TD N°4 : Tests d'hypothèses sur une caractéristique usuelle

<u>Exercice 1</u>

Une machine produit des billes de roulement de diamètre fixe. Si elle fonctionne normalement, il y a une proportion de 5% de billes défectueuses. Si elle est déréglée, la proportion de billes défectueuses est de 10%.

Avant d'envoyer une commande à son client, il teste un lot de 500 billes. Soit X le variable aléatoire représentant le nombre de billes défectueuses sur ce lot. On sait que X suit une loi binomiale b(500,p) où p=0.05 si la machine fonctionne correctement et p=0.1 si la machine est déréglée.

1. Il décide de ne pas livrer son client si la machine est déréglée, c'est-à-dire si le nombre de billes défectueuses est supérieur ou égal à 50 (=500×10%).
   (a) Exprimer les hypothèses H_0 et H_1 ainsi que les risques de $1^{ère}$ et $2^{ème}$ espèces.
   (b) Calculer le risque de $1^{ère}$ espèce.
   (c) Calculer la puissance du test et en déduire si un client doit acheter ou non un lot ayant subi un tel test.
2. Le fabricant re-définit son test en imposant un risque de $1^{ère}$ espèce de 1%.
   (a) Déterminer la région critique du test.
   (b) Calculer sa puissance.
   (c) Enoncer les règles de décision avec les erreurs associées.

- Fonctionnement normal : 5% de billes défectueuses
- Fonctionnement anormal : 10% de billes défectueuses

<u>Variable de décision</u>

Soit X le nombre de billes défectueuses sur un lot de 500 billes. X suit une loi b(n,p). Or n=500 est assez grand pour approcher la loi de X grâce au TCL. On considère donc que X suit une $N(\mu,\sigma^2)$ avec $\mu=np$ et $\sigma^2=np(1-p)$. La valeur de p=0.05 ($\mu=25$) si la machine fonctionne correctement et p=0.1 ($\mu=50$) si elle est déréglée.

1) Test « naif » du fabricant.

(a) <u>Les hypothèses</u>

Etant donné que la machine est déréglée quand il y a 10% de billes défectueuses, le fabricant décide d'appliquer la règle de décision toute simple qui consiste à ne pas livrer s'il y a plus de 10% de billes défectueuses dans le lot, *i.e.* plus de 50 billes défectueuse. Nous avons donc le test suivant :

H_0 : la machine fonctionne correctement (p=0.05)
H_1 : la machine est déréglée (p=0.1)

avec la règle de décision :

- Si $X\geq 50$ j'accepte l'hypothèse H_1, *i.e* je ne livre pas la marchandise

- Si $X<50$ je garde l'hypothèse H_0, *i.e*, je livre la marchandise

Les risques (erreurs) encourus sont

– Risque de 1ère espèce : Accepter H_1 alors que H_0 est vraie $\Leftrightarrow$ ne pas livrer de la marchandise de bonne qualité

– Risque de 2ème espèce : Accepter H_0 alors que H_1 est vraie $\Leftrightarrow$ livrer de la marchandise de mauvaise qualité (perdre client)

(b) <u>Risque de 1ère espèce</u>

On sait que $\alpha=P(\text{Accepter } H_1 \mid H_0 \text{ vraie})=P(X\geq 50 \mid H_0 \text{ vraie})$.

Supposons H_0 vraie alors $p=0.05$ donc X suit une $N(0.05n,\sigma_0^2)$ où $\sigma_0^2=0.95*0.05*n$, d'où

$$\alpha=P(X\geq 50)=P\left(\frac{X-0.05n}{\sqrt{0.95*0.05*n}}\geq\frac{50-0.05n}{\sqrt{0.95*0.05*n}}\sqrt{n}\right)=P(Z\geq 5.13)=1.7*10^{-7}$$

où Z suit une loi $N(0,1)$.

Le fabricant a donc $10^{-5}\%$ de risque de ne pas livrer de la bonne marchandise.

(c) <u>Calcul de la puissance</u>

On sait que $1-\beta=P(\text{Accepter } H_1 \mid H_1 \text{ vraie})=P(X\geq 50 \mid H_0 \text{ vraie})$.

