olytechnique Montréal
Département de mathématiques et de génie industriel
MTH2302D - Probabilités et statistique
Examen final - Automne 2016

Question n° 1 : (8 points) 4

Soit Z_1, Z_2, Z_3, Z_4, Z_5 un échantillon aléatoire de la population $Z \sim N(0,1)$.

a) **(2 points)** On définit $W = 4Z_1 - Z_2 + 3Z_3 - 3Z_4 + Z_5$. Calculer $P(W \leq -9)$.

1

$4Z_1 \sim N(0,4)$ $Z_2 \sim N(0,1)$ $3Z_3 \sim N(0,3)$

$W \sim N(0,4)$ $P(W \leq -9)$

$P(Z \leq \frac{-9-0}{4}) = P(Z \leq -2{,}25) = 1 - P(Z \leq 2{,}25$
$= 1 - \Phi(2{,}25)$
$= 1 - 0{,}98778$
$= 0{,}01222$
$1{,}2\%$

0,98778

b) **(2 points)** Pour vérifier si la moyenne théorique de Z est bien égale à 0, on utilise la moyenne échantillonnale $\bar{Z}$ comme estimateur du paramètre $\theta = E(Z)$. Calculer l'erreur quadratique moyenne de $\bar{Z}$.

0

$EQM(\bar{Z}) = E[(\bar{Z} - Z)^2]$
$= \cancel{E(Z^2 - 2}\ E(\bar{Z}) + (Biais(\bar{Z}))^2$
$\bar{Z} + (Biais(\bar{Z}))^2$

Question n° 1 : (suite)

c) **(2 points)** Déterminer la valeur de $x > 0$ telle que

$$P\left(-x \leq \frac{Z_1}{\sqrt{Z_2^2 + Z_3^2}} \leq x\right) = 0{,}90.$$

1

avec val absolu

$Z_1 \sim N(0,1)$ $\quad Z_2^2 \sim \chi_1^2$ $\quad Z_3^2 \sim \chi_1^2$

$Z_2^2 + Z_3^2 = B \sim \chi_2^2$

$D = \left|\frac{Z_1}{\sqrt{B}}\right| \leq x$

on a $\frac{N}{\sqrt{\chi_2^2}}$

$C = \frac{Z_1}{\sqrt{B}} \sim t_2$

$P(D \leq x)$

$1 - P(D \geq x)$

$t_{0,10,2}$

d) **(2 points)** Déterminer la valeur de d telle que $P(Z_1^2 + Z_2^2 + Z_3^2 \leq d) = 0{,}95$.

2

$Z_1 \sim N(0,1)$ $\quad Z_1^2 + Z_2^2 + Z_3^2 = A \sim (\chi_3^2)$

$Z_1^2 \sim \chi_1^2$

$P(A \leq d) = 1 - P(A \geq d) = 0{,}95$

$P(A \geq d) = 0{,}05$

$\boxed{d = 7{,}81}$

Question n° 2 : (7 points)

0,5

Dans un cours d'informatique, le professeur s'intéresse au temps X (en heures) nécessaire à la réalisation d'un devoir de programmation. Le tableau d'effectifs suivant présente un échantillon aléatoire de X obtenu dans un sondage auprès de $n = 100$ étudiants.

Classe	$[0, 2)$	$[2, 5)$	$[5, 10]$
Effectif	25	18	57

a) (2 points) Calculer approximativement la moyenne et la médiane de cet échantillon en utilisant le point milieu des classes.

0

$E = ((2-0)/2 \times 25 + (5-2)/2 \times 18 + (10-5)/2 \times 57)\frac{1}{100}$

$= (25 + 45 + 142.5)/3 \quad = 70,8 = \bar{X}$

nombre impair

$\tilde{X} = x\left(\frac{n+1}{2}\right) = 18$

b) (2 points) Soit p le paramètre désignant la probabilité de compléter le devoir en moins de 5 heures. Donner une estimation ponctuelle et un intervalle de confiance de niveau 95% pour p.

0

$\alpha = 0,05$

IC

$L = -Z_\alpha \frac{\sigma}{\sqrt{n}} + \bar{X}$

$U = Z_\alpha \frac{\sigma}{\sqrt{n}} + \bar{X}$

Question n° 2 : (suite)

c) (3 points) Peut-on rejeter l'hypothèse que X suit une loi uniforme sur l'intervalle $[0, 10]$? Formuler les hypothèses testées H_0 et H_1. Utiliser un test d'hypothèse approprié au seuil $\alpha = 0{,}05$.

