Chrome Fichier Modifier Présentation Outils Aide
Présentation
Créer
Récemment
Partagé avec moi
Favoris
Corbeille
Stockage
Type
Propriétaire
Modifié
Taille de fichier
Origine
Propriétaires
Activité
Corbeille
Stockage
Nouveau
Présentation
Feuille de calcul
Document
Dossier importé
Importer un fichier
Importer un dossier
Partager
Obtenir le lien
Présentation sans titre
Fichier Modifier Insérer Présentation Outils Aide
Diapo 1 / 1
Mise en page
Thème
Transition
Animer
Partager
Présentation

POLYTECHNIQUE MONTRÉAL

Questionnaire
Examen final

MTH2302D

Sigle du cours

Identification de l'étudiant(e)		
Nom : [illegible]	**Prénom :** Clarence	
Signature : [illegible]	**Matricule :** 2057572	**Groupe :** D

Sigle et titre du cours	Groupe	Trimestre
MTH2302D **Probabilités et statistique**	**TOUS**	**Automne 2017**
Professeur	**Local**	**Téléphone**
Luc Adjengue	**A-520.23**	**4475**
Jour / **Date**	**Durée**	**Heure**
Mercredi / **20 décembre 2017**	**2h30**	**9h30 à 12h00**

Documentation	Calculatrice
☐ Aucune	☐ Aucune
☐ Totale	☐ Totale
☒ Voir directives particulières	☒ Non programmable

Les calculatrices, ajoutes mémoires ou téléavertisseurs sont interdits.

Réservé	
1.	/8
2.	/25
3.	/17
4.	/17
5.	/17
6.	/16
TOTAL	/90

Directives particulières

1. Ne détachez aucune page de l'examen.
2. Les 3 dernières pages (annexes) ne doivent pas être détachées.
3. Documentation permise : une feuille 8,5x11 recto-verso.
4. Chaque réponse doit être complète et accompagnée d'une justification.

Pour ceux d'après-midi seulement : Les étudiants, le professeur ou répondants à aucune question pendant les 30 premières minutes de l'examen. Aucune question ne sera répondue pour diverses raisons (données manquantes, données erronées, etc.) veuillez vous référer à votre chargé de cours pour toute question.

Important

Cet examen contient	8 questions sur un total de 18 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 $X_1, X_2, \ldots, X_{10}$ 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 $P(1 \leq \overline{X} \leq 1.05)$.

$P\left(\frac{1-\mu}{\sigma/\sqrt{n}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{1.05-\mu}{\sigma/\sqrt{n}}\right)$; $P(0 \leq Z \leq 0.95)$; $= P(Z \leq 0)$; $= \Phi(0) - \Phi(-0.95)$; $= 0.5 + 0.3289$

$P(1 \leq \overline{X} \leq 1.05) = 0.8289$

b) On définit deux variables T et U par

$$T = \sum_{i=1}^{10} X_i \quad \text{et} \quad U = \sum_{i=1}^{10} X_i$$

Calculer $P(T > U)$.

$P(T > U) = P(T - U > 0)$; $\mu_{T-U} = \frac{\sigma^2}{n}$; $\sigma_{T-U} = \frac{\sigma^2}{n}$

$P(T > U) = 0.1586$

Question n° 1 : (suite)

c) On définit $Y = \sum_{i=1}^{10} X_i - 17$.

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

Sachant que $Y \sim N(E(Y), \sigma_Y)$ car Y est une somme d'aléatoire de variables aléatoires.

$E(Y) = 0 \Rightarrow \mu_Y$

$V(Y) = 2 \quad \sigma = 100$

$X \sim N(E(X), V(X))$

$X \sim N(0, \frac{\sigma^2}{n})$

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

$\Rightarrow \mu_Y = 0$

$\sigma_Y = \sqrt{10}$

$X \sim N(0, \frac{\sigma^2}{n})$

d) On définit deux variables G et H par

$$G = \sum_{i=1}^{10} X_i - 17 \quad \text{et} \quad H = \sum_{i=1}^{10} X_i - 17.$$

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 fréquences du nombre de défauts X sur le lot. Il n'est plus observé dans un échantillon aléatoire de 100 pièces d'une production.

