POLYTECHNIQUE MONTRÉAL

Questionnaire Examen final

MTH2302D

Sigle du cours

Identification de l'étudiant(e)		
Nom :	Prénom :	
Signature :	Matricule :	Groupe :

Réservé	
1.	/6
2.	/6
3.	/8
4.	/12
5.	/14
6.	/4
TOTAL	/50

Sigle et titre du cours	*Groupe*	*Trimestre*
MTH2302D Probabilités et statistique	TOUS	Automne 2017
Professeur	*Local*	*Téléphone*
Luc Adjengue	A-520.33	4475

Jour	*Date*	*Durée*	*Heures*
Mercredi	20 décembre 2017	2h30	9h30 à 12h00

Documentation	*Calculatrice*	
☐ Aucune ☐ Toute ☒ Voir directives particulières	☐ Aucune ☐ Toutes ☒ Non programmable	Les cellulaires, agendas électroniques ou téléavertisseurs sont interdits.

Directives particulières

1. Ne détachez aucune feuille de ce questionnaire.
2. Les 3 dernières pages constituent une annexe (ne pas détacher).
3. Documentation permise : **une feuille 8,5x11 recto-verso.**
4. Chaque réponse doit être complète et accompagnée d'une justification.

Par souci d'équité envers tous les étudiants, le professeur ne répondra à aucune question durant cet examen. Si vous estimez que vous ne pouvez pas répondre à une question pour diverses raisons (données manquantes, données erronées, etc.), veuillez le justifier (maximum 2 lignes) puis passez à la question suivante.

Important

Cet examen contient 6 questions sur un total de 16 pages **(excluant cette page et l'annexe)**

La pondération de cet examen est de 50 %

Vous devez répondre sur : ☒ le questionnaire ☐ le cahier ☐ les deux

Vous devez remettre le questionnaire : ☒ oui ☐ non

L'étudiant doit honorer l'engagement pris lors de la signature du code de conduite.

$\phi\left(\frac{x-\mu}{\sigma}\right) -$

Question nº 1 : (6 points)

$\boxed{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)}$

Soit $X_1, X_2, \ldots, X_{100}$ un échantillon aléatoire d'une population X de loi normale de moyenne $\mu = 1$ et de variance $\sigma^2 = 1$, c'est-à-dire $X \sim N(1,\ 1)$.

$\mu \quad \sigma^2$

$\boxed{\overline{X} \sim N(1, 1/100)}$

a) Calculer $P\left(1 \leq \overline{X} \leq 1{,}05\right)$.

$$P(1 \leq \overline{X} \leq 1{,}05) = \phi\left(\frac{1{,}05-1}{0{,}10}\right) - \phi\left(\frac{1-1}{0{,}10}\right)$$

$$= \phi(0{,}5) - \phi(0)$$

$$= 0{,}69146 - 0{,}5 = \underline{0{,}19146}$$

b) On définit deux variables T et U par

$$T = \sum_{i=1}^{6} X_i \quad \text{et} \quad U = \sum_{i=7}^{9} X_i.$$

Calculer $P(T > U)$. On trouve $T-U$ pour avoir à calculer $P(T-U>0)$

$\boxed{T \sim N(6,6) \quad U \sim N(3,3) \quad T-U \sim N(6-3, 6+3)}$

$\underline{T-U \sim N(3,9)}$

$$P(T>U) = 1 - \phi\left(\frac{0-3}{3}\right) = 1 - \phi(-1)$$

$$= 1-(1-\phi(1))$$

$$= \phi(1)$$

$$= \underline{0{,}84134}$$

$\boxed{\phi(-z) = 1-\phi(z)}$

Question n° 1 : (suite)

1 a une variance de 0 et une moyenne de 1

$X_i - 1 \sim N(0,1)$?

c) On définit $Y = \sum_{i=1}^{100}(X_i - 1)^2$.

1.c) Donner le nom de la loi exacte de Y ainsi que les valeurs de ses paramètres.

Cela voudrait dire que $Y = \sum_{i=1}^{100} Z^2$ où $Z \sim N(0,1)$

Donc, $Y \sim \chi^2_{100}$

2.c) Donner le nom de la loi approximative de Y selon le théorème central limite ainsi que les valeurs de ses paramètres.

χ^2_{99}

$Z = X_i - 1$

$Z \sim N(0,1)$

$$\sum_{n=1}^{100}\left(\frac{Z_i - \bar{Z}}{1}\right)^2 \sim \chi^2_{99}$$

d) On définit deux variables G et H par

$$G = \sum_{i=1}^{6}(X_i - 1)^2 \quad \text{et} \quad H = \sum_{i=7}^{9}(X_i - 1)^2.$$

Déterminer la valeur de la constante m telle que $P(G > mH) = 0{,}01$.

$G \sim \chi^2_6 \quad H \sim \chi^2_3 \quad \alpha = 0{,}01$

$$F = \frac{\frac{\chi^2_6}{6}}{\frac{\chi^2_3}{3}} = \frac{\chi^2_6}{6} \cdot \frac{3}{\chi^2_3} = \frac{\chi^2_6}{2\chi^2_3} = \frac{G}{2H}$$

$F \sim F_{6,3}$

$$P(G > mH) = P\left(\frac{G}{H} > m\right) = P\left(\frac{G}{2H} > \frac{m}{2}\right) = P\left(F > \frac{m}{2}\right)$$

$$P\left(F > \frac{m}{2}\right) = 0{,}01 = 27{,}91 = \frac{m}{2}$$

$$m = 27{,}91 \cdot 2$$

$$m = 55{,}82$$

Question n° 2 : (6 points)

Le tableau ci-dessous présente la distribution de fréquence du nombre de défauts X sur le fini d'une pièce observés dans un échantillon aléatoire de 100 pièces d'une production.

Nombre de défauts (x_i)	0	1	2	3
Nombre de pièces (O_i)	15	45	25	15

$45+50+45 = 140/100 = 1,4$

En utilisant le test d'ajustement du Khi-deux, on veut tester l'hypothèse selon laquelle la variable X est distribuée selon une loi de Poisson.

a) **(1 point)** Formuler les hypothèses H_0 et H_1 du test à effectuer.

H_0: X est distribuée selon une loi de poisson

H_1: X n'est pas distribuée selon une loi de poisson

b) **(5 points)** Effectuer le test et conclure au seuil $\alpha = 5\%$.

