet du test au seuil de 0.05 est :

$$\boxed{(X_1, \dots, X_n) \in W \iff |\bar{X} - m_0| > 25.18}$$

(b) Est-ce que l'hypothèse de normalité de la population (l'ensemble des valeurs ohmiques de la fabrication) est requise pour l'élaboration de la règle de décision en (a) ?

Solution.

L'hypothèse de normalité de la population n'est pas requise, puisqu'on est dans les conditions de l'application du TCL ($n = 64 \geq 30$).
Toutefois, l'hypothèse est recommandée car elle apporterait plus de précision.

(c) Selon les résultats de l'échantillon de 64 résistances, doit-on rejeter l'hypothèse selon laquelle le procédé est centré à 1000 ohm ?

Solution.

On a
$$|\bar{x} - m_0| = |990 - 1000| = 10$$
$|\bar{x} - m_0| \leq 25.18$, donc $(x_1, \dots, x_n) \notin W$, et alors on ne rejette pas H_0.
On conclut alors que le procédé est centré à 1000 ohm.

(d) Quelle est la probabilité de ne pas rejeter l'hypothèse nulle H_0 : "$m = 1000$ ohm" si le procédé opère en réalité à $m = 1050$ ohm ? Comment appelle-t-on ce risque ?

Solution.

On cherche $1 - \beta(\mu_1)$ avec $\mu_1 = 1050$, c'est-à-dire la probabilité de décider que $\mu = \mu_0$ (ne pas rejeter H_0) alors qu'en réalité $\mu = \mu_1$.

$$\begin{aligned}
1 - \beta(m_1) &= 1 - P((X_1, \dots, X_n) \in W; m_1) \\
&= P((X_1, \dots, X_n) \notin W; m_1) \\
&= P(|\bar{X} - m_0| \leq l_\alpha; m_1) \\
&= P(m_0 - l_\alpha \leq \bar{X} \leq m_0 + l_\alpha; m_1) \\
&= P(\sqrt{n-1}\frac{m_0 - m_1 - l_\alpha}{S} \leq \sqrt{n-1}\,\frac{\bar{X} - m_1}{S} \leq \sqrt{n-1}\frac{m_0 - m_1 + l_\alpha}{S}; m_1) \\
&= P(t_1 \leq Z \leq t_2; m_1)
\end{aligned}$$

en posant

$$
\begin{aligned}
Z &= \sqrt{n-1}\,\frac{\bar{X}-m_1}{S} \\
t_1 &= \sqrt{n-1}\frac{m_0-m_1-l_\alpha}{S} \\
t_2 &= \sqrt{n-1}\frac{m_0-m_1+l_\alpha}{S}
\end{aligned}
$$

On a $n = 64 \geq 30$ et les X_i sont supposées indépendantes et identiquement distribuées, donc avec le TCL et le théorème de Fisher, Z suit approximativement la loi de Student à $n-1$ degrés de liberté.
Alors

$$1-\beta(m_1) = F_{\mathrm{St}(n-1)}(t_2) - F_{\mathrm{St}(n-1)}(t_1)$$

A.N. :

$$
\begin{aligned}
t_1 &= \sqrt{63}\frac{1000-1050-25.18}{100} = -5.967 \\
t_2 &= \sqrt{63}\frac{1000-1050+25.18}{100} = -1.970
\end{aligned}
$$

donc

$$
\begin{aligned}
F_{\mathrm{St}(n-1)}(t_1) &= F_{\mathrm{St}(63)}(-5.967) \simeq 0 \\
F_{\mathrm{St}(n-1)}(t_2) &= F_{\mathrm{St}(63)}(-1.970) = 0.0266
\end{aligned}
$$

donc

$$1-\beta(\mu_1) = 1-0.0266$$

$$\boxed{1-\beta(\mu_1) = 0.9734}$$

Si le procédé est en réalité centré à 1050 ohm, la probabilité de ne pas rejeter H_0 (c'est-à-dire décider qu'il est centré à 1000 ohm) est de 97.34%.
Ce risque est appelé risque de deuxième espèce.