Supposons H_1 vraie alors $p=0.1$ donc X suit une $N(0.1n,\sigma_1^2)$ où $\sigma^2=0.9*0.1*n$, d'où

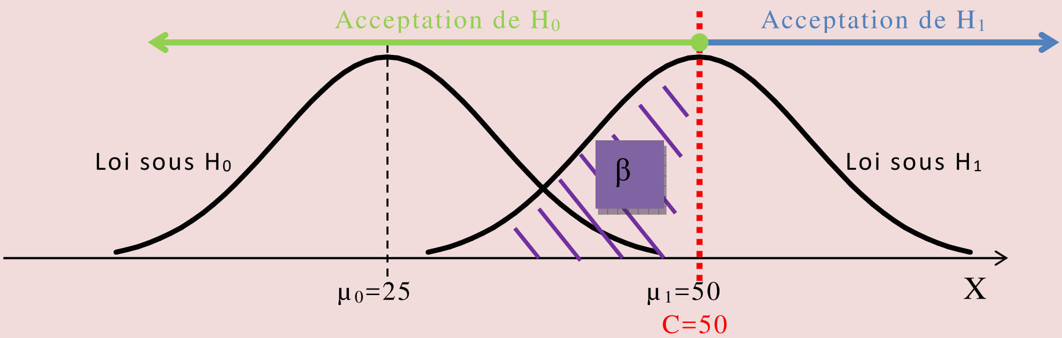
$$1-\beta=P(X\geq 50)=P\left(\frac{X-0.1*n}{\sqrt{0.9*0.1*n}}\geq\frac{50-0.1*n}{\sqrt{0.9*0.1*n}}\right)=P(Z\geq 0)=0.5$$

$\Rightarrow \beta=0.5$

Le fabricant a une chance sur deux de livrer de la marchandise de mauvaise qualité !

Cette règle de décision n'est pas du tout à l'avantage du client et risque de faire perdre sa clientèle au fabricant.

Que se passe-t'il ?



Il suffit de décaler le seuil de décision vers la gauche pour faire diminuer β.

2) Test fabricant

On teste toujours les hypothèses

H_0 : la machine fonctionne correctement ($p=0.05$)

H_1 : la machine est déréglée ($p=0.1$)

mais cette fois la règle de décision n'est pas fixée.

(a) <u>Calcul de la région critique</u>

La région critique W est la région d'acceptation de H_1, donc ici $W=\{X\geq C\}$

On sait que $\alpha=P(W \mid H_0 \text{ vraie})=P(X\geq C \mid H_0 \text{ vraie})$.

Supposons H_0 vraie alors p=0.05 donc X suit une $N(0.05n,\sigma_0^2)$ où $\sigma_0^2=0.95*0.05*n$, d'où

$$\alpha=P(X\geq C)=P(\frac{X-0.05*n}{\sqrt{0.95*0.05*n}}\geq\frac{C-0.05*n}{\sqrt{0.95*0.05*n}})$$

$\Leftrightarrow$ $P(Z\geq C')=0.01$ où Z suit une loi N(0,1)

$\Rightarrow$ $C'=2.33$ $\Rightarrow$ $C=0.05n+2.33*\sqrt{0.95*0.05*n}=36.25$.

La règle de décision est maintanent la suivante :

- Si $X\geq 36.25 \Leftrightarrow X\geq 37$ alors il livre
- Si $X<36.25 \Leftrightarrow X\leq 36$ il ne livre pas

(b) Calcul de la puissance

On sait que $1-\beta=P(W \mid H_1 \text{ vraie})=P(X\geq 37 \mid H_1 \text{ vraie})$.