0.5

0,05

$H_0 : X \sim U(0,10)$ $\quad$ $H_1 : X \nsim U(0,10)$

$\nu = 2$

$\chi^2_{0{,}05,\nu} = \chi^2_{0{,}05,2}$

Question n° 3 : (4 points) 3/4

On a dressé un tableau de contingence afin de déterminer s'il existe un lien entre deux variables qualitatives X et Y. Ce tableau possède deux lignes et deux colonnes (i.e. chaque variable a 2 modalités).

a) **(1 point)** Écrire les hypothèses testées H_0 et H_1.

test independance

H_0 : X et Y sont independente

H_1 : X et Y ne sont pas independante ✓

b) **(2 points)** La statistique du test est égale à 2,4108. Doit-on rejeter H_0 au seuil $\alpha = 0{,}05$?

$\chi_0^2 = 2{,}4108$ $\alpha = 0{,}05$

$\chi^2_{\alpha,v} = \chi^2_{0{,}05,1} = 3{,}84$ $v = (r-1)(c-1) = 1$

le test dit : on rejette H_0 si $\chi_0^2 > \chi^2_{\alpha,v}$

or ici $\chi_0^2 < \chi^2_{0{,}05,1}$

donc on accepte H_0.

on ne rejette pas. ✓

(1 point) À l'aide des centiles fournis à l'annexe, déterminer une borne supérieure et une borne inférieure pour la valeur-P du test.

$\chi_0^2 = 2{,}41$

$[\quad , \chi^2_{0{,}05,1}]$

Question n° 4 : (12 points)

Deux machines (M_1 et M_2) servent à produire un type de pièce dont le diamètre est une caractéristique importante. Le tableau ci-dessous présente les mesures de diamètre et quelques statistiques pour deux échantillons de pièces produites par les deux machines (un échantillon par machine)

Machine	Mesures de diamètre en mm							Taille	Moyenne	Écart-type
M_1	39,5	41,7	38,1	39,5	40,1			5	39,78	1,30
M_2	38,6	40,1	40,2	42,8	39,5	40,8	41,9	7	40,56	1,42

On suppose que les deux échantillons sont indépendants et que les mesures de diamètre X_1 et X_2 des pièces des deux machines suivent respectivement des lois normales $N(\mu_1, \sigma_1^2)$ et $N(\mu_2, \sigma_2^2)$.

Important *: formuler les hypothèses H_0 et H_1 pour chaque test.*

a) **(2 points)** Calculer un intervalle de confiance pour la variance du diamètre des pièces de la machine M_1 au niveau de confiance 95%.

+2

$\alpha = 0,05$

IC pour σ^2 μ inconnu $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

l'intervalle sera sous la forme $\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$

on calcule $S^2 = \frac{1}{4}\left((39,5-39,78)^2 + (41,7-39,78)^2 + (38,1-39,78)^2 + (39,5-39,78)^2 + (40,1-39,78)^2\right)$

$= \frac{1}{4}(0,0784 + 3,6864 + 2.8224 + 0,0784 + 0,1024) =$

$=0,05$

$\cdot 1.692 = 6,768$

$IC = \left[\frac{4\times 1.692}{\chi^2_{0,025,4}}, \frac{4 \times 1,692}{\chi^2_{1-0,025,n-1}}\right] = \left[\frac{6,768}{11.14}, \frac{6,768}{0,48}\right]$

4

$= [0,6075 ; 14.1]$

b) **(2 points)** On compte effectuer une sixième mesure de diamètre d'une nouvelle pièce produite par M_1. Calculer un intervalle de prévision pour cette mesure à un niveau de confiance 95 %; interpréter brièvement le résultat.

$\alpha = 0,05$

pour X_{n+1} 6ème mesure σ^2 n'est pas con

$025,4 = 2,78$

$X_{n+1} = \bar{X} \pm t_{\alpha/2,n-1}\sqrt{S^2\left(1+\frac{1}{n}\right)}$

$= 39.78 \pm 2.78\sqrt{(1.30)^2\left(1+\frac{1}{5}\right)} = 39.78 \pm 3.95$

$IP = [35,83 ; 43,73]$

Question n° 4 : (suite)

c) **(2 points)** Au seuil $\alpha = 0{,}05$, peut-on conclure que les pièces de M_1 présentent en moyenne un diamètre inférieur à 40 mm ?