Exercice 5. La résistance d'un composant électronique doit être, en moyenne, de 400 ohm. Un échantillon de 16 composants, prélevé d'un grand lot, conduit aux résultats suivants :

392	396	386	389
388	387	403	397
401	391	400	402
394	406	406	400

On considère que la distribution de la résistance est celle d'une loi normale.

(a) Peut-on considérer, au seuil de signification $\alpha = 0.05$, que le lot respecte la norme de 400 ohm ?

Solution.

On calcule tout d'abord la moyenne et l'écart-type de l'échantillon :

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{6338}{16} = 396.125$$

$$s^2 = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \bar{x}^2 = \frac{2511322}{16} - 396.125^2 = 42.61 \text{ donc } s = \sqrt{s^2} = 6.528$$

On note $\mu_0 = 400$. On souhaite tester H_0 : "$\mu = \mu_0$" contre H_1 : "$\mu \neq \mu_0$".
Pour le seuil fixé $\alpha = 0.05$, la région critique du test est

$$W = \{(X_1, \ldots, X_n); |\bar{X} - \mu_0| > l_\alpha\}$$

avec l_α réel.
On a la fonction puissance du test :

$$\begin{aligned} \beta : \mathbb{R} &\to [0,1] \\ \mu &\mapsto \beta(\mu) = P((X_1, \ldots, X_n) \in W; \mu) \end{aligned}$$

($\beta(\mu)$ est la probabilité de rejeter H_0 quand la vraie valeur de la moyenne est μ.)
Le seuil du test α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned} \alpha &= \sup_{H_0} \beta(\mu) \\ &= \sup_{H_0} P((X_1, \ldots, X_n) \in W; \mu) \\ &= \sup_{\mu=\mu_0} P(|\bar{X} - \mu_0| > l_\alpha; \mu) \\ &= P(|\bar{X} - \mu_0| > l_\alpha; \mu_0) \\ &= P(\left|\sqrt{n-1}\frac{\bar{X} - \mu_0}{S}\right| > \sqrt{n-1}\frac{l_\alpha}{S}; \mu_0) \end{aligned}$$

On pose

$$T = \sqrt{n-1}\frac{\bar{X} - \mu_0}{S}$$

On suppose que les X_i sont indépendantes. Avec le théorème de Fisher, T suit la loi de Student à $n-1$ degrés de liberté.
On a donc

$$\begin{aligned} \alpha &= P(|T| > \sqrt{n-1}\frac{l_\alpha}{S}; \mu_0) \\ \iff 1 - \alpha &= P(|T| \leq \sqrt{n-1}\frac{l_\alpha}{S}; \mu_0) \end{aligned}$$

À partir de la table, on trouve la valeur $t_{n-1,\alpha} = 2.1315$ correspondant à $n - 1 = 15$ et $1 - \alpha = 0.95$.

On a donc $\sqrt{n-1}\frac{l_\alpha}{s} = t_{n-1,\alpha}$, soit $l_\alpha = \frac{s}{\sqrt{n-1}} t_{n-1,\alpha}$.
A.N. : $l_\alpha = \frac{6.528}{\sqrt{15}} \times 2.1315 = 3.592$.
Donc la règle de rejet du test au seuil de 0.05 est :

$$(X_1, \ldots, X_n) \in W \iff |\bar{X} - \mu_0| > 3.592$$

On a

$$|\bar{x} - \mu_0| = |396.125 - 400| = 3.875$$

$|\bar{x} - \mu_0| > 3.592$, donc $(x_1, \ldots, x_n) \in W$, et alors on rejette H_0.
On conclut que le lot ne respecte pas la norme de 400 ohm.

(b) Avec les résultats de cet échantillon, calculer un intervalle de confiance ayant un niveau de confiance de 95% de contenir la vraie moyenne. Est-ce que cet intervalle contient la norme spécifiée ?

Solution.