Supposons H_1 vraie alors p=0.1 donc X suit une $N(0.1n,\sigma_1^2)$ où $\sigma^2=0.9*0.1*n$, d'où

$$1-\beta=P(X\geq 37)=P(\frac{X-0.1*n}{\sqrt{0.9*0.1*n}}\geq\frac{37-0.1*n}{\sqrt{0.9*0.1*n}})=P(Z\geq -1.94)=P(Z<1.94)=0.9738$$

$\Rightarrow$ $\beta=0.026$

(c) Règles de décision

- Si $X\geq 37$ alors il ne livre pas la marchandise (accepte H_1) avec 1% de risque qu'elle soit de bonne qualité (si H_0 est vraie)
- Si $X\leq 36$ alors il livre la marchandise (garde H_0) avec 2.6% de risque qu'elle soit de mauvaise qualité (si H_1 est vraie).

Cette règle semble plus raisonnable pour le client.

Exercice 2

Les habitants d'une région aéroportuaire se plaignent que le bruit des avions dépasse la limite autorisée de 80 décibels en moyenne imposée par la législation. On admet que l'intensité du bruit causé par les avions est une variable aléatoire X de loi gaussienne d'espérance μ et de variance 64.

On mesure un échantillon journalier de n=16 variables aléatoires indépendantes $X_1, \ldots, X_n$ de l'intensité du bruit, et on effectue le test statistique suivant.

$$\begin{cases} H_0 : \mu=\mu_0=80 \text{ décibels} \\ H_1 : \mu=\mu_1=85 \text{ décibels} \end{cases}$$

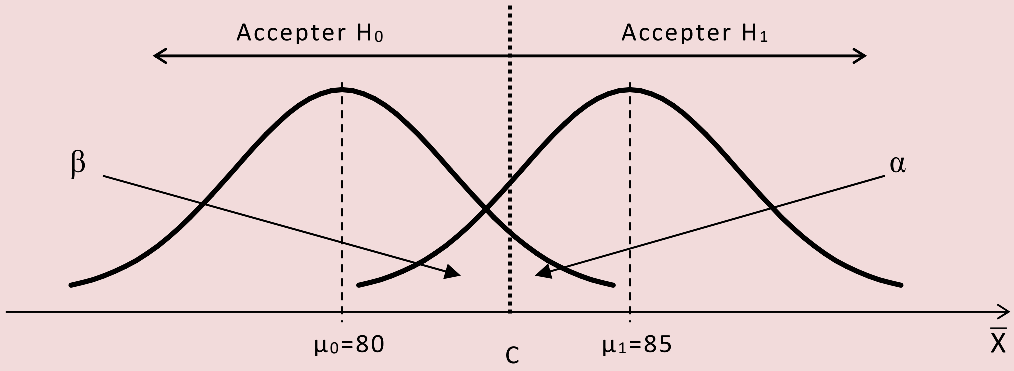
1) Expliciter les risques de première et deuxième espèces. De quel point de vue est fait ce test ? Celui des habitant ou celui des responsables de l'aéroport ?
2) Quelle variable de décision faut-il choisir et quelle est sa loi ?
3) Déterminer graphiquement l'allure de la région critique et représenter sur le graphique les erreurs de première et deuxième espèces.
4) Calculer le seuil de la région critique pour un risque α=5%.
5) Calculer la puissance du test.
6) Enoncer les règles de décision avec les probabilités d'erreur.
7) La moyenne calculée sur l'échantillon est $\overline{x}=83$ décibels. Les habitants ont-ils raison de se plaindre ? Le test d'hypothèses ainsi établi leur est-il favorable ou défavorable ?

8) Combien faudrait-il faire de relevés journaliers, pour que le risque de deuxième espèce soit de 5% ?
9) Quelle serait alors le seuil de décision ?

1)
<u>Risque de 1ère espèce</u> : Accepter H_1 alors que H_0 est vraie $\Leftrightarrow$ « décider que l'aéroport n'est pas aux normes alors qu'il respecte la législation »
<u>Risque de 2nde espèce</u> : Accepter H_0 alors que H_1 est vraie $\Leftrightarrow$ « décider que l'aéroport est aux normes alors que le bruit dépasse le seuil autorisé »
Le test est donc fait du point de vue des responsables de l'aéroport.
2) La moyenne $\overline{X}$ est l'estimateur usuel de μ. L'échantillon est gaussien et $\sigma^2=64$ est connu. Donc