$\frac{1}{2}$

$H_0 : \mu = 40$ $\qquad H_1 : \mu < 40.$

σ inconnu $\qquad T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{39.78 - 0}{1.30\sqrt{5}} = \frac{39.78}{2.9068} = \cancel{13.68}$ $\quad -0{,}5$

$-t_{\alpha; n-1} = -2{,}13$ $\qquad T_0 > -t_{\alpha, n-1}$

$t_{0{,}05; 4} = 2{,}13$ $\qquad$ On accepte H_0.

conclusi $-0{,}5$

d) **(3 points)** Au seuil $\alpha = 0{,}05$, peut-on dire que les variances des diamètres des pièces sont différentes pour les deux machines M_1 et M_2 ?

$\frac{2{,}5}{3}$

$H_0 : \sigma_1^2 = \sigma_2^2$ $\qquad H_1 : \sigma_1^2 \neq \sigma_2^2$

$F_0 = \frac{S_1^2}{S_2^2} = \left(\frac{1.30}{1.42}\right)^2 = 0{,}838$

$F_{\alpha/2, n_1-1; n_2-1}$ $\qquad F_{0{,}025; 4, 6} = 6{,}23$.

$F_0 < F_{\alpha/2, n_1-1, n_2-1}$

$F_{1-\alpha/2, n_1-1, n_2-1}$ $\qquad F_{0{,}975, 4, 6} = \frac{1}{F_{0{,}025, 6, 4}} = \frac{1}{9{,}20}$

$= 0{,}108$

$F_0 > F_{1-\alpha/2, n_1-1, n_2-1}$

On ne rejette pas H_0.

conclusi $-0{,}5$

Question n° 4 : (suite)

e) (3 points) Au seuil $\alpha = 0{,}05$, peut-on dire que les moyennes des diamètres des pièces sont différentes pour les deux machines M_1 et M_2 ?

$\frac{2,5}{3}$

faux

OK !

~~d'après le test de la question precedente $\sigma_1^2 = \sigma_2^2$ utile pour ce test~~

$$\cancel{H_0 : \mu_1 - \mu_2 = 0 \qquad H_1 : \mu_1 - \mu_2 \neq 0}$$

$$\cancel{S_p = \sqrt{\frac{(m+1)S_x^2 + (n-1)S_y^2}{m+n-2}} = \sqrt{\frac{6(1.30)^2 + 6(1.42)^2}{10}} = \sqrt{\frac{10{,}14 + 12.0984}{10}} = 1{,}491}$$

$$\cancel{T_0 = \frac{(\bar{X} + \bar{Y} - \Delta)}{S_p\sqrt{\frac{1}{m} + \frac{1}{n}}} = \frac{39{,}78 + 40{,}56 - 0}{1{,}491\sqrt{\frac{1}{5} + \frac{1}{7}}} = \frac{80{,}34}{0{,}873040205} = 92.023}$$

~~Si $|T_0| > t_{\alpha/2, m+n-2}$ $\quad |T_0| > t_{\alpha/2; m+n-2}$~~

~~$t_{0{,}025, 10} = 2.23$~~

~~On rejette H_0.~~

la difference n'est pas le seul moyen, le meilleur c'est deux moyenne $\quad H_0 : \mu_1 = \mu_2 \qquad H_1 : \mu_1 \neq \mu_2$

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

On verifie si $|T_0| > t_{\alpha/2, n_1 + n_2 - 2}$

$$T_0 = \frac{39{,}78 - 40{,}56}{1.37\sqrt{\frac{1}{5} + \frac{1}{7}}} = \frac{-0{,}78}{0{,}80218986} = -0{,}972 \qquad S_p^2 = \frac{4 \times (1{,}30)^2 + 6(1.42)^2}{10}$$

$$S_p = \sqrt{1{,}88584} = \underline{1{,}37}$$

$t = 2{,}23$

$\underset{0{,}972}{|T_0|} < 2{,}23$

On accepte H_0.

Conclusion (−0,5

Question n° 5 : (14 points)

8

Un ingénieur étudie le temps de transfert de fichiers de grande taille sur un réseau informatique. Il construit un modèle de régression linéaire simple d'équation

$$Y = \beta_0 + \beta_1 x + \epsilon,$$

où x représente la taille d'un fichier (en Go, giga-octets) et Y, le temps de transfert (en minutes). De plus, β_0 et β_1 sont des paramètres, et ε une erreur aléatoire que l'on suppose de loi $N(0,\ \sigma^2)$. Les données recueillies pour estimer les paramètres du modèle sont fournies dans le tableau suivant.