On a

$$\alpha = P(|\bar{X} - \mu_0| > l_\alpha; \mu_0)$$
$$\iff 1 - \alpha = P(|\bar{X} - \mu_0| \leq l_\alpha; \mu_0)$$

donc un intervalle de confiance ayant un niveau de confiance de 95% pour μ est

$$\boxed{[\bar{x} - l_\alpha, \bar{x} + l_\alpha] = [392.53, 399.72]}$$

Cet intervalle de confiance ne contient pas la norme spécifiée de 400 ohm.

Exercice 6. Une entreprise fabrique des dispositifs électroniques dont la durée de vie moyenne est de 800 heures. La durée de vie des dispositifs est distribuée normalement avec un écart-type $\sigma = 50$ heures. Pour vérifier la qualité des dispositifs, un échantillon de 25 dispositifs est soumis à un essai de fiabilité et on adopte la règle de décision suivante :

Les dispositifs sont de qualité inacceptable si la durée de vie moyenne de 25 dispositifs est inférieure à 783.55 heures ; on les considère de qualité acceptable si la durée de vie moyenne est supérieure ou égale à 783.55 heures.

(a) Quelles sont les hypothèses statistiques que l'on veut tester avec cette règle de décision ?

Solution.

On note $\mu_0 = 800$. Les hypothèses statistiques que l'on souhaite tester avec cette règle de décision sont H_0 : "$\mu \geq \mu_0$" contre H_1 : "$\mu < \mu_0$".

(b) Quelle est la probabilité de rejeter à tort un lot de dispositifs de qualité acceptable ?

Solution.

On cherche le seuil du test α.
La région critique du test est

$$W = \{(X_1, \dots, X_n); \bar{X} < l_\alpha\}$$

avec $l_\alpha = 783.55$ selon l'énoncé.
On a la fonction puissance du test

$$\begin{aligned}\beta : \mathbb{R} &\to [0,1] \\ \mu &\mapsto \beta(\mu) = P((X_1, \dots, X_n) \in W; \mu)\end{aligned}$$

($\beta(\mu)$ est la probabilité de rejeter H_0 quand la vraie valeur de la moyenne est μ.)
α est la probabilité maximale de rejeter H_0 alors que H_0 est vraie :

$$\begin{aligned}\alpha &= \sup_{H_0} \beta(\mu) \\ &= \sup_{H_0} P((X_1, \dots, X_n) \in W; \mu) \\ &= \sup_{\mu \geq \mu_0} P(\bar{X} < l_\alpha; \mu) \\ &= \sup_{\mu \geq \mu_0} P(\sqrt{n}\frac{\bar{X} - \mu}{\sigma} < \sqrt{n}\frac{l_\alpha - \mu}{\sigma}; \mu) \\ &= \sup_{\mu \geq \mu_0} P(U < \sqrt{n}\frac{l_\alpha - \mu}{\sigma}; \mu)\end{aligned}$$

en posant

$$U = \sqrt{n}\frac{\bar{X} - \mu}{\sigma}$$

On suppose que les X_i sont indépendantes. U suit donc la loi normale centrée réduite.
On a donc

$$\alpha = \sup_{\mu \geq \mu_0} \Phi(\sqrt{n}\frac{l_\alpha - \mu}{\sigma}; \mu)$$

$\Phi(x)$ est bien sûr croissante en x, donc $\Phi(\sqrt{n}\frac{l_\alpha - \mu}{\sigma})$ est décroissante en μ.
On a donc

$$\sup_{\mu \geq \mu_0} \Phi(\sqrt{n}\frac{l_\alpha - \mu}{\sigma}; \mu) = \Phi(\sqrt{n}\frac{l_\alpha - \mu_0}{\sigma}; \mu_0)$$

et alors

$$\begin{aligned}\alpha &= \Phi(\sqrt{n}\frac{l_\alpha - \mu_0}{\sigma}; \mu_0) \\ &= 1 - \Phi(\sqrt{n}\frac{\mu_0 - l_\alpha}{\sigma}; \mu_0)\end{aligned}$$

A.N. :

$$\sqrt{n}\frac{\mu_0 - l_\alpha}{\sigma} = \sqrt{25}\frac{800 - 783.55}{50} = 1.645$$

À partir de la table, on trouve

$$\Phi(1.645) = 0.95$$

donc

$$\alpha = 1 - 0.95$$

$$\boxed{\alpha = 0.05}$$

(c) Quel est le risque de deuxième espèce pour chacune des valeurs suivantes de μ : 750, 760, 770, 780, 800.

