MTH2302D - TD 10

Vincent Perreault

POLYTECHNIQUE
MONTRÉAL

Ex. 1 - Énoncé

On sait que le diamètre d'un certain type de boulon présente un écart-type de 0,001 cm, et que ce diamètre suit une loi normale. Le responsable du contrôle de la qualité croit que le diamètre moyen diffère de la norme en vigueur, soit 0,255 cm. Il effectue un test d'hypothèse au seuil de 1 % à partir d'un échantillon aléatoire de 10 boulons. Il obtient une valeur *P* égale à 0,028.

a) Quelles sont les hypothèses statistiques appropriées pour répondre aux interrogations du responsable ?

b) Quelle conclusion le responsable a-t-il tiré de son expérience ?

c) Pouvez-vous déduire la valeur observée de la statistique du test (z_0) ?

d) Sachant que le diamètre moyen obtenu dans cet échantillon est supérieur à la norme en vigueur, déterminez sa valeur à la lumière des informations que vous possédez.

- Si $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, $X \sim N(\mu, \sigma^2)$ et σ^2 connue,
  alors $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$, on rejette H_0 si $|Z_0| > z_{\alpha/2}$
  et $p-value = 2P(Z > |z_0|) = 2(1 - \Phi(|z_0|))$

Ex. 1 - Réponse - a), b)

Quelles sont les hypothèses?

$$H_0 : \mu = 0.255, \quad H_1 : \mu \neq 0.255$$

Quelle est la conclusion de l'expérience?

On ne peut pas rejeter H_0 car
$p-value = 0.028 > 0.01 = \alpha.$

Le diamètre des boulons ne diffère pas significativement de 0.255.

$X \sim \mathrm{N}(\mu, 0.001^2)$
norme : 0.255

test $\alpha = 0.01$, $n = 10$

$\Rightarrow p-value = 0.028$

$H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$,
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue

$$\Rightarrow Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$\Rightarrow p-value = 2(1 - \Phi(|z_0|))$

Ex. 1 - Réponse - c), d)

Quel est le Z_0 correspondant?

$$p\!-\!value = 2(1 - \Phi(|z_0|))$$

$$\Leftrightarrow \Phi(|z_0|) = 1 - \frac{1}{2}p\!-\!value = 1 - \frac{1}{2}0.028 = 0.986$$

$$\Rightarrow |z_0| \approx 2.20 \;\; \Rightarrow z_0 \approx \pm 2.20$$

Quelle était $\bar{X}$ sachant que $\bar{X} > 0.255$?

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$\Leftrightarrow \bar{X} = \mu_0 + \frac{\sigma}{\sqrt{n}} z_0 > \mu_0 \Rightarrow z_0 > 0$$

$$= 0.255 + \frac{0.001}{\sqrt{10}} 2.20 \approx 0.2557$$

$X \sim \mathrm{N}(\mu, 0.001^2)$
norme : 0.255

test $\alpha = 0.01$, $n = 10$
$\Rightarrow p\!-\!value = 0.028$

$H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$,
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue
$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$
$\Rightarrow p\!-\!value = 2(1 - \Phi(|z_0|))$

Ex. 2 - Énoncé

- Si $H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$, $X \sim \mathrm{N}(\mu, \sigma^2)$ et σ^2 inconnue,
  alors $T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$, on rejette H_0 si
  $T_0 > t_{\alpha;n-1}$

Un article du *Journal of Construction Engineering and Management* (vol. 125, nº 1, 1999, p. 39) présente quelques données sur le nombre d'heures de travail perdues par jour, dans le cas d'un projet de construction, en raison de problèmes dus au mauvais temps. Voici le nombre d'heures de travail perdues sur 11 jours ouvrables.

8,8	8,8	9,1
12,5	12,2	14,7
5,4	13,3	2,2
12,8	6,9	

Dans ces données, on a $\bar{x} = 9{,}7$ et $s^2 = 14{,}62$. Supposez que le nombre d'heures de travail perdues est normalement distribué.

a) Y a-t-il une preuve permettant de conclure que le nombre moyen d'heures de travail perdues par jour est supérieur à 8 au seuil de 5 % ? Énoncez clairement les hypothèses que vous testez et la règle de rejet de H_0.

b) La probabilité d'erreur de deuxième espèce (β) serait-elle supérieure :

i) si la taille d'échantillon était de 25 au lieu de 11 ?

ii) si le seuil du test était de $\alpha = 1\,\%$ au lieu de $\alpha = 5\,\%$?

iii) si le véritable nombre moyen d'heures perdues était de $\mu_1 = 10$ heures ou de $\mu_1 = 12$ heures ?