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

3)



On en déduit que la région critique (région d'acceptation de H_1) est définie par
$$W=\{\overline{X}>C\}.$$
4) $\alpha = P(W|H_0)$ or sous l'hypothèses H_0, $\overline{X}$ suit une loi normale $N(80,8^2/n)$, d'où

$$\alpha = P(W|H_0) \Leftrightarrow 0.05 = P(\overline{X} > C) \Leftrightarrow 0.05 = P(\sqrt{n}\frac{\overline{X}-80}{8} > \sqrt{n}\frac{C-80}{8})$$
$$\Leftrightarrow 0.05 = P(Z > C')$$

où Z suit une loi N(0,1).
A l'aide de la table de la loi normale, on obtient C'=1.65 puis
$$C = 80 + 1.65 * \frac{8}{\sqrt{n}} \approx 83.3.$$
5) Sous l'hypothèse H_1, $\overline{X}$ suit une loi $N(85,8^2/n)$, d'où
$$1-\beta = P(W|H_1) = P(\overline{X} > C) = P(\sqrt{n}\frac{\overline{X}-85}{8} > \sqrt{n}\frac{C-85}{8})$$
$$= P(Z > -0.85) = 1 - P(Z > 0.85) = 0.2$$
6)
- Si $\overline{x} > 83.3$ décibels alors on accepte H_1, *i.e* on considère que l'aéroport n'est pas aux normes, avec 5% de chance de se tromper.
- Si $\overline{x} < 83.3$ alors on garde H_0, *i.e* on considère que l'aéroport est aux normes, avec 20% chance de se tromper.

7) La moyenne calculée sur l'échantillon est $\bar{x} = 83 < 83.3$. Donc les habitants n'ont pas de raison de se plaindre. Cependant, le risque de deuxième espèce étant très élevé, ils sont en droit de remettre en cause le test.
8) Supposons que la taille de l'échantillon ne soit pas fixée. Alors le risque de 1ère espèce donne l'équation

$$C = 80 + 1.65 * \frac{8}{\sqrt{n}}$$

Et le risque de deuxième espèce donne

$$1-\beta = 0.95 = P(\sqrt{n}\frac{\overline{X}-85}{8} > \sqrt{n}\frac{C-85}{8}) = P(Z > C') = P(Z < -C') = 0.95$$

A l'aide de la table de la loi normale, on obtient C'=-1.65 puis

$$C = 85 - 1.65 * \frac{8}{\sqrt{n}}.$$

D'où

$$80 + 1.65 * \frac{8}{\sqrt{n}} = 85 - 1.65 * \frac{8}{\sqrt{n}} \Leftrightarrow \sqrt{n} = \frac{2*1.65*8}{85-80} \approx 5.3 \Rightarrow n \approx 27$$

et

$$C = 80 + 1.65 * \frac{8}{5.3} \approx 81$$

Il faudrait donc tester un échantillon de taille 27 et on aurait le seuil de décision à 81 décibels.

<u>Exercice 3</u>

Un fabricant de conserves de petits pois produit des boîtes où l'étiquette annonce un poids net égoutté de 560gr. Il souhaite construire un test pour s'assurer, d'une part qu'il n'aura pas d'ennui à l'issu d'un contrôle éventuel, et d'autre part, que le poids moyen des boîtes n'est pas excédentaire. Pour ce faire, il compte prélever un lot de 25 boites et relever le poids moyen ainsi que l'écart-type.

1) Déterminer les hypothèses et expliciter les risques de 1ère et 2ème espèces.
2) Quelle est la variable de décision ? Préciser sa loi.
3) Déterminer graphiquement l'allure de la région critique.
4) Calculer les seuils de la région critique sachant que le risque de 1ère espèce est 10%.
5) Peut-on calculer la puissance du test ?
6) Il prélève un lot de 25 boites et il pèse un poids moyen de $\bar{x} = 556\text{gr}$ avec un écart-type empirique $s^* = 10\text{gr}$. Quelle décision doit-il prendre ?