$X \sim Poi(\lambda)$

$\lambda = EX = moy =$

Rappel : si $X \sim \text{Poisson}(c)$, alors $P(X = x) = \dfrac{c^x e^{-c}}{x!}$, pour $x = 0, 1, \ldots$

$$\chi_0^2 = \sum_{i=1}^{4} \left(\frac{(O_i - E_i)^2}{E_i} \right)$$

$$P(X=0) = \frac{c^0 e^{-c}}{0!} = e^{-c}$$

$$E_1 = 100e^{-c} = 100e^{-1,4} = \underline{24,66}$$

$$P(X=1) = ce^{-c}$$

$$E_2 = 100ce^{-c} = 100(1,4)e^{-1,4} = \underline{34,52}$$

$$P(X=2) = \frac{c^2 e^{-c}}{2}$$

$$E_3 = 50c^2e^{-c} = 50(1,4)^2e^{-1,4} = 24,17$$

$$E_4 = \frac{50}{3}1,4^3e^{-1,4} = 11,28$$

$$P(X=3) = \frac{c^3 e^{-c}}{6}$$

$\chi_0^2 = 8,22$ on compare avec

$$\chi^2_{4-1-1} = \chi^2_2(5\%)$$

$\hookrightarrow 5,99$

$\chi_0^2 > \chi_2^2$

<u>on rejette H_0</u>

Question n° 2 : (suite)

Question n° 3 : (8 points)

Le service du contrôle de la qualité veut étudier (estimer et tester) la proportion actuelle d'unités non conformes d'une production. Pour cette étude, le service compte utiliser un échantillon de n unités prélevées au hasard de la production.

a) **(2 points)** Quelle doit être la plus petite valeur de n pour que, avec un niveau de confiance de 95%, l'erreur d'estimation soit inférieure à 0,045 ?

$$n \geq \left(\frac{Z_{\alpha/2}}{\underset{0{,}045}{e}}\right)^2 (0{,}25) \rightarrow n \geq \left(\frac{Z_{0{,}025}}{0{,}045}\right)^2 (0{,}25)$$

$$n \geq \left(\frac{1{,}96}{0{,}045}\right)^2 \cdot 0{,}25$$

$$n \geq 474{,}27$$

$$n_{min} = 475$$

Dans le passé, la proportion d'unités non conformes de la production était de 10%. Le service du contrôle de la qualité veut effectuer un test statistique, de seuil $\alpha = 5\%$, afin de vérifier si cette proportion est à présent plus élevée.

b) **(1 point)** Définir le paramètre d'intérêt et formuler les hypothèses H_0 et H_1 du test à effectuer.

Paramètre d'intérêt : proportion d'unités non conformes de la production

H_0 : demeure à 10%

H_1 : est plus élevée

$\alpha = 0{,}05$

Question n° 3 : (suite)

c) **(3 points)** Avec un échantillon de 400 unités prélevées de la production, on obtient une valeur-p (ou «*p-value*») de 0,0024 pour le test des hypothèses définies en b).

- Quelle doit être la conclusion du test ?

Comme p-value est petite, ($< 0{,}05$) on rejette H_0

- Quel est le nombre d'unités non conformes observées dans cet échantillon ?

p-value $\rightarrow P(Z > Z_0) = 0{,}0024 \rightarrow \Phi(Z_0) = 1 - 0{,}0024 = 0{,}9976$

$\Phi(Z_0) = 0{,}9976$

$Z_0 = 2{,}82$

$$Z_0 = \frac{\hat{p} - 0{,}10}{\sqrt{\frac{0{,}10 \cdot 0{,}9}{400}}} \rightarrow 2{,}82 = \frac{\hat{p} - 0{,}10}{0{,}015}$$

$0{,}0423 = \hat{p} - 0{,}10$

$\hat{p} = 0{,}1423$

$n_{nc} = 0{,}1423(400) \approx 57$

d) **(2 points)** Supposons que la proportion réelle d'unités non conformes soit actuellement de 11% et qu'on utilise un échantillon aléatoire de 500 unités pour le test. Quelle serait alors la probabilité de conclure à tort que la proportion d'unités non conformes actuelle n'est pas plus élevée que dans le passé ?

accepter H_0 / H_0 est fausse : Cela correspond à $\beta(p = 0{,}11)$

$$\beta = \Phi\left(\frac{0{,}1 - 0{,}11 + 1{,}96\sqrt{\frac{0.1 \cdot 0.9}{500}}}{\sqrt{\frac{0.1 \cdot 0.9}{500}}}\right)$$

$= \Phi(1{,}2146)$

$= \boxed{0{,}887}$

Question n° 4 : (12 points)

Dans un réseau informatique, un ingénieur mesure les délais de communication (en millisecondes) entre deux ordinateurs avec un protocole standard (population X_1) et avec un protocole expérimental (population X_2). L'ingénieur prélève un échantillon aléatoire de taille $n_1 = 16$ de la population X_1 et un échantillon aléatoire de taille $n_2 = 9$ de la population X_2. On suppose que les deux échantillons sont indépendants. Les modèles utilisés et les résultats obtenus sont décrits dans le tableau suivant.

Population X_i	Taille n_i	Moyenne $\bar{x}_i$	Variance s_i^2
$X_1 \sim N(\mu_1,\ \sigma_1^2)$	16	52,4	0,37
$X_2 \sim N(\mu_2,\ \sigma_2^2)$	9	50,3	2,18

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

a) **(2 points)** Donner un intervalle de confiance de niveau 90% pour le paramètre μ_2. $\alpha = 0,1$

σ^2 inconnue et $X \sim N(\mu, \sigma^2)$

$$\bar{X} \pm t_{\alpha/2;\, n-1} \frac{S}{\sqrt{n}}$$

$S_2 = 1,4765$

$\sqrt{n} = 3$

$\alpha/2 = 0,05$

$\bar{X}_2 = 50,3$

$n-1 = 8$

$$50,3 \pm t_{0,05;8} \cdot \frac{1,4765}{3}$$

$$50,3 \pm 1,86 \cdot 0,4922$$

$$50,3 \pm 0,91543$$

$$[49,38457,\ 51,21453]$$

Question n° 4 : (suite)

b) (2 points) Donner un intervalle de prévision de niveau 90% pour une dixième mesure de délai du protocole expérimental.

$$X_{10} \in \overline{X_2} \pm t_{0,05;8}\sqrt{2{,}18(1+1/9)}$$

$$\in 50{,}3 \pm 1{,}86 \cdot 1{,}556$$

$$\in 50{,}3 \pm 2{,}89466$$

$$[47{,}40534 ; 53{,}19466]$$

c) (2 points) Peut-on rejeter l'affirmation selon laquelle $\sigma_1^2 = 0{,}5$? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