i	1	2	3	4	5	6	7	8	9	10
x_i	50	50	100	100	150	150	200	200	250	250
y_i	7,4	7,3	14,9	15,1	21,9	22,2	29,4	29,1	36,5	37,0
i	11	12	13	14	15	16	17	18	19	20
x_i	300	300	350	350	400	400	450	450	500	500
y_i	43,8	44,3	51,7	52,8	61,5	59,2	66,9	70,7	73,6	78,4

L'ingénieur trouve que

$$\bar{x} = 275{,}00\ ; \quad S_{xx} = 412\,500\ ; \quad SS_R = 9\,656{,}2125;$$
$$\bar{y} = 41{,}185\ ; \quad S_{yy} = 9\,685{,}0255.$$

a) **(4 points)** Calculer $\hat{\beta}_0$ et $\hat{\beta}_1$. Donner l'équation de la droite des moindres carrés.

1.5

$\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}} =$ vs

$SS_R = \hat{\beta}_1 S_{xx}$

$\hat{\beta}_1 = \dfrac{SS_R}{S_{xx}} = \dfrac{9656.2125}{412500} = 0{,}023409$

$0{,}023$

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 41{,}185 - 0{,}023409 \times (275) = 41{,}185 - 6{,}437475$
$= 34{,}747525$

$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x + \cancel{\varepsilon}$

$\hat{Y} = 34{,}747525 + 0{,}023409\, x + \cancel{\varepsilon}$

Question n° 5 : (suite)

b) **(3 points)** Compléter le tableau d'analyse de la variance ci-dessous. Quelle conclusion peut-on en tirer au seuil $\alpha = 0{,}05$?

3

$SS_R = 9656.2125$

$SS_T = 9685.0255$

Source de variation	Somme des carrés	Degrés de liberté	Moyenne des carrés	F_0
Régression	9 656,2125	1	9656.2125	6035,1328
Erreur	28.813	18	1.60	
Totale	9 685,0255	19		

$SS_E = SS_T - SS_R = 9685{,}0255 - 9656{,}2125$ $\qquad n-2 = 18$

$MS_E = \frac{SS_E}{n-2} = \frac{28.813}{18} = 1.60$ $\qquad MS_R = \frac{SS_R}{1} = 9656.2125$

$F_0 = MS_R / MS_E = 9656.2125 / 1.60$

$H_0 : \beta_1 = 0$ $\qquad H_1 : \beta_1 \neq 0$.

$F_0 = 6035{,}1328$ $\qquad F_{0{,}05,1,18} = 4{,}41$ $\qquad f_0 > F_{0{,}05,1,18}$

On rejette H_0, c-a-d que la variation de x influe y (regression significative)

c) **(2 points)** En négligeant des délais minimaux indépendants de la taille des fichiers, l'ingénieur s'attendrait à ce que $\beta_0 = 0$. Peut-on rejeter cette hypothèse ? Justifier à l'aide d'un test approprié au seuil $\alpha = 0{,}05$.

2

on a $H_0 = \beta_0 = 0$ $\qquad H_1 : \beta_0 \neq 0$.

la statistique : $t_0 = \frac{(\hat{\beta}_0 - \beta_{0,0})}{\sqrt{MS_E\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} = \frac{34{,}74}{\sqrt{1.60\left(\frac{1}{20} + \frac{(275)^2}{412500}\right)}}$ $\qquad \beta_{0,0} = 0$.

$= \frac{34{,}74}{0{,}611010092} = 56{,}86$

$\alpha = 0{,}05$

$\alpha/2 = 0{,}025$

$t_{\alpha/2 ; n-2} : t_{\alpha/2, 18} = 2.10$.

On rejette cette hypothèse $\quad t_0 > t_{\alpha/2, n-2}$

Question n° 5 : (suite)

d) **(2 points)** Quelle valeur ponctuelle de temps de transfert l'ingénieur peut-il prévoir pour un fichier de 5 giga-octets ? Interpréter ce résultat.

0,5

$x_0 = 5$

$y_0 = \hat{y}_0 \pm t_{\alpha/2, n-2} \sqrt{MS_E\left(1/n \; \frac{(\bar{X}-x_0)^2}{S_{XX}}\right.}$

$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 = 34,7476125 + 0,023 \times 5 = 34,8625$

$y_0 = 34,86 \pm 2.10\sqrt{1.60\left(1/n \left(\frac{39.78-5}{412500}\right)\right)}$

On obtiendra l'intervale de prediction, si on veut un unilateral ça sera $\hat{Y}_0 + t_{1-\alpha/2, n-2}\sqrt{MS_E\left(1/n \; \frac{\bar{X}-x_0}{S_{xx}}\right)}$ à dr

$\hat{Y}_0 - t_{1-\alpha/2, n-2}\sqrt{MS_E\left(1/n \; \frac{\bar{X}-x_0}{S_{xx}}\right)}$ à ga

e) **(1 point)** Calculer le coefficient de détermination R^2 et interpréter ce résultat.