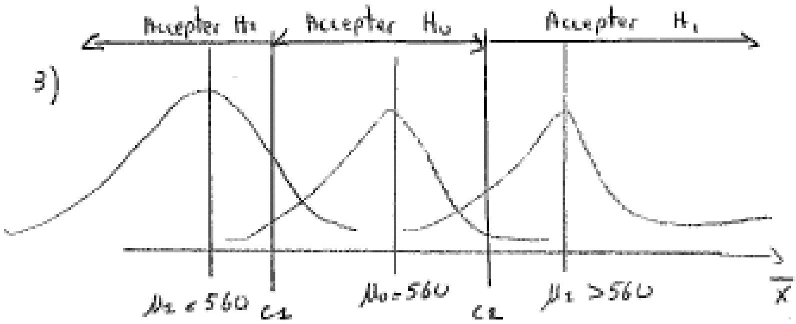
1) $\begin{cases} H_0 : \text{" poids conforme à celui annoncé "} : \mu = \mu_0 = 560 \\ H_1 : \text{" poids non conforme "} : \mu = \mu_1 < 560 \text{ ou } \mu = \mu_1 > 560 \end{cases}$

<u>Risque de 1ère espèce</u> : Accepter H_1 alors que H_0 est vraie
$\Leftrightarrow$ Revoir la production alors qu'elle est conforme
<u>Risque de 2ème espèce</u> : Accepter H_0 alors que H_1 vraie
$\Leftrightarrow$ considérer que le poids est conforme alors qu'il ne l'est pas

2) la moyenne $\overline{X}$ est un estimateur usuel de μ et sera donc la variable de décision. L'échantillon est de taille moyenne et l'écart-type est inconnu. On suppose donc que l'échantillon est gaussien, i.e que le poids d'une boite suit une loi normale $\mathcal{N}(\mu, \sigma^2)$ où σ^2 inconnu.
On a alors

$$Z = \sqrt{n}\,\frac{\overline{X}-\mu}{s^*} \leadsto t_{n-1}$$

3)



<u>Région critique</u> :
$W = \{ \overline{X} \leq C_1 \text{ ou } \overline{X} \geq C_2 \}$
<u>Région d'acceptation de H_0</u>
$\overline{W} = \{ C_1 \leq \overline{X} \leq C_2 \}$

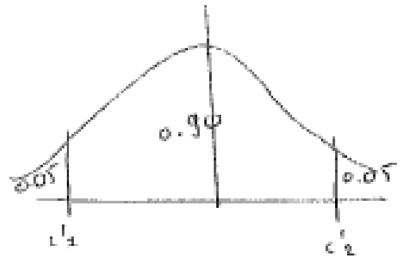
4) $\alpha = P[W \mid H_0 \text{ vraie}] \Leftrightarrow 1-\alpha = P[\overline{W} \mid H_0 \text{ vraie}]$ artifice pour éliminer le "ou"
Supposons H_0 vraie alors $\sqrt{n}\,\frac{\overline{X}-560}{s^*} \leadsto t_{n-1}$, d'où

$$1-\alpha = P[\overline{W}] = P[C_1 \leq \overline{X} \leq C_2] = P\left[\sqrt{n}\,\frac{C_1-560}{s^*} \leq \sqrt{n}\,\frac{\overline{X}-560}{s^*} \leq \sqrt{n}\,\frac{C_2-560}{\sqrt{n}}\right]$$

$$\Leftrightarrow \quad 0.90 = P[C'_1 \leq Z \leq C'_2]$$

On suppose un risque symétrique
De plus la loi de Student est symétrique par rapport à Oy d'où

$$\begin{cases} C'_1 = -C'_2 \\ P[Z \geq C'_2] = 0.05 \end{cases}$$



$$\underset{\text{table } t_{24}}{\Rightarrow} \begin{cases} C'_1 = -1.711 \\ C'_2 = 1.711 \end{cases} \Rightarrow \begin{cases} C_1 = 560 - \dfrac{1.711 * s^*}{\sqrt{n}} = 563.42 \\ C'_2 = 560 + \dfrac{1.711 * s^*}{\sqrt{n}} = 556.58 \end{cases}$$

5) On ne peut pas calculer la puissance du test car la loi de $\overline{X}$ sous l'hypothèse H_1 est inconnue car μ_1 n'est pas fixé.
6) $\overline{x} = 556 < C_1$, donc il accepte H_1, *ie* il revoir sa cycle de production, avec 10% de risque que celui-ci soit en bonne état.