Solution.

Le risque de deuxième espèce est $1 - \beta(\mu_1)$, c'est-à-dire la probabilité de décider que $\mu \geq \mu_0$ (ne pas rejeter H_0) alors qu'en réalité $\mu = \mu_1$.

$$\begin{aligned}
1 - \beta(\mu_1) &= 1 - P((X_1, \ldots, X_n) \in W; \mu_1) \\
&= P((X_1, \ldots, X_n) \notin W; \mu_1) \\
&= P(\bar{X} \geq l_\alpha; \mu_1) \\
&= P(\sqrt{n}\frac{\bar{X} - \mu_1}{\sigma} \geq \sqrt{n}\frac{l_\alpha - \mu_1}{\sigma}; \mu_1) \\
&= 1 - P(Z \leq \sqrt{n}\frac{l_\alpha - \mu_1}{\sigma}; \mu_1)
\end{aligned}$$

en posant

$$Z = \sqrt{n}\frac{\bar{X} - \mu_1}{\sigma}$$

Les X_i sont supposées indépendantes, donc Z suit la loi normale centrée réduite. Alors

$$\boxed{1 - \beta(\mu_1) = 1 - \Phi(\sqrt{n}\frac{l_\alpha - \mu_1}{\sigma})}$$

A.N. :

$$\begin{aligned}
1 - \beta(750) &= 1 - \Phi(\sqrt{25}\frac{783.55 - 750}{50}) = 1 - \Phi(3.355) = 1 - 0.9996 = 0.0004 \\
1 - \beta(760) &= 1 - \Phi(\sqrt{25}\frac{783.55 - 760}{50}) = 1 - \Phi(2.355) = 1 - 0.9907 = 0.0093 \\
1 - \beta(770) &= 1 - \Phi(\sqrt{25}\frac{783.55 - 770}{50}) = 1 - \Phi(1.355) = 1 - 0.9123 = 0.0877 \\
1 - \beta(780) &= 1 - \Phi(\sqrt{25}\frac{783.55 - 780}{50}) = 1 - \Phi(0.355) = 1 - 0.6387 = 0.3613 \\
1 - \beta(800) &= 1 - \Phi(\sqrt{25}\frac{783.55 - 800}{50}) = 1 - \Phi(-1.645) = \Phi(1.645) = 0.95
\end{aligned}$$

(Pour la dernière valeur, on la sait déjà sans calcul.)

(d) Le responsable du contrôle a décidé de modifier son plan de contrôle en prélevant 36 dispositifs de la production (au lieu de 25). Quelle règle de décision doit-il alors adopter pour tester les hypothèses spécifiques en (a) ?

Solution.

Avec $n = 36$, et avec le même seuil $\alpha = 0.05$, on a

$$1 - \alpha = \Phi(\sqrt{n}\frac{\mu_0 - l_\alpha}{\sigma}; \mu_0)$$

À partir de la table, on trouve la valeur $u_\alpha = 1.645$ correspondant à $1 - \alpha = 0.95$.
On a donc $\sqrt{n}\frac{\mu_0 - l_\alpha}{\sigma} = u_\alpha$, soit $l_\alpha = \mu_0 - \frac{\sigma}{\sqrt{n}}u_\alpha$.
A.N. : $l_\alpha = 800 - \frac{50}{\sqrt{36}} \times 1.645 = 786.29$.
Donc la règle de décision au seuil 0.05 pour la nouvelle taille de l'échantillon est

$$\boxed{(X_1, \ldots, X_n) \in W \iff \bar{X} < 786.29}$$