$H_0: \sigma^2 = \sigma_1^2$
$H_1: \sigma^2 \neq \sigma_1$

Statistique $\chi_0^2 = \frac{(n-1)S_1^2}{\sigma_0^2} = \frac{15(0{,}37)}{0{,}5} = 11{,}1$

On rejette H_0 si $\chi_0^2 > \chi^2_{0{,}025;15}$ ou $\chi_0^2 < \chi^2_{0{,}975;15}$

$\chi_0^2 < 6{,}27$: Non

On ne peut pas rejeter H_0.
On ne peut donc pas nier le fait que $\sigma_1^2 = 0{,}5$, sans pouvoir le rejeter

Question n° 4 : (suite)

d) **(3 points)** Peut-on dire que les variances σ_1^2 et σ_2^2 sont différentes ? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

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

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{0{,}37}{2{,}18} = 0{,}17$$

* Attention de tester les 2 !!!

On rejette H_0 si

$F_0 < F_{0,975;15;8}$ ou si $F_0 > F_{0,025;15;8}$

ok

$$F_{0,975;15;8} = \frac{1}{F_{0,025;15;8}} = \frac{1}{4{,}10}$$

$F_0 < 0{,}244$ oui !

Donc on rejette H_0

Question n° 4 : (suite)

1 standard
2 exp

e) **(3 points)** Peut-on affirmer qu'en moyenne les délais du protocole expérimental sont plus courts que ceux du protocole standard ? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

$H_0 : \mu_1 = \mu_2 \quad H_1 : \mu_1 > \mu_2$

σ_1^2 et σ_2^2 inconnues

$\sigma_1^2 \neq \sigma_2^2$

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{2{,}1}{\sqrt{\frac{0{,}37}{16} + \frac{2{,}18}{9}}}$$

$$= \frac{2{,}1}{0{,}515} = \underline{4{,}078}$$

$$V = \frac{(S_1^2/n_1 + S_2^2/n_1)^2}{\frac{(S_1^2/n_1)^2}{17} + \frac{(S_2^2/n_2)^2}{10}} - 2 = \cancel{2{,}31}$$

$T_0 > t_{0{,}05;2{,}31}$

$T_0 > \approx 2{,}92$?

Oui

On rejette H_0. Les 2 moy ne sont pas =. ~~[illegible]~~

Il se peut que

Question nº 5 : (14 points)

Pour des raisons de santé publique, on cherche à déterminer le lien entre la concentration d'ozone O_3 dans l'air d'une journée et la température à midi de la même journée. Pour ce faire, on envisage un modèle de régression linéaire simple d'équation

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

où X représente la température à midi (en °C) et Y, la concentration d'ozone O_3 dans l'air (en microgrammes par millilitre), 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 un échantillon de 12 jours sont présentées dans le tableau suivant.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	10,0	23,8	16,3	27,2	7,1	25,1	27,5	19,4	15,0	19,8	32,2	20,7
y_i	~~71,3~~	~~115,4~~	~~76,8~~	~~113,8~~	~~81,6~~	~~115,4~~	~~125,0~~	~~83,6~~	~~84,5~~	~~75,2~~	~~136,8~~	102,8

L'ajustement du modèle aux 12 observations a permis d'obtenir les résultats partiels suivants :

$$\sum_{i=1}^{12} x_i = 244{,}10\ ; \quad \sum_{i=1}^{12} x_i^2 = 5\,567{,}37\ ; \quad \sum_{i=1}^{12} x_i y_i = 25\,631{,}34\ ;$$

$$\sum_{i=1}^{12} y_i = 1\,182{,}20\ ; \quad \sum_{i=1}^{12} y_i^2 = 121\,916{,}60.$$

$\bar{X} = 20{,}34$

$\bar{Y} = 98{,}52$

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

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{25\,631{,}34 - 12(20.34 \cdot 98{,}52)}{5\,567{,}37 - 12(20{,}34^2)} = \frac{1584{,}58}{602{,}7828}$$

$\hat{\beta}_1 = 2{,}63$

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$

$\hat{\beta}_0 = 98{,}52 - [2{,}63(20{,}34)]$

$\hat{\beta}_0 = 45{,}0258$

$Y = 45{,}0258 + 2{,}63X + \varepsilon$

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

Question nº 5 : (suite)

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

Source de variation	Somme des carrés	Degrés de liberté	Moyenne des carrés	F_0
Régression	4167,44	1	4167,44	32,49
Erreur	1282,76	10	128,276	
Totale	5 450,197	11		

$SS_R = \hat{\beta}_1 \cdot S_{xy} = \beta_1 \cdot \sum x_i y_i - n\bar{x}\bar{y} = 4167{,}44$

$MR = SSR/n = 347{,}29$

$ME = SSE/_{n-2} = 128{,}276$

$F_0 = MSR/MSE$

$F_{0{,}05;1;10} = 4{,}67$

$F_0 > F_{0{,}05;1;10}$ oui
Rejet H_0

c) **(2 points)** Donner une estimation ponctuelle de la variance σ^2.

Donc les 2 cas s'intéressent!

Estimation ponctuelle de la variance = $MSE = 128{,}276$

Question n° 5 : (suite)

d) **(3 points)** À un niveau de 95%, donner un intervalle de prévision pour la concentration d'ozone O_3 d'une journée dont la température à midi est de 20 °C.

$$IP = \hat{Y} \pm t_{\alpha/2;\,n-2} \cdot \sqrt{MSE\left[1 + \frac{1}{n} + \frac{(20-\bar{x})^2}{S_{xx}}\right]}$$

$\hat{Y} =$

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

$R^2 = \frac{SSR}{SST}$ ...

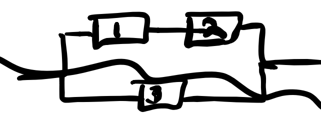
Question n° 6 : (4 points)

Cette question est constituée de deux parties a) et b) indépendantes l'une de l'autre.

a) **(2 points)** Un système est constitué de deux composants placés en série qui fonctionnent indépendamment l'un de l'autre. De plus, un troisième composant est utilisé en redondance passive, précisément :

- il est mis en fonction pour remplacer le premier composant actif qui tombe en panne et,
- il fonctionne indépendamment du composant encore en fonction.

On suppose que la durée de vie de chacun des trois composants est distribuée selon une loi exponentielle Exp(2).