<u>Exercice 4</u>

1) Sur un échantillon de 900 naissances, on constate qu'il y a 470 garçons. Un généticien décide d'utiliser ces données pour savoir si la proportion de garçons est significativement plus importante que la proportion de filles dans cette population.

2) Etablir le même test pour savoir si les proportions sont différentes.

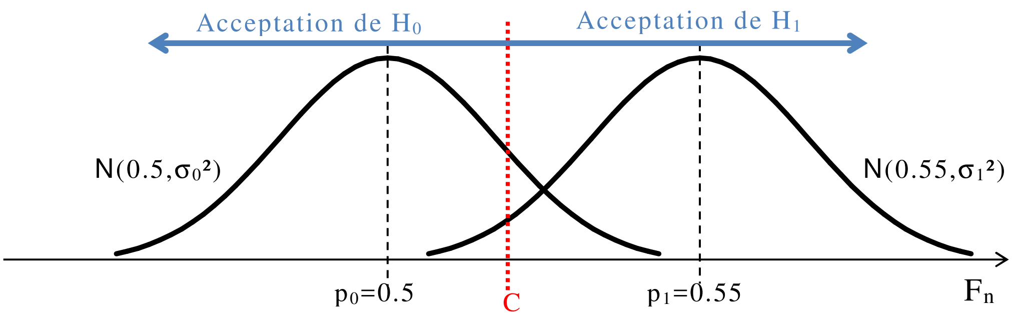
1)

$$H_0 : p=0.5$$
$$H_1 : p=>0.5$$

(a) Variable de décision

La fréquence empirique F_n est un estimateur de p et avec un échantillon de taille 900, on peut considérer grâce au TCL que F_n suit une loi normale N(p,p(1-p)/n) où p=0.5 sous l'hypothèse H_0 et p inconnu sous l'hypothèse H_1 .

(b) Allure de la région critique



La région critique W est la région d'acceptation de H_1 d'où $W=\{F_n \geq C\}$.

(c) Calcul du seuil

On sait que $\alpha = P(W \mid H_0 \text{ vraie}) = P(F_n \geq C \mid H_0 \text{ vraie})$.

Supposons H_0 vraie alors F_n suit une $N(0.5,\sigma_0^2)$ où $\sigma_0^2=0.5*0.5/n$, d'où

$$\alpha = P(F_n \geq C) = P\left(\frac{F_n - 0.5}{\sqrt{0.5*0.5}}\sqrt{n} \geq \frac{C-0.5}{0.5}\sqrt{n}\right) \Leftrightarrow P(Z \geq C')=0.05 \text{ où Z suit une loi } N(0,1)$$

$$\Rightarrow C'=1.64 \Rightarrow C=0.5+1.64*\frac{0.5}{\sqrt{n}}.$$

$$\Rightarrow C \approx 0.5+1.64*0.5/30 \approx 0.53$$

(d) Règles de décision

- Si $f_n \geq 0.53$ alors on accepte H_1, i.e on considère qu'il y a plus de garçons que de filles avec 5% de risque de se tromper.
- Si $f_n < 0.53$ alors on garde H_0, i.e on considère qu'il y a autant de filles que de garçons (sans connaitre le risque se tromper).

L'échantillon considéré indique $f_n=470/900=0.52$ donc le généticien conclut qu'il y autant de garçons que de filles.