1 — $R^2 = SS_R / S_{yy} = 9656.2125 / 9685,0255 = 0,9970$

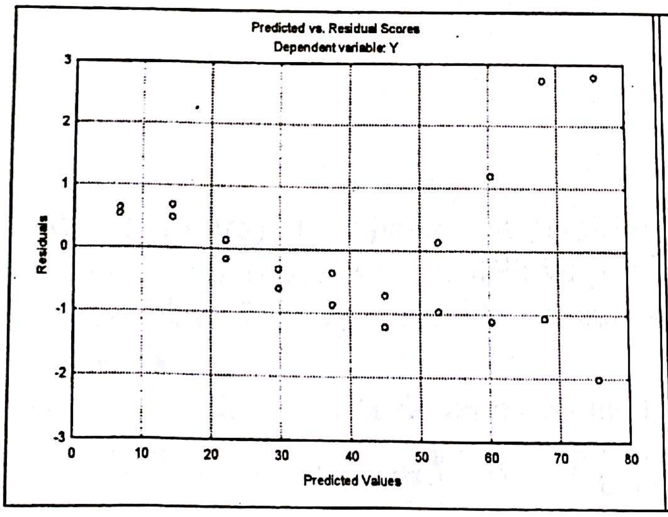
d'après le calcul fait on peut déduire que le modèle traité est adéquat et la regression est significative parce que la regle dit que quand le R^2 s'approche de 1 ce qui est le cas 0,99,7 (99,7%), la regression est significative.

adéquate

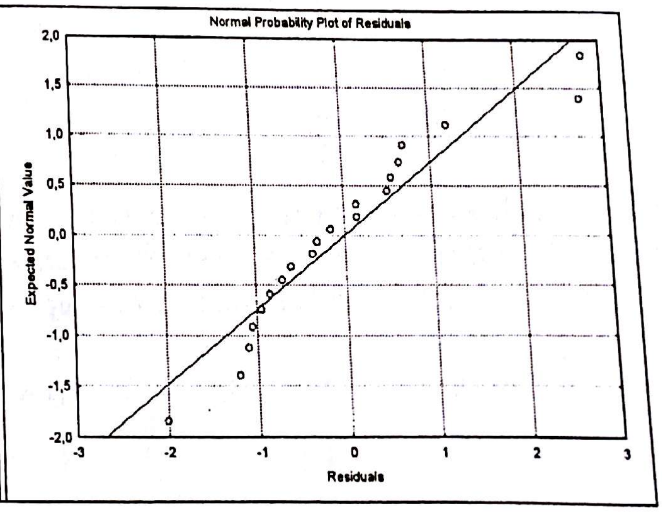
f) **(2 points)** Les graphiques de la page suivante ont été produits avec Statistica. Que peut-on dire sur le modèle de régression en analysant les résidus ?

0

Question nº 5 : (suite)



1) Résidus en fonction des valeurs prédites.



2) Graphe de probabilité normale des résidus.

1

Question nº 6 : (5 points)

Cette question est constituée de deux parties indépendantes l'une de l'autre.

a) **(2 points)** Des clients se présentent à un guichet automatique en formant une file d'attente $M/M/1$ avec $\lambda = 27$ clients par heure. On estime que 90% des clients sont présents plus de 3 minutes au total pour l'attente et le service.

0.5

Déterminer le nombre moyen de clients dans le système à l'équilibre, $\bar{N}$.

$\cdot E(T_n) = 1/\lambda$

$\bar{N}$

$\lambda_e = 27\,c/h$

$\bar{T} = 3$ min

Selon la loi de little ~~[illegible]~~ $\bar{N} = \lambda_e \bar{T}$

$\bar{N}_Q = \lambda_e \bar{T}_Q$ et $\bar{N}_S = \lambda_e \bar{T}_S$

On doit tout d e m faire la conversion 27 client/~~min~~ h

US

$\bar{T} = 3$ min

$T = 3$ min

$\lambda = 27 \times \frac{1}{60} = 0{,}45\,c/min$

$\bar{N} = \lambda_e \bar{T} = 0{,}45 \times 3 = 1{,}35 \simeq \cancel{2}$

le nombre de client total dans le système en equilibre. est 2.

b) **(3 points)** Un système en parallèle est constitué de trois composants dont les durées de vie respectives sont modélisées par les variables aléatoires indépendantes $T_1 \sim \text{Exp}(2)$, $T_2 \sim \text{Exp}(5)$ et $T_2 \sim \text{Exp}(8)$. On suppose que ces composants sont utilisés en redondance active.