Ex. 2 - Réponse - a)

Peut-on prouver avec $\alpha = 0.05$ que $\mu > 8$?

$$H_0 : \mu = 8, \quad H_1 : \mu > 8$$

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \text{ rejeter } H_0 \text{ si}$$
$$T_0 > t_{\alpha;n-1} = t_{0.05;10} = 1.8125$$

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{9.7 - 8}{\sqrt{14.62}/\sqrt{11}} \approx 1.4746$$

Non, on ne peut pas rejeter l'hypothèse $\mu = 8$
(ou éventuellement moins).

$X \sim \mathrm{N}(\mu, \sigma^2)$

$\bar{x} = 9.7, \quad s^2 = 14.62, \quad n = 11$

$H_0 : \mu = \mu_0, \quad H_1 : \mu > \mu_0,$
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 inconnue
$\Rightarrow T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$
$\Rightarrow$ rejeter H_0 si $T_0 > t_{\alpha;n-1}$

Ex. 2 - Réponse - b)

β serait-il supérieur si $n = 25$ au lieu de $n = 11$?

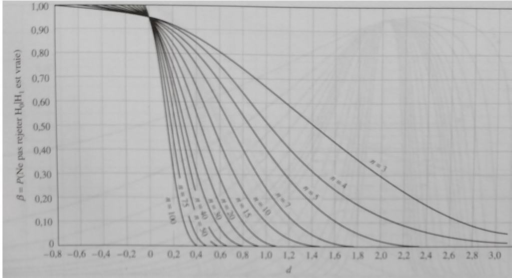
Non, β diminue si n augmente.

β serait-il supérieur si $\alpha = 0.01$ au lieu de $\alpha = 0.05$?

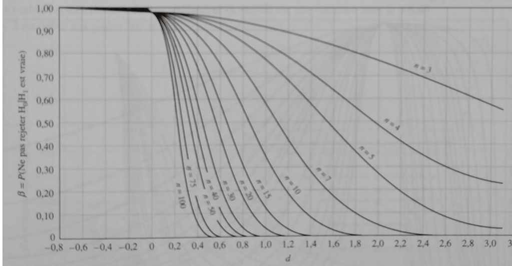
Oui, β augmente si α diminue.

β serait-il supérieur si $\mu_1 = 10$ ou si $\mu_1 = 12$?

$\mu_1 = 10$, *β diminue si $d = \mu_1 - \mu_0$ augmente.*



Courbes caractéristiques du test *t* **unilatéral** avec $\boldsymbol{\alpha = 0{,}05}$
- Test sur une moyenne : *d* est défini à la page 300.
- Test sur une différence de moyennes : *d* est défini à la page 314.



Courbes caractéristiques du test *t* **unilatéral** avec $\boldsymbol{\alpha = 0{,}01}$
- Test sur une moyenne : *d* est défini à la page 300.
- Test sur une différence de moyennes : *d* est défini à la page 314.

Ex. 3 - Énoncé

- Si $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$, $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$
et σ_1^2, σ_2^2 connues,
alors $Z_0 = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$,
on rejette H_0 si $|Z_0| > z_{\alpha/2}$
$\beta =$
$$\Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$$

Deux machines remplissent des bouteilles en plastique d'un volume net de 16 centilitres. On suppose raisonnablement que les procédés de remplissage sont normaux avec pour écarts-types $\sigma_1 = 0{,}015$ et $\sigma_2 = 0{,}018$. Le service de contrôle de la qualité soupçonne que les deux machines ne versent pas le même volume net, que ce volume soit de 16 centilitres ou pas. Un échantillon aléatoire est prélevé de la production de chaque machine.