1.a) Déterminer la durée de vie moyenne du système.

Durée de vie moyenne = $\sum R(t)$ ou $\int R(t)$

partie en série

$R_1(t) = e^{-4t}$

MTTF = $\int$

partie en //

$R_2(t) = \sim e^{-2t}$

2.a) Donner le nom et les valeurs des paramètres de la loi de probabilité qui modélise la durée de vie du système.

Question nº 6 : (suite)

b) **(2 points)** Un commerce possède deux caisses modélisées par une file d'attente $M/M/1$. La première caisse sert à payer les achats : les clients s'y présentent au taux $\lambda_1 = 45$ clients/heure et ils sont servis au taux $\mu_1 = 50$ clients/heure. La deuxième caisse sert à traiter les retours et échanges : les clients s'y présentent au taux $\lambda_2 = 4$ clients/heure et ils sont servis au taux $\mu_2 = 5$ clients/heure. On suppose que les deux files fonctionnent indépendamment l'une de l'autre.

1.b) Déterminer le nombre moyen de clients dans le système (aux deux caisses) à l'équilibre.

$\overline{N_1} = \overline{N_Q} + \overline{N_S}$ $\quad \rho = \frac{1/\mu}{1/\lambda}$ $\quad \rho_1 = 0,9$ $\quad \overline{N_2} = \overline{N_Q} + \overline{N_S}$ $\quad \rho_2 = \frac{1/\mu}{1/\lambda} = 0,8$

$= \frac{\rho^2}{1-\rho} + \rho$ $\quad = \frac{\rho^2}{1-\rho} + \rho$

$= 9$ $\quad \overline{N_2} = 4$

$= \underline{13}$

2.b) Déterminer la probabilité que le système soit vide à l'équilibre.

$\pi_0^1 \cdot \pi_0^2 = (1-0,9)\cdot(1-0,8) = 2\%$

Intervalles 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 : $S_\mu^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

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

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

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

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

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

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

- 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

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

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

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

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

Quelques centiles

$\chi^2_{0,01;1}=6,63$	$\chi^2_{0,025;1}=5,02$	$\chi^2_{0,05;1}=3,84$	$\chi^2_{0,10;1}=2,71$	$\chi^2_{0,25;1}=1,32$	$\chi^2_{0,50;1}=0,45$
$\chi^2_{0,025;2}=7,38$	$\chi^2_{0,05;2}=5,99$	$\chi^2_{0,025;3}=9,35$	$\chi^2_{0,05;3}=7,81$	$\chi^2_{0,025;4}=11,14$	$\chi^2_{0,05;4}=9,49$
$\chi^2_{0,025;5}=12,83$	$\chi^2_{0,05;5}=11,07$	$\chi^2_{0,025;6}=14,45$	$\chi^2_{0,05;6}=12,59$	$\chi^2_{0,025;9}=19,02$	$\chi^2_{0,05;9}=16,92$
$\chi^2_{0,95;1}=0,004$	$\chi^2_{0,95;2}=0,10$	$\chi^2_{0,975;3}=0,22$	$\chi^2_{0,975;4}=0,48$	$\chi^2_{0,975;15}=6,27$	$\chi^2_{0,025;15}=27,49$
$t_{0,05;2}=2,92$	$t_{0,025;4}=2,78$	$t_{0,05;8}=1,86$	$t_{0,025;8}=2,31$	$t_{0,025;9}=2,26$	$t_{0,05;9}=1,83$
$t_{0,025;10}=2,23$	$t_{0,05;10}=1,81$	$t_{0,05;11}=1,80$	$t_{0,025;11}=2,20$	$t_{0,05;14}=1,76$	$t_{0,025;14}=2,14$
$t_{0,025;15}=2,13$	$t_{0,05;15}=1,75$	$t_{0,025;17}=2,11$	$t_{0,025;22}=2,07$	$t_{0,05;22}=1,72$	$t_{0,025;23}=2,07$
$t_{0,05;23}=1,71$	$t_{0,05;24}=1,71$	$t_{0,05;25}=1,71$	$t_{0,05;26}=1,71$	$t_{0,05;28}=1,70$	$t_{0,05;30}=1,70$
$F_{0,25;1,1}=5,83$	$F_{0,025;1,10}=6,94$	$F_{0,025;1,18}=5,98$	$F_{0,05;1,10}=4,67$	$F_{0,025;4,6}=6,23$	$F_{0,025;6,4}=9,20$
$F_{0,01;2,4}=18,00$	$F_{0,01;6,3}=27,91$	$F_{0,01;3,6}=9,78$	$F_{0,05;3,15}=3,29$	$F_{0,05;4,10}=3,47$	$F_{0,05;4,12}=3,26$
$F_{0,025;7,9}=4,20$	$F_{0,025;9,7}=4,82$	$F_{0,025;9,11}=3,59$	$F_{0,025;11,9}=3,91$	$F_{0,025;15,8}=4,10$	$F_{0,025;8,15}=3,20$

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

Question n° 1 : (6 points)

Soit $X_1, X_2, \ldots, X_{100}$ un échantillon aléatoire d'une population X de loi normale de moyenne $\mu = 1$ et de variance $\sigma^2 = 1$, c'est-à-dire $X \sim N(1,\ 1)$.

a) Calculer $P\left(1 \leq \overline{X} \leq 1{,}05\right)$.

0.5 $\bar{X} \sim N\left(1, \frac{1}{100}\right)$

0.25 $P\left(1 \leq \bar{X} \leq 1{,}05\right) = \Phi\left(\frac{1{,}05-1}{0{,}10}\right) - \Phi\left(\frac{1-1}{0{,}10}\right)$

1

0.25
$$= \Phi(0{,}5) - \Phi(0)$$
$$= 0{,}69146 - 0{,}50$$
$$= 0{,}19146$$

b) On définit deux variables T et U par

$$T = \sum_{i=1}^{6} X_i \quad \text{et} \quad U = \sum_{i=7}^{9} X_i.$$

1,5

Calculer $P(T > U)$.

0.75 $T - U \sim N(3, 9)$

$$\underbrace{P(T>U) = P(T-U>0) =}_{0.5} \underbrace{1 - \Phi\left(\frac{0-3}{3}\right)}_{} = \underbrace{1 - \Phi(-1) = \Phi(1) = 0{,}84134}_{0.25}$$

Question n° 1 : (suite)

c) On définit $Y = \sum_{i=1}^{100}(X_i - 1)^2$.

1.c) Donner le nom de la loi exacte de Y ainsi que les valeurs de ses paramètres.