0,5

Déterminer la durée de vie moyenne du système. $\lambda_1 = 2$ $\lambda_2 = 5$ $\lambda_3 = 8$

Redondance active $R(t) = 1 - \prod_{k=1}^{n} (1 - R_k(t))$

on désire $T = E(T)$

Question n° 6 : (suite)

On est face a des loi continues

$$J = E(T) = \int_0^\infty t\, f(t)\, dt = \int_0^\infty R(t)$$

$f(t) = \lambda e^{-\lambda t}$ $\quad R(t) = 1 - \int_0^t f(s)\, ds$.

$E(T) = 1/\lambda$ $\quad P(T_1 \cup T_2 \cup T_3)$

$$E(T_i) = \sum 1/\lambda_i$$

$$E(T) = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = 1/2 + 1/5 + 1/8$$

$$= 0{,}5 + 0{,}2 + 0{,}125$$

$$= 0{,}825$$

Scanned by CamScanner

[Inter]valles de confiance $(1-\alpha)$

Une moyenne μ

- σ^2 connue, $X \sim N(\mu,\sigma^2)$ ou n grand : $\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
- σ^2 inconnue et $X \sim N(\mu,\sigma^2)$: $\overline{X} \pm t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}$
- σ^2 inconnue et n grand : $\overline{X} \pm z_{\alpha/2}\frac{S}{\sqrt{n}}$

Une variance σ^2

- $X \sim N(\mu,\sigma^2)$, μ connue : $S_\mu^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2$

$$\left[\frac{nS_\mu^2}{\chi^2_{\alpha/2;n}};\ \frac{nS_\mu^2}{\chi^2_{1-\alpha/2;n}}\right]$$

- $X \sim N(\mu,\sigma^2)$, μ inconnue : $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\overline{X})^2$

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}};\ \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}\right]$$

- pour σ, avec n grand : $\left[S/(1+\frac{z_{\alpha/2}}{\sqrt{2n}});\ \ S/(1-\frac{z_{\alpha/2}}{\sqrt{2n}})\right]$

Une proportion p

- n est très grand : $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Différence de deux moyennes $\mu_1-\mu_2$

- σ_1^2, σ_2^2 connues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ ou n_1,n_2 grands : $\overline{X}_1-\overline{X}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
- σ_1^2, σ_2^2 inconnues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ $(\sigma_1^2=\sigma_2^2)$

$$\overline{X}_1-\overline{X}_2 \pm t_{\alpha/2;n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}};\ S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

- σ_1^2, σ_2^2 inconnues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ $(\sigma_1^2\neq\sigma_2^2)$

$$\overline{X}_1-\overline{X}_2 \pm t_{\alpha/2;\nu}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}};\ \nu = \frac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1}+\frac{(S_2^2/n_2)^2}{n_2+1}}-2$$

- n_1, n_2 grands

$$\overline{X}_1-\overline{X}_2 \pm z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

- Données couplées, $D_i = X_{1,i}-X_{2,i}, i=1,\ldots,n$

$$\overline{D} \pm t_{\alpha/2;n-1}\frac{S_D}{\sqrt{n}}$$

Rapport de deux variances σ_1^2/σ_2^2

- $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$

$$\left[\frac{S_1^2}{S_2^2}F_{1-\alpha/2;n_2-1;n_1-1};\ \ \frac{S_1^2}{S_2^2}F_{\alpha/2;n_2-1;n_1-1}\right]$$

Différence de deux proportions p_1-p_2

- n_1 et n_2 grands

$$\hat{p}_1-\hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Tests paramétriques (seuil α)

Une moyenne. Critères de rejet de $H_0: \mu=\mu_0$ contre H_1

$H_1: \mu<\mu_0$	$H_1: \mu>\mu_0$	$H_1: \mu\neq\mu_0$
• σ^2 connue, $X \sim N(\mu,\sigma^2)$ ou n grand : statistique $Z_0 = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$
• σ^2 inconnue et $X \sim N(\mu,\sigma^2)$: statistique du test $T_0 = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$		
si $T_0 < -t_{\alpha;n-1}$	si $T_0 > t_{\alpha;n-1}$	si $\lvert T_0\rvert > t_{\alpha/2;n-1}$
• σ^2 inconnue et n grand : statistique du test $Z_0 = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$