	Machine 1		**Machine 2**	
	16,03	16,01	16,02	16,03
	16,04	15,96	15,97	16,04
	16,05	15,98	15,96	16,02
	16,05	16,02	16,01	16,01
	16,02	15,99	15,99	16,00
Moyenne	16,015		16,005	
Écart-type	0,0303		0,0255	

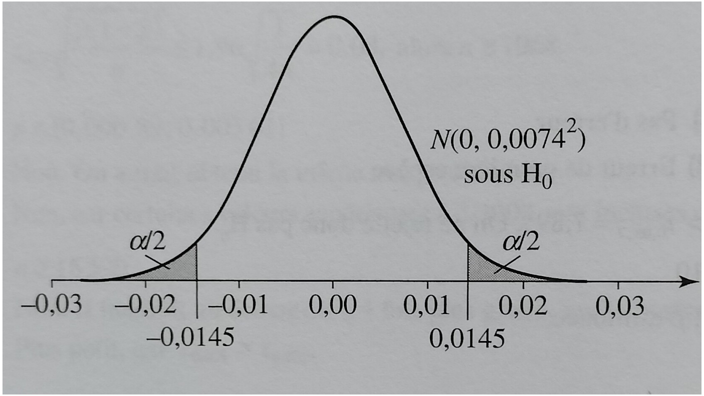
a) Tracez la distribution de la différence de moyennes échantillonnales sous H_0, et identifiez la zone de rejet de H_0.

b) Le service de contrôle de la qualité a-t-il raison ? Utilisez $\alpha = 0{,}05$.

c) Calculez la puissance du test pour une véritable différence des moyennes de 0,01 centilitre.

Ex. 3 - Réponse - a)

Tracer la distribution de $\bar{X}_1 - \bar{X}_2$ sous H_0.



$X_1 \sim \mathrm{N}(\mu_1, 0.015^2)$, $X_2 \sim \mathrm{N}(\mu_2, 0.018^2)$
$\bar{X}_1 = 16.015$, $\bar{X}_2 = 16.005$
$S_1 = 0.0303$, $S_2 = 0.0255$

$n_1 = n_2 = 10$

$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$,
$X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, σ_1^2, σ_2^2 connues,

$$\Rightarrow Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$$

Ex. 3 - Réponse - b)

Tester si $\mu_1 = \mu_2$.

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$$

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \text{ rejeter } H_0 \text{ si}$$

$$|Z_0| > z_{\alpha/2} = z_{0.05/2} = 1.96$$

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{16.015 - 16.005}{\sqrt{\frac{0.015^2}{10} + \frac{0.018^2}{10}}} \approx 1.35$$

On ne peut pas rejeter l'hypothèse $\mu_1 = \mu_2$. Les moyennes ne diffèrent pas significativement. (ou éventuellement moins).

$X_1 \sim \mathrm{N}(\mu_1, 0.015^2)$, $X_2 \sim \mathrm{N}(\mu_2, 0.018^2)$
$\bar{X}_1 = 16.015$, $\bar{X}_2 = 16.005$
$S_1 = 0.0303$, $S_2 = 0.0255$

$n_1 = n_2 = 10$

$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$,
$X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, σ_1^2, σ_2^2 connues,

$$\Rightarrow Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$$

Ex. 3 - Réponse - c)

Calculer la puissance $1-\beta$ du test si $\mu_1-\mu_2 = 0.01$.

$$\beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right)$$

$$= \Phi\left(1.96 - \frac{0.01}{\sqrt{\frac{0.015^2}{10}+\frac{0.018^2}{10}}}\right) - \Phi\left(-1.96 - \frac{0.01}{\sqrt{\frac{0.015^2}{10}+\frac{0.018^2}{10}}}\right)$$

$$= \Phi(0.610) - \Phi(-3.310)$$

$$= \Phi(0.610) - (1-\Phi(3.310))$$

$$= 0.72907 - (1-0.99953) = 0.7286$$

$$\Rightarrow 1-\beta = 0.2714$$

$X_1 \sim \mathrm{N}(\mu_1, 0.015^2)$, $X_2 \sim \mathrm{N}(\mu_2, 0.018^2)$
$\bar{X}_1 = 16.015$, $\bar{X}_2 = 16.005$
$S_1 = 0.0303$, $S_2 = 0.0255$
$n_1 = n_2 = 10$

$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$,
$X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, σ_1^2, σ_2^2 connues,

$$\Rightarrow Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}},$$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right)$$