Nombre de défauts (x_i)	0	1	2	3
Nombre de pièces (N_i)	15	40	30	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.

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

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

$H_0 : X \sim Po(\lambda)$

$H_1 : X \not\sim Po(\lambda)$

b) (3 points) Effectuer le test et conclure au seuil $\alpha = 1\%$.

On rejette H_0 si $T_0 \geq k_{\alpha}$

comme on ignore λ, on $\hat{\lambda}$

$\hat{\lambda} = \frac{\sum x_i n_i}{n} = 1.5$

$P(X=0) = 0.223$

$P(X=1) = 0.335$

$P(X=2) = 0.251$

$P(X=3) = 0.188$

H_0 : 0 1 2 3

H_1 : 15 40 30 15

E : 22.3 33.5 25.1 18.8

$T_0 = \sum \frac{(n_i - E_i)^2}{E_i} = 0.74$

$\alpha = 1\% \Rightarrow 1 - \alpha = 0.99$

$\nu = 4 - 1 - 1 = 2$

$k_{\alpha} = 9.210$

On rejette H_0 car $T_0 > k_{\alpha}$

$X \not\sim Po(\lambda)$

Question n° 2 : (suite)

Question n° 3 : (8 points)

Le service du contrôle de la qualité veut étudier (estimer et tester) la proportion acceptable 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 0,04?

$\alpha = 5\%$ comme il s'agit de 95% soit 0,05

$\frac{\sigma}{\sqrt{n}} \leq 0.04$ tout mis à 0,04 à la 2

$n \geq \left(\frac{z_{\alpha/2}\,\sigma}{0.04}\right)^2 = \left(\frac{1.96 \times 0.5}{0.04}\right)^2 = 1225 = n^*$

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

$H_0 : p = 0.05$

$H_1 : p > 0.05$

proportion de défectueux : p

Question n° 3 : (suite)

c) (3 points) Avec un échantillon de 400 unités prélevées de la production, on obtient une valeur p (ou $\hat{p}$) égale à 0,0625 pour le test. Les deux hypothèses définies en b):

- Quelle doit être la conclusion du test ?

On rejette H_0 car pseudo α

$k_\alpha = \hat{p}$ valeur pseudo α ; $\alpha = 0.05$

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

Il y a plus de 50 unités non conformes car la proportion d'unités non conformes est plus grande que 0,05

d) (2 points) Supposons que la proportion réelle d'unités non conformes soit actuellement de 12% 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 ?

$P(\text{non-rejet } H_0 | H_0 \text{ est faux})$; $= \beta$

$\beta = \Phi\left(\frac{0.05 - 0.12 - \frac{1}{2(500)}}{\sqrt{\frac{0.12(0.88)}{500}}}\right)$; $= \Phi\left(\frac{0.12 - 0.05 - \frac{1}{1000}}{0.0145}\right)$; $= \Phi(-2.9655)$; $= 1 - \Phi(2.9655)$; $= 0.0052$; $= 0.0148$

Question n° 4 : (12 points)

Dans un réseau informatique, un ingénieur mesure les délais de communication (en millisecondes) entre deux ordinateurs avec un chronomètre électronique. Il obtient un échantillon aléatoire de taille $n = 10$ de la variable X. L'ingénieur mesure ensuite un échantillon aléatoire de taille $n = 10$ de la variable Y à partir d'un ordinateur de taille $n = 8$ 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.

Échantillon	n	Taille n	Moyenne $\bar{X}$	Variance S^2
$X_1 \sim N(\mu_1, \sigma_1^2)$	10	10	52,1	0,37
$X_2 \sim N(\mu_2, \sigma_2^2)$	8	10	50,2	3,18

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_1.