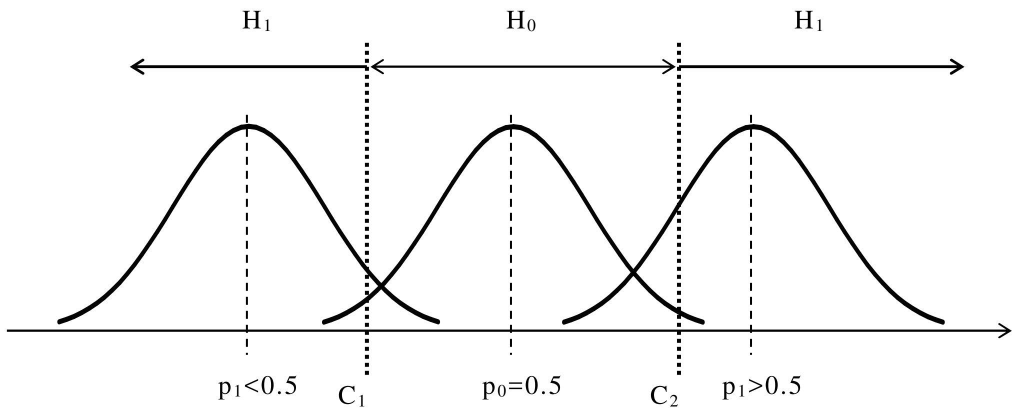
2)

$$H_0 : p=0.5$$
$$H_1 : p \neq 0.5$$

Le test est maintenant bilatéral. Même estimateur, même lois.

(a) Région critique

La région critique W est la région d'acceptation de H_1 d'où $W=\{F<C_1 \text{ ou} F>C_2\}$



et $\overline{W}$ est la région d'acceptation de H_0 d'où $\overline{W}=\{C_1 \le F \le C_2\}$ (cf. dessin).

(b) Calcul des seuils

On sait que $\alpha=P(W \mid H_0 \text{ vraie})=P(F<C_1 \text{ ou } F>C_2 \mid H_0 \text{ vraie})$. Afin de simplifier le calcul de probabilité, on passe à $1-\alpha=P(\overline{W} \mid H_0 \text{ vraie})=P(C_1 \le F \le C_2 \mid H_0 \text{ vraie})$.

Supposons H_0 vraie alors F suit une $N(0.5,\sigma_0^2)$ où $\sigma_0^2=0.5*0.5/n$, d'où

$$1-\alpha=P(C_1 \le F \le C_2)=P\left(\frac{C_1-0.5}{\sqrt{0.5*0.5}}\sqrt{n} \le \frac{F-0.5}{\sqrt{0.5*0.5}}\sqrt{n} \le \frac{C_2-0.5}{\sqrt{0.5*0.5}}\sqrt{n}\right)$$

$\Leftrightarrow P(C'_1 \le Z \le C_2')=0.95$.

On suppose que le risque est symétrique, or la loi normale centrée est aussi symétrique par rapport à 0, on a donc $C'_1=-C'_2$ et $P(Z \ge C_2')=0.025 \Rightarrow C'_2=1.96 \Rightarrow$ $C_2=0.5+1.96*\sqrt{0.5*0.5}/\sqrt{n}=0.53$ et $C_1=0.5-1.96*\sqrt{0.5*0.5}/\sqrt{n}=0.47$

(c) Règles de décision

- Si $f_n<0.47$ ou $f_n>0.53$, on accepte H_1, *i.e* on considère qu'il n'y a pas autant de filles que de garçons avec 5% de chance de se tromper.
- Dans le cas contraire, on accepte H_0, i.e on considère que la proportion de filles et de garçons est la même, mais on ne connaît pas le risque encourus car on ne connaît pas la loi sous H_1.

Exercice 5

Un fabricant produit des piles dont la durée moyenne de vie annoncée est 80h avec un écart-type de 7,44h. Suite à des réclamations, il veut vérifier si la qualité a baissé ou non. Pour l'hypothèse alternative, il choisit de tester si la durée de vie moyenne est 75h. Le cas échéant, il sera obligé de baisser le prix de vente.

1) Déterminer les hypothèses et expliciter les risques de $1^{ère}$ et $2^{ème}$ espèces.
2) Quelle est la statistique du test ? Préciser sa loi.
3) Déterminer graphiquement l'allure de la région critique.
4) Enoncer la règle de décision sachant que le fabricant souhaite maîtriser les deux erreurs avec $\alpha=0.025$ et $\beta=0.05$.