Ex. 4 - Énoncé

Exercice n° 4 Une machine (i.e., un tour) sert à usiner des billes dont le diamètre est distribué selon une loi normale avec un écart type de 0,1 *cm*. Lorsque la machine est correctement ajustée, le diamètre des billes produites présente alors une moyenne de 2,0 *cm* et une variabilité stable. Afin de vérifier si la machine est correctement ajustée, on compte effectuer un test statistique à l'aide échantillon aléatoire de n billes prélevé dans la production. Le test devrait être tel que la probabilité de conclure à tort que la machine n'est pas correctement ajustée est de 0,05 ; et celle de conclure que la machine est correctement ajustée alors que le diamètre moyen est de 2,25 *cm* est de 0,10.

a) Formuler les hypothèses à tester en précisant les paramètres.

b) Déterminer la taille d'échantillon n nécessaire.

c) Un échantillon de 36 billes de la production donne une moyenne de 2,18 *cm*.
Peut-on conclure que la machine n'est pas correctement ajustée ?
Quelle est la valeur p (i.e. la «*p-value*») du test utilisé ?

- Si $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, $X \sim \mathrm{N}(\mu, \sigma^2)$ et σ^2 connue, alors $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$, on rejette H_0 si $|Z_0| > z_{\alpha/2}$

$$p\text{-}value = 2P(Z > |z_0|) = 2(1 - \Phi(|z_0|))$$

et $n \approx \frac{(z_{\alpha/2}+z_\beta)^2\sigma^2}{(\mu_1-\mu_0)^2}$

Ex. 4 - Réponse - a), b)

Quels sont les hypothèses et les paramètres.

$$H_0 : \mu = 2.0, \quad H_1 : \mu \neq 2.0$$

où

$$\alpha = P(\text{rejeter } \mu = \mu_0 | \mu = \mu_0) = 0.05$$

$$\beta = P(\text{accepter } \mu = \mu_0 | \mu = 2.25) = 0.10$$

Déterminer la taille n nécessaire.

$$n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2} = \frac{(z_{0.05/2} + z_{0.1})^2 \sigma^2}{(\mu_1 - \mu_0)^2}$$

$$\approx \frac{(1.96 + 1.285)^2 0.1^2}{(2.25 - 2)^2} \approx 1.684$$

$\Rightarrow n = 2$

$X \sim \mathrm{N}(\mu, 0.1^2)$

norme : $\mu_0 = 2.0$

$P(\text{rejeter } \mu = \mu_0 | \mu = \mu_0) = 0.05$

$P(\text{accepter } \mu = \mu_0 | \mu = 2.25) = 0.10$

$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$

$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue

$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$\Rightarrow p-value = 2(1 - \Phi(|z_0|))$

$\Rightarrow n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2}$

Ex. 4 - Réponse - c)

Si $\bar{X} = 2.18$ pour $n = 36$, la machine est-elle non correctement ajustée? Quelle est la p-valeur du test?

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.18 - 2.0}{0.1/\sqrt{36}} = 10.8$$

On rejette H_0 si $|Z_0| > z_{\alpha/2} = z_{0.05/2} = 1.96$

Oui, on rejette l'hypothèse que la machine est correctement ajustée

$$p\text{−}value = 2(1 - \Phi(|z_0|)) = 2(1 - \Phi(|10.8|))$$
$$\approx 2(1 - 1) = 0. < 0.05$$

$X \sim \mathrm{N}(\mu, 0.1^2)$
norme : $\mu_0 = 2.0$

$P(\text{rejeter } \mu = \mu_0 | \mu = \mu_0) = 0.05$
$P(\text{accepter } \mu = \mu_0 | \mu = 2.25) = 0.10$

$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue
$\Rightarrow Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$
$\Rightarrow p\text{−}value = 2(1 - \Phi(|z_0|))$
$\Rightarrow n \approx \frac{(z_{\alpha/2}+z_\beta)^2\sigma^2}{(\mu_1-\mu_0)^2}$

Ex. 5 - Énoncé

Exercice nº 5 Des expériences planifiées ont permis de mesurer la résistance à l'éclatement (en lb/po^2) pour un échantillon de 41 bouteilles d'une production en série. Les données obtenues se trouvent dans le fichier `td10ex5.csv`.