• loi exacte $Y \sim \chi^2_{100}$ ou Gamma $(r = 50, \lambda = \frac{1}{2})$

1

• ~~loi approximative $Y \underset{\text{approx.}}{\sim} N(100, 200)$~~

2.c) Donner le nom de la loi approximative de Y selon le théorème central limite ainsi que les valeurs de ses paramètres.

loi approximative $Y \underset{\text{app.}}{\sim} N(100, 200)$

1

d) On définit deux variables G et H par

$$G = \sum_{i=1}^{6}(X_i - 1)^2 \quad \text{et} \quad H = \sum_{i=7}^{9}(X_i - 1)^2.$$

Déterminer la valeur de la constante m telle que $P(G > mH) = 0{,}01$.

1,5

0,75 $\dfrac{G/6}{H/3} = \dfrac{G}{2H} \sim F_{6,3}$

0.50 $P(G > mH) = 0{,}01 \iff P\left(\dfrac{G}{2H} > \dfrac{m}{2}\right) = 0{,}01$

0.25 Donc $\dfrac{m}{2} = F_{0,01;6,3}$ i.e $m = 2 \times F_{0,01;6,3}$

$= 2 \times 27{,}91$

$= 55{,}82$

Question n° 3 : (8 points)

Le service du contrôle de la qualité veut étudier (estimer et tester) la proportion actuelle d'unités non conformes d'une production. Pour cette étude, le service compte utiliser un échantillon de n unités prélevées au hasard de la production.

a) **(2 points)** Quelle doit être la plus petite valeur de n pour que, avec un niveau de confiance de 95%, l'erreur d'estimation soit inférieure à 0,045 ?

$e = 0{,}045$

1.5 $\quad n \geq \left(\frac{z_{\alpha/2}}{e}\right)^2 \times \frac{1}{4}$ formule (9.20) p. 262

0.5 $\quad$ i.e. $n \geq \left(\frac{1{,}96}{0{,}045}\right)^2 \times \frac{1}{4} = 474{,}27$ soit $n_{min} = 475$

Dans le passé, la proportion d'unités non conformes de la production était de 10%. Le service du contrôle de la qualité veut effectuer un test statistique, de seuil $\alpha = 5\%$, afin de vérifier si cette proportion est à présent plus élevée.

b) **(1 point)** Définir le paramètre d'intérêt et formuler les hypothèses H_0 et H_1 du test à effectuer.

p : proportion actuelle d'unités non conformes

On teste $H_0: p = 0{,}10$ contre $H_1: p > 0{,}10$ $\quad (\alpha = 0{,}05)$

Question n° 3 : (suite)

c) **(3 points)** Avec un échantillon de 400 unités prélevées de la production, on obtient une valeur-p (ou «*p-value*») de 0,0024 pour le test des hypothèses définies en b).

- Quelle doit être la conclusion du test ?

1 → Puisque p-value = 0,0024 est petite (< 0,05) on rejette H_0

conclusion = 0,5 On peut conclure que la proportion actuelle est plus élevée

- Quel est le nombre d'unités non conformes observées dans cet échantillon ?

Puisque p-value = $P(Z > z_0) = 0{,}0024 \Rightarrow \Phi(z_0) = 0{,}9976$

$\Rightarrow z_0 = 2{,}82$

Donc,

1,5

$$z_0 = \frac{\hat{p} - 0{,}10}{\sqrt{\frac{0{,}10 \times 0{,}90}{400}}} = 2{,}82 \Rightarrow \hat{p} = 0{,}10 + 2{,}82 \times \sqrt{\frac{0{,}10 \times 0{,}90}{400}} = 0{,}1423$$

Le nombre d'unités non conformes est $400 \times \hat{p} = 0{,}1423 \times 400$
$= 56{,}92$

soit 57

d) **(2 points)** Supposons que la proportion réelle d'unités non conformes soit actuellement de 11% et qu'on utilise un échantillon aléatoire de 500 unités pour le test. Quelle serait alors la probabilité de conclure à tort que la proportion d'unités non conformes actuelle n'est pas plus élevée que dans le passé ?

$$.5 \; \{ \; \beta(p=0{,}11) = \Phi\left(\frac{0{,}10 - 0{,}11 + 1{,}645\sqrt{\frac{0{,}10 \times 0{,}90}{500}}}{\sqrt{\frac{0{,}11 \times 0{,}89}{500}}} \right)$$

$$0{,}25 \; \} \; = \Phi(0{,}86)$$

$$0{,}25 \; \} \; = 0{,}805$$

POLYTECHNIQUE MONTRÉAL

Questionnaire
Examen final

MTH2302D

Sigle du cours

Identification de l'étudiant(e)		
Nom :	**Prénom :**	
Signature :	**Matricule :**	**Groupe :**

Sigle et titre du cours		***Groupe***	***Trimestre***
MTH2302D Probabilités et statistique		**TOUS**	**Automne 2017**
Professeur		***Local***	***Téléphone***
Luc Adjengue		**A-520.33**	**4475**
Jour	***Date***	***Durée***	***Heures***
Mercredi	**20 décembre 2017**	**2h30**	**9h30 à 12h00**

Documentation	***Calculatrice***	
☐ Aucune	☐ Aucune	**Les cellulaires, agendas électroniques ou téléavertisseurs sont interdits.**
☐ Toute	☐ Toutes	
☒ Voir directives particulières	☒ Non programmable	

Réservé	
1.	**/6**
2.	**/6**
3.	**/8**
4.	**/12**
5.	**/14**
6.	**/4**
TOTAL	**/50**

Directives particulières

1. Ne détachez aucune feuille de ce questionnaire.
2. Les 3 dernières pages constituent une annexe (ne pas détacher).
3. Documentation permise : **une feuille 8,5x11 recto-verso.**
4. Chaque réponse doit être complète et accompagnée d'une justification.

Par souci d'équité envers tous les étudiants, le professeur ne répondra à aucune question durant cet examen. Si vous estimez que vous ne pouvez pas répondre à une question pour diverses raisons (données manquantes, données erronées, etc.), veuillez le justifier (maximum 2 lignes) puis passez à la question suivante.

Important

Cet examen contient **6** questions sur un total de **16** pages **(excluant cette page et l'annexe)**

La pondération de cet examen est de **50** %

Vous devez répondre sur : ☒ le questionnaire ☐ le cahier ☐ les deux

Vous devez remettre le questionnaire : ☒ oui ☐ non

L'étudiant doit honorer l'engagement pris lors de la signature du code de conduite.