Une variance. Critères de rejet de $H_0: \sigma^2=\sigma_0^2$ contre H_1

$H_1: \sigma^2<\sigma_0^2$	$H_1: \sigma^2>\sigma_0^2$	$H_1: \sigma^2\neq\sigma_0^2$
• $X \sim N(\mu,\sigma^2)$: la statistique du test $\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$		
si $\chi_0^2 < \chi^2_{1-\alpha;n-1}$	$\chi_0^2 > \chi^2_{\alpha;n-1}$	si $\chi_0^2 < \chi^2_{1-\alpha/2;n-1}$ ou si $\chi_0^2 > \chi^2_{\alpha/2;n-1}$
• n est grand $(n\geq 40)$: la statistique du test $Z_0 = \frac{S-\sigma_0}{\sigma_0/\sqrt{2n}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$

Une proportion. Critères de rejet de $H_0: p=p_0$ contre H_1

$H_1: p<p_0$	$H_1: p>p_0$	$H_1: p\neq p_0$
• si n est très grand : la statistique du test $Z_0 = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$

Deux moyennes. Critères de rejet de $H_0: \mu_1=\mu_2$ contre H_1

$H_1: \mu_1<\mu_2$	$H_1: \mu_1>\mu_2$	$H_1: \mu_1\neq\mu_2$
• σ_1^2, σ_2^2 connues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ ou n_1,n_2 grands ; statistique du test $Z_0 = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$
• σ_1^2, σ_2^2 inconnues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ $(\sigma_1^2=\sigma_2^2)$; la statistique du test $T_0 = \frac{\overline{X}_1-\overline{X}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$; avec $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$		
si $T_0 < -t_{\alpha;n_1+n_2-2}$	si $T_0 > t_{\alpha;n_1+n_2-2}$	si $\lvert T_0\rvert > t_{\alpha/2;n_1+n_2-2}$
• σ_1^2, σ_2^2 inconnues, $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$ $(\sigma_1^2\neq\sigma_2^2)$; la statistique du test $T_0 = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$; avec $\nu = \frac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1}+\frac{(S_2^2/n_2)^2}{n_2+1}}-2$		
si $T_0 < -t_{\alpha;\nu}$	si $T_0 > t_{\alpha;\nu}$	si $\lvert T_0\rvert > t_{\alpha/2;\nu}$
• n_1, n_2 grands : la statistique du test $Z_0 = \frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$		
si $Z_0 < -z_\alpha$	si $Z_0 > z_\alpha$	si $\lvert Z_0\rvert > z_{\alpha/2}$
• Données couplées, $D_i = X_{1,i}-X_{2,i}, i=1,\ldots,n$: statistique $T_0 = \frac{\overline{D}}{S_D/\sqrt{n}}$		
si $T_0 < -t_{\alpha;n-1}$	si $T_0 > t_{\alpha;n-1}$	si $\lvert T_0\rvert > t_{\alpha/2;n-1}$

Deux variances. Critères de rejet de $H_0: \sigma_1^2=\sigma_2^2$ contre H_1

$H_1: \sigma_1^2<\sigma_2^2$	$H_1: \sigma_1^2>\sigma_2^2$	$H_1: \sigma_1^2\neq\sigma_2^2$
• $X_i \sim N(\mu_i,\sigma_i^2), i=1,2$: la statistique du test $F_0 = \frac{S_1^2}{S_2^2}$		
si $F_0 < F_{1-\alpha;n_1-1;n_2-1}$	si $F_0 > F_{\alpha;n_1-1;n_2-1}$	si $F_0 < F_{1-\alpha/2;n_1-1;n_2-1}$ ou si $F_0 > F_{\alpha/2;n_1-1;n_2-1}$

La fonction de répartition d'une loi $N(0,1)$: $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\{-u^2/2\}du.$

z	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	0,91308	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
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992

α	0,25	0,10	0,05	0,025	0,010	0,005	0,001	0,0005
z_α	0,674	1,282	1,645	1,960	2,326	2,576	3,090	3,291
$z_{\alpha/2}$	1,150	1,645	1,960	2,241	2,576	2,807	3,291	3,481

Calcul de β et n (cas d'une moyenne μ)

$X \sim N(\mu,\ \sigma^2)$ avec σ^2 connue et niveau (seuil) critique α.