Mesures de résistance à l'éclatement (en lb/po^2)

189,32	201,25	195,38	195,47	199,86	185,42	190,18	197,53
199,07	195,74	190,41	193,99	197,73	205,68	192,95	192,95
195,85	198,00	200,78	179,70	202,81	195,70	191,88	203,32
191,70	201,50	184,38	200,97	197,13	197,33	188,32	190,31
192,52	197,67	198,81	180,48	204,20	187,43	193,80	193,78
188,15							

a) L'hypothèse d'une distribution normale (i.e. gaussienne) des mesures de résistance est-elle plausible ? Utiliser un test et illustrer votre conclusion à l'aide d'un graphique approprié.

a) $\rightarrow$ oui

b) Sur la base d'un test d'hypothèses, peut-on conclure que l'écart type de la résistance à l'éclatement des bouteilles est inférieur à 6,5 lb/po^2 ? Justifier brièvement votre raisonnement.

c) Le cahier des charges exige que la résistance moyenne à l'éclatement des bouteilles soit de plus de 190 lb/po^2.

1.c Sur la base d'un test d'hypothèses, peut-on conclure que les bouteilles de cette production respectent les exigences du cahier des charges ? Justifier brièvement votre raisonnement.

2.c) Reprendre la question (1.d) ci-dessus en se basant sur la valeur p (i.e. «*p-value*») du test.

Ex. 5 - Réponse - a)

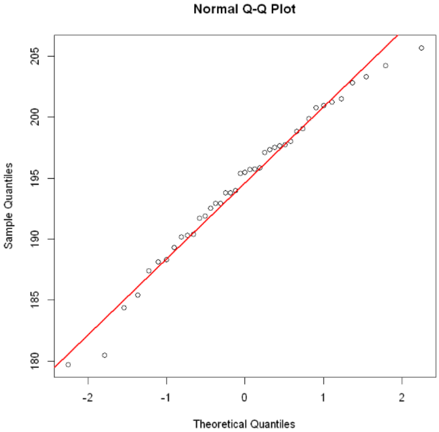
Normalité plausible?

```r
resist <- read.csv("Td10ex5.csv", header = TRUE, sep = ";",dec = ",")
resist <- resist$RESIST
shapiro.test(resist)
```

```
        Shapiro-Wilk normality test

data:  resist
W = 0.97679, p-value = 0.5563
```

p-value >> 0.05 **---> normalité plausible**

```r
qqnorm(resist)
qqline(resist, col="red", lwd=2)
```



Ex. 5 - Réponse - b)

Écart-type inférieur à 6.5?

$$H_0 : \sigma = 6.5, \quad H_1 : \sigma < 6.5$$

avec $X \sim \mathrm{N}(\mu, \sigma^2)$,

alors $\chi_0^2 = (n-1)\dfrac{S^2}{\sigma_0^2}$, on rejette H_0 si

$$\chi_0^2 < \chi^2_{1-\alpha;n-1} = \chi^2_{1-0.05;40} = 26.51$$

c'est-à-dire si $p-value < \alpha = 0.05$

```r
#install.packages("EnvStats")
library(EnvStats)

varTest(resist, sigma.squared = 6.5^2, alternative = "less", conf.level = 0.95)
```

```
Results of Hypothesis Test
--------------------------

Null Hypothesis:                 variance = 42.25

Alternative Hypothesis:          True variance is less than 42.25

Test Name:                       Chi-Squared Test on Variance

Estimated Parameter(s):          variance = 37.65445

Data:                            resist

Test Statistic:                  Chi-Squared = 35.64918

Test Statistic Parameter:        df = 40

P-value:                         0.3336103

95% Confidence Interval:         LCL =  0.00000
                                 UCL = 56.81695
```

X^2 >= 26.51

p-value >> 0.05 **---> non, on ne peut pas rejeter l'hypothèse que *σ* = 6.5 (ou éventuellement plus)**

Ex. 5 - Réponse - c)

Moyenne supérieure à 190?