Question n° 1 : (6 points)

Soit $X_1, X_2, \ldots, X_{100}$ un échantillon aléatoire d'une population X de loi normale de moyenne $\mu = 1$ et de variance $\sigma^2 = 1$, c'est-à-dire $X \sim N(1,\ 1)$.

a) Calculer $P\left(1 \leq \overline{X} \leq 1{,}05\right)$.

$$\frac{\overline{X}-1}{1/\sqrt{100}} \sim N(0,1)$$

$$\frac{\overline{X}-1}{0{,}10} \sim N(0,1)$$

$P(1 \leq \overline{X} \leq 1{,}05)$

$= \phi\left(\frac{1{,}05-1}{0{,}1}\right) - \phi(0)$

$= \phi(0{,}5) - \phi(0) = 0{,}69146 - 0{,}5 = \boxed{0{,}19146}$

b) On définit deux variables T et U par

$$T = \sum_{i=1}^{6} X_i \quad \text{et} \quad U = \sum_{i=7}^{9} X_i.$$

Calculer $P(T > U)$.

$T \sim N(6,6)$ $\quad U \sim N(3,3)$

$P(T > U) = P(T-U > 0)$ $\quad \underbrace{T-U}_{} \sim N(3,9)$

$P(T-U>0) = \phi\left(\frac{0-3}{3}\right) = \phi(-1)$ $\quad Z \sim N(3,9)$

$= \phi(-1)$ $\quad \frac{Z-3}{3} \sim N(0,1)$

$= 1 - \phi(1)$

$= 1 - 0{,}84134 = 0{,}15686$

Question n° 1 : (suite)

c) On définit $Y = \sum_{i=1}^{100}(X_i - 1)^2$.

1.c) Donner le nom de la loi exacte de Y ainsi que les valeurs de ses paramètres.

$X_i - 1 \sim N(0,1)$, car 1 a une moyenne de 1 et une variance de 0

$$\sum_{i=1}^{100} (X_i - 1)^2 \sim \chi^2_{100}$$

2.c) Donner le nom de la loi approximative de Y selon le théorème central limite ainsi que les valeurs de ses paramètres.

On a vu que $\sum (X_i - 1)^2$ était un match avec une χ^2. Selon le TCL cela peut être approximé avec une loi normale ayant une moyenne égale à l'espérance de la χ^2 et une variance égale à la variance de la χ^2.

$E(Y) = n = 100$

$V(Y) = 2n = 200$

donc $Y \underset{\text{approx}}{\sim} N(100, 200)$

d) On définit deux variables G et H par

$$G = \sum_{i=1}^{6}(X_i - 1)^2 \quad \text{et} \quad H = \sum_{i=7}^{9}(X_i - 1)^2.$$

Déterminer la valeur de la constante m telle que $P(G > mH) = 0{,}01$.

Question n° 2 : (6 points)

Le tableau ci-dessous présente la distribution de fréquence du nombre de défauts X sur le fini d'une pièce observés dans un échantillon aléatoire de 100 pièces d'une production.

Nombre de défauts (x_i)	0	1	2	3
Nombre de pièces (O_i)	15	45	25	15

En utilisant le test d'ajustement du Khi-deux, on veut tester l'hypothèse selon laquelle la variable X est distribuée selon une loi de Poisson.

a) (1 point) Formuler les hypothèses H_0 et H_1 du test à effectuer.

b) (5 points) Effectuer le test et conclure au seuil $\alpha = 5\%$.

Rappel : si $X \sim \text{Poisson}(c)$, alors $P(X = x) = \dfrac{c^x e^{-c}}{x!}$, pour $x = 0, 1, \ldots$

Question n° 2 : (suite)

Question n° 3 : (8 points)

Le service du contrôle de la qualité veut étudier (estimer et tester) la proportion actuelle d'unités non conformes d'une production. Pour cette étude, le service compte utiliser un échantillon de n unités prélevées au hasard de la production.

a) (2 points) Quelle doit être la plus petite valeur de n pour que, avec un niveau de confiance de 95%, l'erreur d'estimation soit inférieure à 0,045 ?

Dans le passé, la proportion d'unités non conformes de la production était de 10%. Le service du contrôle de la qualité veut effectuer un test statistique, de seuil $\alpha = 5\%$, afin de vérifier si cette proportion est à présent plus élevée.

b) (1 point) Définir le paramètre d'intérêt et formuler les hypothèses H_0 et H_1 du test à effectuer.

Question nº 3 : (suite)

c) **(3 points)** Avec un échantillon de 400 unités prélevées de la production, on obtient une valeur-p (ou «*p-value*») de 0,0024 pour le test des hypothèses définies en b).

- Quelle doit être la conclusion du test ?

- Quel est le nombre d'unités non conformes observées dans cet échantillon ?

d) **(2 points)** Supposons que la proportion réelle d'unités non conformes soit actuellement de 11% et qu'on utilise un échantillon aléatoire de 500 unités pour le test. Quelle serait alors la probabilité de conclure à tort que la proportion d'unités non conformes actuelle n'est pas plus élevée que dans le passé ?

Question n° 4 : (12 points)

Dans un réseau informatique, un ingénieur mesure les délais de communication (en millisecondes) entre deux ordinateurs avec un protocole standard (population X_1) et avec un protocole expérimental (population X_2). L'ingénieur prélève un échantillon aléatoire de taille $n_1 = 16$ de la population X_1 et un échantillon aléatoire de taille $n_2 = 9$ de la population X_2. On suppose que les deux échantillons sont indépendants. Les modèles utilisés et les résultats obtenus sont décrits dans le tableau suivant.

Population X_i	Taille n_i	Moyenne $\bar{x}_i$	Variance s_i^2
$X_1 \sim N(\mu_1,\ \sigma_1^2)$	16	52,4	0,37
$X_2 \sim N(\mu_2,\ \sigma_2^2)$	9	50,3	2,18

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

a) **(2 points)** Donner un intervalle de confiance de niveau 90% pour le paramètre μ_2.

Question n° 4 : (suite)

b) **(2 points)** Donner un intervalle de prévision de niveau 90% pour une dixième mesure de délai du protocole expérimental.

c) **(2 points)** Peut-on rejeter l'affirmation selon laquelle $\sigma_1^2 = 0{,}5$? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

Question n° 4 : (suite)

d) **(3 points)** Peut-on dire que les variances σ_1^2 et σ_2^2 sont différentes ? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

Question n° 4 : (suite)

e) **(3 points)** Peut-on affirmer qu'en moyenne les délais du protocole expérimental sont plus courts que ceux du protocole standard ? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0{,}05$.