Hypothèses	valeur de β	valeur de n
$H_0 : \mu = \mu_0$ contre $H_1 : \mu < \mu_0$	$\beta(\mu) = \Phi\left(z_\alpha + \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$
$H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$	$\beta(\mu) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$
$H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$	$\beta(\mu) = \Phi\left(z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu - \mu_0)\sqrt{n}}{\sigma}\right)$	$n \approx \frac{(z_{\alpha/2} + z_\beta)^2\sigma^2}{(\mu - \mu_0)^2}$

Calcul de β et n (cas de deux moyennes $\mu_1,\ \mu_2$)

$X_1 \sim N(\mu_1,\ \sigma_1^2)$; $X_2 \sim N(\mu_2,\ \sigma_2^2)$ avec σ_1^2, σ_2^2 connues et niveau (seuil) critique α.

Hypothèses	valeur de β	$n = n_1 = n_2$
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_\alpha + \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$	$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_\alpha - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$	$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$
$H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$	$\beta(\mu_1, \mu_2) = \Phi\left(z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$	$n \approx \frac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\mu_1 - \mu_2)^2}$

Quelques centiles

$\chi^2_{0{,}01;1} = 6{,}63$	$\chi^2_{0{,}025;1} = 5{,}02$	$\chi^2_{0{,}05;1} = 3{,}84$	$\chi^2_{0{,}10;1} = 2{,}71$	$\chi^2_{0{,}25;1} = 1{,}32$	$\chi^2_{0{,}50;1} = 0{,}45$
$\chi^2_{0{,}025;2} = 7{,}38$	$\chi^2_{0{,}05;2} = 5{,}99$	$\chi^2_{0{,}025;3} = 9{,}35$	$\chi^2_{0{,}05;3} = 7{,}81$	$\chi^2_{0{,}025;4} = 11{,}14$	$\chi^2_{0{,}05;4} = 9{,}49$
$\chi^2_{0{,}025;5} = 12{,}83$	$\chi^2_{0{,}05;5} = 11{,}07$	$\chi^2_{0{,}025;6} = 14{,}45$	$\chi^2_{0{,}05;6} = 12{,}59$	$\chi^2_{0{,}025;9} = 19{,}02$	$\chi^2_{0{,}05;9} = 16{,}92$
$\chi^2_{0{,}95;1} = 0{,}004$	$\chi^2_{0{,}95;2} = 0{,}10$	$\chi^2_{0{,}975;3} = 0{,}22$	$\chi^2_{0{,}975;4} = 0{,}48$	$\chi^2_{0{,}95;6} = 1{,}64$	$\chi^2_{0{,}975;9} = 2{,}70$
$t_{0{,}05;2} = 2{,}92$	$t_{0{,}025;4} = 2{,}78$	$t_{0{,}05;4} = 2{,}13$	$t_{0{,}025;8} = 2{,}31$	$t_{0{,}025;9} = 2{,}26$	$t_{0{,}05;9} = 1{,}83$
$t_{0{,}025;10} = 2{,}23$	$t_{0{,}05;10} = 1{,}81$	$t_{0{,}05;11} = 1{,}80$	$t_{0{,}025;12} = 2{,}18$	$t_{0{,}05;12} = 1{,}78$	$t_{0{,}025;14} = 2{,}14$
$t_{0{,}025;15} = 2{,}13$	$t_{0{,}025;18} = 2{,}10$	$t_{0{,}05;20} = 1{,}73$	$t_{0{,}025;22} = 2{,}07$	$t_{0{,}025;25} = 2{,}06$	$t_{0{,}025;26} = 2{,}06$
$F_{0{,}05;1,13} = 4{,}67$	$F_{0{,}025;1,16} = 6{,}12$	$F_{0{,}025;1,18} = 5{,}98$	$F_{0{,}05;1,18} = 4{,}41$	$F_{0{,}025;4,6} = 6{,}23$	$F_{0{,}025;6,4} = 9{,}20$
$F_{0{,}05;2,9} = 4{,}26$	$F_{0{,}05;3,10} = 3{,}71$	$F_{0{,}05;3,12} = 3{,}49$	$F_{0{,}05;3,15} = 3{,}29$	$F_{0{,}05;4,10} = 3{,}47$	$F_{0{,}05;4,12} = 3{,}26$
$F_{0{,}025;7,9} = 4{,}20$	$F_{0{,}025;9,7} = 4{,}82$	$F_{0{,}025;9,11} = 3{,}59$	$F_{0{,}025;11,9} = 3{,}91$	$F_{0{,}025;9,15} = 3{,}12$	$F_{0{,}025;13,15} = 2{,}92$

Rappel : $F_{1-\alpha;\nu_1,\nu_2} = \dfrac{1}{F_{\alpha;\nu_2,\nu_1}}$