$$H_0 : \mu = 190, \quad H_1 : \mu > 190$$

$$\text{avec } X \sim N(\mu, \sigma^2),$$

$$\text{alors } T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \text{ on rejette } H_0 \text{ si}$$

$$T_0 > t_{\alpha;n-1} = t_{0.05;40} = 1.6839$$

c'est-à-dire si $p-value < \alpha = 0.05$

```
t.test(resist, mu = 190, alternative = "greater")
```

```
        One Sample t-test

data:  resist
t = 4.8216, df = 40, p-value = 1.043e-05
alternative hypothesis: true mean is greater than 190
95 percent confidence interval:
 193.007     Inf
sample estimates:
mean of x
 194.6207
```

$t > 1.6839$

p-value < 0.05 **---> on rejette H0 pour l'alternative μ > 190**

Ex. 6 - Énoncé

Exercice nº 6 Lors d'un test sur la moyenne d'une population normale de variance 9, la règle de décision est de rejeter H_0 si $\overline{X} \notin [13{,}77, \quad 16{,}23]$ en se basant sur les résultats d'un échantillon de taille 25.

a) Énoncer les hypothèses à confronter.

b) Quelle est la probabilité de commettre l'erreur de première espèce (type I) ?

c) Quelle est la probabilité de commettre l'erreur de deuxième espèce (type II) si en réalité la moyenne est de 16,5 ? Répondre à cette question en procédant de deux manières :

1.c) En utilisant directement les formules de la table de β et n.

2.c) En utilisant la définition de l'erreur de type II.

d) Quelle doit-être la taille de l'échantillon prélevé si on veut détecter une moyenne théorique de 16,5 au seuil α calculé en b) avec une probabilité de 0,90 ?

- Si $H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$ $X \sim \mathrm{N}(\mu, \sigma^2)$ et σ^2 connue, alors $Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$, on rejette H_0 si $|Z_0| > z_{\alpha/2}$

$$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$$

$$\beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right)$$

et $n \approx \dfrac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2}$

Ex. 6 - Réponse - a)

Quelles sont les hypothèses?

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$$

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}, \text{ rejeter } H_0 \text{ si } |Z_0| > z_{\alpha/2}$$

$$|Z_0| > z_{\alpha/2}$$
$$\Leftrightarrow Z_0 \notin [-z_{\alpha/2}, z_{\alpha/2}]$$
$$\Leftrightarrow \bar{X} \notin [\underbrace{\mu_0 - \tfrac{\sigma}{\sqrt{n}} z_{\alpha/2}}_{L}, \underbrace{\mu_0 + \tfrac{\sigma}{\sqrt{n}} z_{\alpha/2}}_{U}]$$

$$\Rightarrow \mu_0 = \frac{L+U}{2} = \frac{13.77 + 16.23}{2} = 15$$

$$H_0 : \mu = 15, \quad H_1 : \mu \neq 15$$

$X \sim \mathrm{N}(\mu, 9)$

test : $n = 25$, rejeter H_0 si
$\bar{X} \notin [13.77, 16.23]$

$H_0 : \mu = \mu_0, \;\; H_1 : \mu \neq \mu_0,$
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue
$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$
$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right) -$
$\Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right)$
$\Rightarrow n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2}$

$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$

Ex. 6 - Réponse - b)

Quel est α?

$$U = \mu_0 + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

$$\Leftrightarrow z_{\alpha/2} = \frac{U - \mu_0}{\sigma/\sqrt{n}} = \frac{16.23 - 15}{\sqrt{9}/\sqrt{25}} = 2.05$$

$$\Rightarrow 1 - \alpha/2 = 0.97982$$

$$\Rightarrow \alpha \approx 0.0404$$

$X \sim \mathrm{N}(\mu, 9)$

test : $n = 25$, rejeter H_0 si

$\bar{X} \notin [13.77, 16.23]$

$H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$,

$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue

$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right) -$
$\Phi\left(-z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right)$

$\Rightarrow n \approx \frac{(z_{\alpha/2}+z_\beta)^2\sigma^2}{(\mu_1-\mu_0)^2}$

$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$

Ex. 6 - Réponse - 1.c)

$X \sim \mathrm{N}(\mu, 9)$

Quel est β si $\mu_1 = 16.5$?