Question n° 5 : (14 points)

Pour des raisons de santé publique, on cherche à déterminer le lien entre la concentration d'ozone O_3 dans l'air d'une journée et la température à midi de la même journée. Pour ce faire, on envisage un modèle de régression linéaire simple d'équation

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

où X représente la température à midi (en °C) et Y, la concentration d'ozone O_3 dans l'air (en microgrammes par millilitre), 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 un échantillon de 12 jours sont présentées dans le tableau suivant.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	10,0	23,8	16,3	27,2	7,1	25,1	27,5	19,4	15,0	19,8	32,2	20,7
y_i	71,3	115,4	76,8	113,8	81,6	115,4	125,0	83,6	84,5	75,2	136,8	102,8

L'ajustement du modèle aux 12 observations a permis d'obtenir les résultats partiels suivants :

$$\sum_{i=1}^{12} x_i = 244{,}10 \ ; \quad \sum_{i=1}^{12} x_i^2 = 5\,567{,}37 \ ; \quad \sum_{i=1}^{12} x_i y_i = 25\,631{,}34 \ ;$$
$$\sum_{i=1}^{12} y_i = 1\,182{,}20 \ ; \quad \sum_{i=1}^{12} y_i^2 = 121\,916{,}60.$$

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

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 $\alpha = 0{,}05$?

Source de variation	Somme des carrés	Degrés de liberté	Moyenne des carrés	F_0
Régression				
Erreur				
Totale	5 450,197			

c) **(2 points)** Donner une estimation ponctuelle de la variance σ^2.

Question n° 5 : (suite)

d) **(3 points)** À un niveau de 95%, donner un intervalle de prévision pour la concentration d'ozone O_3 d'une journée dont la température à midi est de 20 °C.

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

Question n^{o} 6 : (4 points)

Cette question est constituée de deux parties a) et b) indépendantes l'une de l'autre.

a) **(2 points)** Un système est constitué de deux composants placés en série qui fonctionnent indépendamment l'un de l'autre. De plus, un troisième composant est utilisé en redondance passive, précisément :

- il est mis en fonction pour remplacer le premier composant actif qui tombe en panne et,
- il fonctionne indépendamment du composant encore en fonction.

On suppose que la durée de vie de chacun des trois composants est distribuée selon une loi exponentielle Exp(2).

1.a) Déterminer la durée de vie moyenne du système.

2.a) Donner le nom et les valeurs des paramètres de la loi de probabilité qui modélise la durée de vie du système.

Question n° 6 : (suite)

b) **(2 points)** Un commerce possède deux caisses modélisées par une file d'attente $M/M/1$. La première caisse sert à payer les achats : les clients s'y présentent au taux $\lambda_1 = 45$ clients/heure et ils sont servis au taux $\mu_1 = 50$ clients/heure. La deuxième caisse sert à traiter les retours et échanges : les clients s'y présentent au taux $\lambda_2 = 4$ clients/heure et ils sont servis au taux $\mu_2 = 5$ clients/heure. On suppose que les deux files fonctionnent indépendamment l'une de l'autre.

1.b) Déterminer le nombre moyen de clients dans le système (aux deux caisses) à l'équilibre.

2.b) Déterminer la probabilité que le système soit vide à l'équilibre.

Intervalles 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 : $\left[S/(1+\frac{z_{\alpha/2}}{\sqrt{2n}});\quad S/(1-\frac{z_{\alpha/2}}{\sqrt{2n}})\right]$

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

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

- 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

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

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

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

$\chi^2_{0,01;1} = 6{,}63$	$\chi^2_{0,025;1} = 5{,}02$	$\chi^2_{0,05;1} = 3{,}84$	$\chi^2_{0,10;1} = 2{,}71$	$\chi^2_{0,25;1} = 1{,}32$	$\chi^2_{0,50;1} = 0{,}45$
$\chi^2_{0,025;2} = 7{,}38$	$\chi^2_{0,05;2} = 5{,}99$	$\chi^2_{0,025;3} = 9{,}35$	$\chi^2_{0,05;3} = 7{,}81$	$\chi^2_{0,025;4} = 11{,}14$	$\chi^2_{0,05;4} = 9{,}49$
$\chi^2_{0,025;5} = 12{,}83$	$\chi^2_{0,05;5} = 11{,}07$	$\chi^2_{0,025;6} = 14{,}45$	$\chi^2_{0,05;6} = 12{,}59$	$\chi^2_{0,025;9} = 19{,}02$	$\chi^2_{0,05;9} = 16{,}92$
$\chi^2_{0,95;1} = 0{,}004$	$\chi^2_{0,95;2} = 0{,}10$	$\chi^2_{0,975;3} = 0{,}22$	$\chi^2_{0,975;4} = 0{,}48$	$\chi^2_{0,975;15} = 6{,}27$	$\chi^2_{0,025;15} = 27{,}49$
$t_{0,05;2} = 2{,}92$	$t_{0,025;4} = 2{,}78$	$t_{0,05;8} = 1{,}86$	$t_{0,025;8} = 2{,}31$	$t_{0,025;9} = 2{,}26$	$t_{0,05;9} = 1{,}83$
$t_{0,025;10} = 2{,}23$	$t_{0,05;10} = 1{,}81$	$t_{0,05;11} = 1{,}80$	$t_{0,025;11} = 2{,}20$	$t_{0,05;14} = 1{,}76$	$t_{0,025;14} = 2{,}14$
$t_{0,025;15} = 2{,}13$	$t_{0,05;15} = 1{,}75$	$t_{0,025;17} = 2{,}11$	$t_{0,025;22} = 2{,}07$	$t_{0,05;22} = 1{,}72$	$t_{0,025;23} = 2{,}07$
$t_{0,05;23} = 1{,}71$	$t_{0,05;24} = 1{,}71$	$t_{0,05;25} = 1{,}71$	$t_{0,05;26} = 1{,}71$	$t_{0,05;28} = 1{,}70$	$t_{0,05;30} = 1{,}70$
$F_{0,25;1,1} = 5{,}83$	$F_{0,025;1,10} = 6{,}94$	$F_{0,025;1,18} = 5{,}98$	$F_{0,05;1,10} = 4{,}67$	$F_{0,025;4,6} = 6{,}23$	$F_{0,025;6,4} = 9{,}20$
$F_{0,01;2,4} = 18{,}00$	$F_{0,01;6,3} = 27{,}91$	$F_{0,01;3,6} = 9{,}78$	$F_{0,05;3,15} = 3{,}29$	$F_{0,05;4,10} = 3{,}47$	$F_{0,05;4,12} = 3{,}26$
$F_{0,025;7,9} = 4{,}20$	$F_{0,025;9,7} = 4{,}82$	$F_{0,025;9,11} = 3{,}59$	$F_{0,025;11,9} = 3{,}91$	$F_{0,025;15,8} = 4{,}10$	$F_{0,025;8,15} = 3{,}20$

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

Question n° 1 : (6 points)

Soit $X_1, X_2, \ldots, X_{100}$ un échantillon aléatoire d'une population X de loi normale de moyenne $\mu = 1$ et de variance $\sigma^2 = 1$, c'est-à-dire $X \sim N(1,\ 1)$.

a) Calculer $P\left(1 \leq \overline{X} \leq 1{,}05\right)$.