$X \sim N(\mu_1, \sigma_1^2)$; $\sigma_1^2 = 0.37$; $\bar{X} = 52.1$

$IC(\mu_1) = \bar{X} \pm z_{\alpha/2}\frac{\sigma_1}{\sqrt{n}}$

$52.1 \pm 1.645 \times \frac{0.37}{\sqrt{10}}$

52.1 ± 0.19

Question n° 4 : (suite)

b) (3 points) Donner un intervalle de prévision de niveau 90% pour une dixième nouvelle mesure de délai du protocole expérimental.

$X_{11} \in \bar{X} \pm t_{\alpha/2}(8)\sqrt{S^2(\frac{1}{10})}$

$X_{11} \in 63,5 \pm 1,86 \times 2,42$

$X_{11} \in [58,995 \; , \; 68,005]$

c) (3 points) Peut-on rejeter l'affirmation selon laquelle $\sigma_1^2 = 0,5$? Donner la conclusion d'un test d'hypothèse au seuil $\alpha = 0,05$.

$H_0 : \sigma_1^2 = 0,5$

$H_1 : \sigma_1^2 \neq 0,5$

On rejette H_0 si: $F_0 \geq F_{\alpha/2,n-1,n-1}$ ou $F_0 \leq F_{1-\alpha/2,n-1,n-1}$

$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(15)\times 0,37}{0,5} = 11,1$ on ne $\chi^2_{\alpha/2,n-1}$

$\chi^2_{0,975}(14) = 26,119$ $\quad \chi^2_{0,025}(14) = 5,629$

$\chi^2_{0,975}(14) = 6,571$ $\quad \chi^2_{0,025}(14) = 23,685$

Donc $\chi_0^2 \in [5,629 \; , \; 23,685]$ on ne peut pas rejeter H_0

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 $\alpha = 0,05$.

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$F_0 = \frac{S_1^2}{S_2^2}$ on rejette H_0 si: $F_0 \geq F_{\alpha/2,n_1-1,n_2-1}$ ou $F_0 \leq F_{1-\alpha/2,n_1-1,n_2-1}$

$F_0 = \frac{0,37}{0,05} = 0,6471$

$F_{0,975,14,14} = F_{0,025,14,14} = \frac{1}{F_{0,025,14,14}} = \frac{1}{3,1} = 0,3225$ $\leftarrow$ On rejette H_0

$F_{0,025,14,14} = 3,1$

Oui on peut dire que σ_1^2 et σ_2^2 sont différentes car on a rejeté l'hypothèse, puisque elles sont égales.

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 $\alpha = 0,05$.

$Y = \bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \sigma^2(\frac{1}{n_1} + \frac{1}{n_2}))$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 < \mu_2$ donc $\mu_1 - \mu_2 < 0$

On rejette H_0 si:

$T_0 \leq t_{\alpha,\nu}$

$T_0 = \frac{\bar{Y} - 0 - \Delta_0}{\sqrt{S^2(\frac{1}{n_1} + \frac{1}{n_2})}} = 5,06873$

$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} = 27$

$t_{0,05,27} = t_{1-\alpha,\nu} = 1,681$

On rejette H_0 car $5,0687 > 1,681$

Donc oui on peut affirmer que moyenne les délais du protocole expérimental sont plus courts que ceux du protocole standard.

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 pièce et la température à midi de la même journée. Pour ce faire, on envisage un modèle de régression linéaire simple d'équation

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

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	15,7	18,8	18,5	17,1	17,3	21,1	17,1	19,4	15,6	19,9	16,2	20,7
y_i	71,5	111,4	78,6	112,6	80,6	214,1	108,9	164,4	64,5	191,5	68,8	222,3

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

$\sum_{i=1}^{12} x_i = 216,13$; $\sum_{i=1}^{12} y_i^2 = 5687,27$; $\sum_{i=1}^{12} y_i y_i = 25421,34$;

$\sum_{i=1}^{12} y_i = 1152,20$; $\sum_{i=1}^{12} y_i^2 = 121\,916,60$.