test : $n = 25$, rejeter H_0 si $\bar{X} \notin [13.77, 16.23]$

$$\beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right)$$

$$= \Phi\left(2.05 - \frac{(16.5 - 15)\sqrt{25}}{\sqrt{9}}\right) - \Phi\left(-2.05 - \frac{(16.5 - 15)\sqrt{25}}{\sqrt{9}}\right)$$

$$= \Phi(-0.45) - \Phi(-4.55) = 1 - \Phi(0.45) - (1 - \Phi(4.55))$$

$$= \Phi(4.55) - \Phi(0.45) = 0.9999973 - 0.67364 \approx 0.326$$

$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue
$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$
$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$
$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right)$
$\Rightarrow n \approx \frac{(z_{\alpha/2}+z_\beta)^2\sigma^2}{(\mu_1-\mu_0)^2}$

$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$

Ex. 6 - Réponse - 2.c)

Quel est β si $\mu_1 = 16.5$?

$$
\begin{aligned}
\beta &= P(\text{accepter } H_0 | \mu = \mu_1) \\
&= P(\bar{X} \in [L, U] | \mu = \mu_1) \\
&= P(\bar{X} \leq U | \mu = \mu_1) - P(\bar{X} \leq L | \mu = \mu_1) \\
&= P\left(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} \leq \frac{U-\mu_1}{\sigma/\sqrt{n}} \mid |\mu = \mu_1\right) - P\left(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} \leq \frac{L-\mu_1}{\sigma/\sqrt{n}} \mid |\mu = \mu_1\right) \\
&= P\left(Z \leq \frac{U-\mu_1}{\sigma/\sqrt{n}}\right) - P\left(Z \leq \frac{L-\mu_1}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(\frac{U-\mu_1}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{L-\mu_1}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(\frac{16.23-16.5}{\sqrt{9}/\sqrt{25}}\right) - \Phi\left(\frac{13.77-16.5}{\sqrt{9}/\sqrt{25}}\right) \\
&= \Phi(-0.45) - \Phi(-4.55) = \ldots \approx 0.326
\end{aligned}
$$

$X \sim \mathrm{N}(\mu, 9)$

test : $n = 25$, rejeter H_0 si $\bar{X} \notin [13.77, 16.23]$

$H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$,
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue

$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1-\mu_0)\sqrt{n}}{\sigma}\right)$

$\Rightarrow n \approx \frac{(z_{\alpha/2}+z_\beta)^2\sigma^2}{(\mu_1-\mu_0)^2}$

$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$

Ex. 6 - Réponse - d)

Quelle taille n pour pouvoir détecter $\mu \neq \mu_0$ avec proba 0.9 si $\mu = 16.5$?

puissance $1 - \beta = P(\text{refuser } \mu = \mu_0 | \mu = \mu_1) = 0.9$

$\Rightarrow \beta = 0.1$

$$n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2} = \frac{(z_{\alpha/2} + z_{0.1})^2 \sigma^2}{(\mu_1 - \mu_0)^2}$$

$$\approx \frac{(2.05 + 1.285)^2 \cdot 9}{(16.5 - 15)^2} \approx 44.49$$

$\Rightarrow n = 45$

$X \sim \mathrm{N}(\mu, 9)$

test : $n = 25$, rejeter H_0 si $\bar{X} \notin [13.77, 16.23]$

$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0,$
$X \sim \mathrm{N}(\mu, \sigma^2)$, σ^2 connue

$\Rightarrow Z_0 = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

$\Rightarrow$ rejeter H_0 si $|Z_0| > z_{\alpha/2}$

$\Rightarrow \beta = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_0)\sqrt{n}}{\sigma}\right)$

$\Rightarrow n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{(\mu_1 - \mu_0)^2}$

$\beta = P(\text{accepter } H_0 | \mu = \mu_1)$

Ex. 7 - Énoncé

Exercice nº 7 Une compagnie utilise deux procédés (méthodes) pour produire un type de rondelle (joint) d'étanchéité. Au service du contrôle de la qualité de la compagnie, on aimerait vérifier si les rondelles produites par les deux procédés possèdent les mêmes caractéristiques. Les données suivantes sont des mesures d'épaisseur obtenues pour deux échantillons aléatoires : un de 12 rondelles provenant du premier procédé, et l'autre de 18 rondelles provenant du deuxième procédé.