0.5 $\bar{X} \sim N\left(1, \frac{1}{100}\right)$

0.25 $P\left(1 \leq \bar{X} \leq 1{,}05\right) = \Phi\left(\frac{1{,}05-1}{0{,}10}\right) - \Phi\left(\frac{1-1}{0{,}10}\right)$

1

0.25
$= \Phi(0{,}5) - \Phi(0)$
$= 0{,}69146 - 0{,}50$
$= 0{,}19146$

b) On définit deux variables T et U par

$$T = \sum_{i=1}^{6} X_i \quad \text{et} \quad U = \sum_{i=7}^{9} X_i.$$

1,5

Calculer $P(T > U)$.

0.75 $T - U \sim N(3, 9)$

$\underbrace{P(T>U) = P(T-U>0) =}_{0.5} \underbrace{1 - \Phi\left(\frac{0-3}{3}\right) = 1 - \Phi(-1) = \Phi(1) = 0{,}84134}_{0.25}$

Question n° 1 : (suite)

c) On définit $Y = \sum_{i=1}^{100} (X_i - 1)^2$.

1.c) Donner le nom de la loi exacte de Y ainsi que les valeurs de ses paramètres.

· loi exacte $Y \sim \chi^2_{100}$ ou Gamma $(r = 50, \alpha = \frac{1}{2})$

1

· ~~loi approximative $Y \underset{approx.}{\sim} N(100, 200)$~~

2.c) Donner le nom de la loi approximative de Y selon le théorème central limite ainsi que les valeurs de ses paramètres.

loi approximative $Y \underset{app.}{\sim} N(100, 200)$

1

d) On définit deux variables G et H par

$$G = \sum_{i=1}^{6}(X_i - 1)^2 \quad \text{et} \quad H = \sum_{i=7}^{9}(X_i - 1)^2.$$

Déterminer la valeur de la constante m telle que $P(G > mH) = 0{,}01$.

1,5

0,75 $\frac{G/6}{H/3} = \frac{G}{2H} \sim F_{6,3}$

0.50 $P(G > mH) = 0{,}01 \iff P\left(\frac{G}{2H} > \frac{m}{2}\right) = 0{,}01$

0.25 Donc $\frac{m}{2} = F_{0{,}01;6,3}$ i.e $m = 2 \times F_{0{,}01;6,3}$

$= 2 \times 27{,}91$

$= 55{,}82$

Question n° 3 : (8 points)

Le service du contrôle de la qualité veut étudier (estimer et tester) la proportion actuelle d'unités non conformes d'une production. Pour cette étude, le service compte utiliser un échantillon de n unités prélevées au hasard de la production.

a) **(2 points)** Quelle doit être la plus petite valeur de n pour que, avec un niveau de confiance de 95%, l'erreur d'estimation soit inférieure à 0,045 ?

$e = 0{,}045$

1.5 $n \geq \left(\frac{z_{\alpha/2}}{e}\right)^2 \times \frac{1}{4}$ formule (9.20) p. 262

0.5 i.e. $n \geq \left(\frac{1{,}96}{0{,}045}\right)^2 \times \frac{1}{4} = 474{,}27$ soit $n_{min} = 475$

Dans le passé, la proportion d'unités non conformes de la production était de 10%. Le service du contrôle de la qualité veut effectuer un test statistique, de seuil $\alpha = 5\%$, afin de vérifier si cette proportion est à présent plus élevée.

b) **(1 point)** Définir le paramètre d'intérêt et formuler les hypothèses H_0 et H_1 du test à effectuer.

p : proportion actuelle d'unités non conformes

On teste $H_0: p = 0{,}10$ contre $H_1: p > 0{,}10$ $(\alpha = 0{,}05)$

Question n° 3 : (suite)

c) (3 points) Avec un échantillon de 400 unités prélevées de la production, on obtient une valeur-p (ou *«p-value»*) de 0,0024 pour le test des hypothèses définies en b).

- Quelle doit être la conclusion du test ?

1 → Puisque p-value = 0,0024 est petite (< 0,05) on rejette H_0

conclusion = 0,5 On peut conclure que la proportion actuelle est plus élevée

- Quel est le nombre d'unités non conformes observées dans cet échantillon ?

1,5

Puisque p-value $= P(Z > z_0) = 0{,}0024 \Rightarrow \Phi(z_0) = 0{,}9976$

$\Rightarrow z_0 = 2{,}82$

Donc,

$$z_0 = \frac{\hat{p} - 0{,}10}{\sqrt{\frac{0{,}10 \times 0{,}90}{400}}} = 2{,}82 \Rightarrow \hat{p} = 0{,}10 + 2{,}82 \times \sqrt{\frac{0{,}10 \times 0{,}90}{400}} = 0{,}1423$$

Le nombre d'unités non conformes est $400 \times \hat{p} = 0{,}1423 \times 400$
$= 56{,}92$

soit 57

d) (2 points) Supposons que la proportion réelle d'unités non conformes soit actuellement de 11% et qu'on utilise un échantillon aléatoire de 500 unités pour le test. Quelle serait alors la probabilité de conclure à tort que la proportion d'unités non conformes actuelle n'est pas plus élevée que dans le passé ?

,5 $\beta(p=0{,}11) = \Phi\left(\dfrac{0{,}10 - 0{,}11 + 1{,}645\sqrt{\frac{0{,}10 \times 0{,}90}{500}}}{\sqrt{\frac{0{,}11 \times 0{,}89}{500}}}\right)$

0,25 $= \Phi(0{,}86)$

0,25 $= 0{,}805$