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

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = 2,0354$ $\quad \hat{\beta}_0 = \frac{\sum y_i}{n} - \hat{\beta}_1 \frac{\sum x_i}{n} = -32,513$

$= \frac{1525,31}{601,75} = 2,535$

$\hat{\beta}_0 = 96,08 - (2,535)(18,01) = -45,65$

$Y_i = 2,034 \, x_i - 32,513$

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	43,74	1	43,74	0,7283
Erreur	180,93	10	17,983	
Totale	1 224,20	11		

$SS_R = \hat{\beta}_1 \cdot S_{xy} = \hat{\beta}_1 \cdot \sum x_i y_i - n\bar{x}\bar{y} = 2,03 \cdot 25421,34 - 12 \cdot 18,01 \cdot 96,08 = 43,74$

$SS_E = SS_T - SS_R = 1224,20 - 43,74 = 1180,46$

$F_0 = \frac{MS_R}{MS_E} = \frac{43,74}{17,983} = 2,43$

$H_0 : \beta_1 = 0$ $\quad F_{0,05,1,10} = F_{0,95,1,10} = 4,96$

$H_1 : \beta_1 \neq 0$ $\quad F_0 \geq F_{\alpha,\nu} \Rightarrow$ on ne rejette pas H_0

On peut en conclure que la régression n'est pas significative.

c) (3 points) Donner une estimation ponctuelle de la variance σ^2.

$S^2 = \frac{SS_E}{n-2} = \frac{1180,46}{10} = 118,05$

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.

$IP(x_0) \in \left(\hat{Y} \pm t_{\alpha/2,n-2}\sqrt{MS_E(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}})}\right)$

$\hat{Y}_0 = 2,034 \times 20 - 32,513 = 8,54$

$t_{0,025,10} = 2,23$

$\sqrt{MS_E(1 + \frac{1}{12} + \frac{(20 - 18,01)^2}{1224,20})} = 11,72$

$IP(x_0) \in [-3,18 \; , \; 19,27]$

e) (3 points) Calculer le coefficient de détermination R^2 de interpréter ce résultat.

$R^2 = \frac{SS_R}{SS_T} = \frac{43,74}{1224,20} = 0,0357$

R^2 varie entre 0 et 1 il désigne le pourcentage de la variance expliquée par la droite de régression. Ici avec $R^2 = 0,035$ notre régression est relativement faible.

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 $Exp(\beta)$.

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

$\lambda = 2 \quad \mu = 0,5$

$\Sigma = \bar{X} \pm t$

Soient $X_1, X_2, X_3 \sim Exp(\beta)$ indépendantes

$R_1(t) = \int_t^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx$

$= \int_t^\infty \frac{1}{\Gamma(1)\beta^1} x^{1-1} e^{-x/\beta} dx$

$= \int_t^\infty \frac{1}{\beta} e^{-x/\beta} dx$

$= -e^{-x/\beta}\Big|_t^\infty$

$= e^{-t/\beta}$

Cumulative $R_1(t) = 1 - \sum_{k=0}^{2} \frac{1}{k!} R_1(t)^k (1 - R_1(t))^{2-k}$

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

b

Question n° 6 : (suite)

b) (2 points) Un commerce possède deux caisses modélisées par une file d'attente M/M/1. Les clients arrivent au taux $\lambda_1 = 10$ clients/heure et les clients y prennent au taux $\mu_1 = 40$ clients/heure et le seul service au taux $\mu_2 = 10$ 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 $\mu_2 = 8$ 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.

$\lambda_1 = 5 \quad \mu_1 = 20$

$\lambda_2 = 2 \quad \mu_2 = 5$

$E[N] = \frac{\lambda_1}{\mu_1 - \lambda_1} + \frac{\lambda_2}{\mu_2 - \lambda_2} = \frac{5}{20-5} + \frac{2}{5-2} = \frac{1}{3} + \frac{2}{3} = 1$

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

$P(X = 0) = (1 - \frac{1}{3}) \cdot \frac{1}{3} = 0,1$

$P_1(x = 0) = (1 - \frac{1}{3}) \cdot \frac{1}{3} = 0,2$

$P(0) = P_1(0) = P(0) = 0,02$