Mesures d'épaisseur en mm

Procédé 1		Procédé 2		
4,1	4,6	2,7	5,6	5,7
5,1	4,6	5,4	4,3	6,3
4,9	4,3	7,7	3,7	4,2
5,5	4,5	6,4	4,2	4,1
4,2	4,8	5,0	5,8	4,0
3,9	4,5	3,2	4,4	4,9

Remarque : Ces données se trouvent dans le fichier `td10ex7.csv`.

a) Montrer que l'hypothèse d'une distribution normale pour l'épaisseur des rondelles est plausible pour chacun des deux procédés.

b) Tester au seuil critique de 5% si les variances sont égales dans les deux populations. Calculer la valeur p («p-value») du test utilisé et commenter brièvement.

c) En se basant sur le résultat en b), tester au seuil critique de 5% si l'épaisseur des rondelles est la même en moyenne pour les deux procédés. Calculer la valeur p («p-value») du test utilisé et commenter brièvement.

Ex. 7 - Réponse - a)

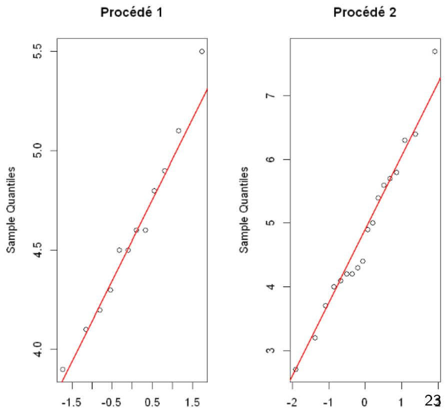
Normalités plausibles?

```r
donnees <- read.csv("Td10ex7.csv", header = TRUE, sep = ";",dec = ",")
proc1 <- donnees$EPAR[donnees$PROC=="P1"]
proc2 <- donnees$EPAR[donnees$PROC=="P2"]

shapiro.test(proc1)
shapiro.test(proc2)
```

```
        Shapiro-Wilk normality test

data:  proc1
W = 0.97618, p-value = 0.9637


        Shapiro-Wilk normality test

data:  proc2
W = 0.97496, p-value = 0.884
```

p-value >> 0.05 **---> normalité plausible pour P1 et P2**

```r
layout(matrix(1:2,1,2)) #divise la sortie graphique en deux
# groupe A
qqnorm(proc1, main="Procédé 1")
qqline(proc1, col="red", lwd=2)

# groupe B
qqnorm(proc2, main="Procédé 2")
qqline(proc2, col="red", lwd=2)
```



Ex. 7 - Réponse - b)

Les variances sont-elles égales?

$$H_0 : \sigma_1 = \sigma_2, \quad H_1 : \sigma_1 \neq \sigma_2$$

avec $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$,

alors $F_0 = \dfrac{S_1^2}{S_2^2}$, on rejette H_0 si

$$F_0 < F_{1-\alpha/2; n_1-1, n_2-1} = \frac{1}{F_{0.05/2;11,17}} = 0.3484$$

ou

$$F_0 > F_{\alpha/2; n_1-1, n_2-1} = F_{0.05/2;11,17} = 2.870$$

c'est-à-dire si $p\text{-}value < \alpha = 0.05$

```
var.test(proc1, proc2, alternative = "two.sided", conf.level = 0.95)
```

```
        F test to compare two variances

data:  proc1 and proc2
F = 0.1284, num df = 11, denom df = 17, p-value = 0.001375
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.04474368 0.42135658
sample estimates:
ratio of variances
         0.1283982
```

F < 0.3484

p-value < 0.05 ---> **non, on rejette l'hypothèse que leurs variances sont égales**

Ex. 7 - Réponse - c)

Les moyennes sont-elles égales?

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$$

avec $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ et $\sigma_1 \neq \sigma_2$,

alors $T_0 = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$, on rejette H_0 si

$$|T_0| > t_{\alpha/2;\nu}$$

$$\text{où } \nu = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$$

c'est-à-dire si $p\text{-}value < \alpha = 0.05$

```
t.test(proc1, proc2, var.equal = FALSE)
```

```
        Welch Two Sample t-test

data:  proc1 and proc2
t = -0.88263, df = 22.868, p-value = 0.3866
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.9476028  0.3809362
sample estimates:
mean of x mean of y 
 4.583333  4.866667 
```

p-value >> 0.05 **---> on ne peut pas rejeter l'hypothèse que leurs moyennes sont égales**