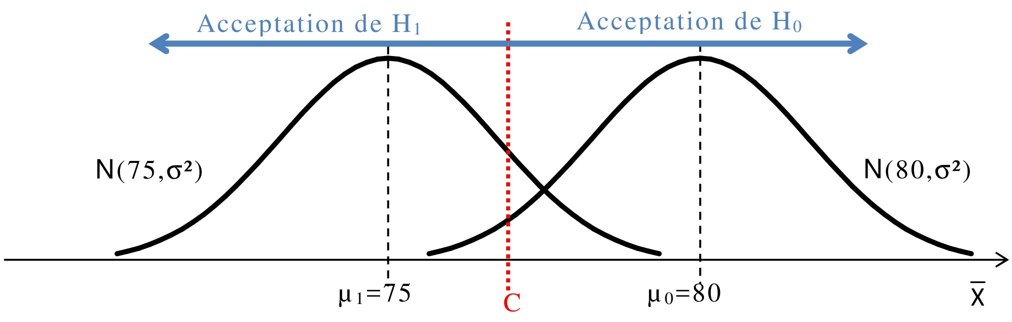
1) Soit μ la durée de vie moyenne des piles

$$H_0 : \mu=80h$$
$$H_1 : \mu=75h$$

<u>Risque de 1ère espèce</u> : Accepter H_1 alors que H_0 est vraie ⇔ « baisser le prix de vente alors que la marchandise est conforme »
<u>Risque de 2nde espèce</u> : Accepter H_0 alors que H_1 est vraie ⇔ « vendre au même prix des piles non conformes »

2) La statistique du test est la moyenne de la durée de vie des piles testées de l'échantillon. Nous allons supposer a priori que l'échantillon est assez grand pour approcher sa loi par une loi normale $N(\mu,\sigma^2/n)$ où σ=7,44h.

3) La région critique est de la forme : $W=\{\overline{X}<C\}$



4)
- Equation 1 : $\alpha=P(W|H_0)$

or sous l'hypothèses H_0, $\overline{X}$ suit une loi normale $N(80,7.44^2/n)$, d'où

$$\alpha=P(W|H_0)\Leftrightarrow 0.025=P(\overline{X}<C)\Leftrightarrow 0.025=P(\sqrt{n}\frac{\overline{X}-80}{7,44}<\sqrt{n}\frac{C-80}{7,44})$$
$$\Leftrightarrow 0.025=P(Z<C')$$

où Z suit une loi N(0,1).
A l'aide de la table de la loi normale, on obtient C'=-1.96 puis

$$\boxed{C=80-1.96*\frac{7,44}{\sqrt{n}}}.$$

- Equation 2 : $\beta=P(\text{Accepter } H_0|H_1)=P(\overline{X}>C|H_1)$

or sous l'hypothèses H_1, $\overline{X}$ suit une loi normale $N(75,7.44^2/n)$, d'où

$$\beta=P(\overline{X}>C|H_1)\Leftrightarrow 0.05=P(\overline{X}>C)\Leftrightarrow 0.05=P(\sqrt{n}\frac{\overline{X}-75}{7,44}>\sqrt{n}\frac{C-75}{7,44})$$
$$\Leftrightarrow 0.05=P(Z>C')$$

où Z suit une loi N(0,1).
A l'aide de la table de la loi normale, on obtient C'=1.65 puis

$$\boxed{C=75+1.65*\frac{7.44}{\sqrt{n}}}.$$

On a donc deux inconnues C et n et deux équations. On en déduit $\sqrt{n}$=6.1 et C=77.
Avec un échantillon de taille 38 (arrondi à l'entier supérieur), on a la régle de décision suivante :

- Si $\overline{x} < 77h$ alors on accepte H_1, *i.e* on considère que la marchandise n'est pas conforme, avec 2,5% de chance de se tromper.
- Si $\overline{x} > 77h$ alors on garde H_0, *i.e* on considère que la marchandise est conforme, avec 5% chance de se tromper.

On valide a posteriori que l'échantillon est assez grand pour le TCL