Documents autorisés : cours, TD, notes manuscrites, calculatrice. Barème indicatif : 3+2+4+3+4+4 Durée : 1h 30.

Les résultats sont présentés avec trois chiffres significatifs, sauf indication particulière.

Exercice 1

Changement de variable et intégrale généralisée

1. y est un réel positif quelconque. Calculer $\int_0^y \frac{1}{4x^2+4x+2}\,dx$ en utilisant le changement de variable $t = 2x+1$.
2. En déduire la nature de $\int_0^{+\infty} \frac{1}{4x^2+4x+2}\,dx$ (convergente ou divergente) et son éventuelle valeur.

Exercice 2

Espérance et variance

La loi de probabilité d'une variable aléatoire X est définie par

1. $X(\Omega) = \{0, 1, 2, 3, 4, 5\}$
2.

i	0	1	2	3	4	5
$P(X=i)$	0,10	0,05	0,30	0,30	0,20	0,05

Préciser l'espérance $E(X)$ et la variance $V(X)$.
Reproduire et compléter le tableau suivant.

$E(X)$	$V(X)$

Exercice 3

Intervalle de fluctuation et loi binomiale

Un directeur commercial affirme que 90 % des consommateurs sont satisfaits de la qualité des produits commercialisés par son entreprise. On réalise une étude de satisfaction sur un échantillon de 80 personnes. Parmi les personnes interrogées, 68 déclarent être satisfaites des produits.
Est-ce qu'on doit remettre en question l'affirmation du directeur commercial ?

Pour répondre à la question, on va déterminer l'intervalle de fluctuation au seuil de risque α (ou de niveau $1-\alpha$) de la fréquence de personnes satisfaites par une détermination directe. On prendra $\alpha = 0,05$.
On note X_i $(1 \leq i \leq 80)$ la variable aléatoire prenant la valeur 1 pour une personne satisfaite et la valeur 0 pour une personne non satisfaite. On suppose $X_i \sim \mathcal{B}(1;\, 0,90)$ et que les X_i sont indépendantes $(1 \leq i \leq 80)$.

On pose $X = \sum_{i=1}^{80} X_i$ (nombre de personnes satisfaites).
X suit la loi binomiale $\mathcal{B}(80;\, 0,90)$: $X \sim \mathcal{B}(80;\, 0,90)$.

1. Déterminer l'intervalle de fluctuation $[\![n_1, n_2]\!]$ au seuil de risque $\alpha = 0,05$ (ou de niveau $1-\alpha$) qui est le plus petit intervalle $[\![n_1, n_2]\!]$ tel que $P(X < n_1) \leq \frac{\alpha}{2}$ et $P(X > n_2) \leq \frac{\alpha}{2}$.
   Indication :
   n_1 est l'entier vérifiant $P(X \leq n_1 - 1) \leq \frac{\alpha}{2}$ et $P(X \leq n_1) > \frac{\alpha}{2}$,
   n_2 est l'entier vérifiant $P(X \leq n_2 - 1) < 1 - \frac{\alpha}{2}$ et $P(X \leq n_2) \geq 1 - \frac{\alpha}{2}$.
   On pourra utiliser la table donnée en annexe.
2. Est-ce que $68 \in [\![n_1, n_2]\!]$?
3. Est-ce qu'on doit remettre en question l'affirmation du directeur commercial ?
4. Préciser
   (a) $P(X < n_1)$
   (b) $P(n_1 \leq X \leq n_2)$
   (c) $P(X > n_2)$

   Reproduire et compléter le tableau suivant.

$P(X < n_1)$	
$P(n_1 \leq X \leq n_2)$	
$P(X > n_2)$	

Exercice 4

Intervalle de confiance et loi de Student
On a construit un algorithme, noté *A1*, dont on souhaite déterminer les performances en terme de temps de calcul. Soit X_i le temps de réalisation de l'algorithme *A1* sur la i-ème simulation.
On suppose les X_i indépendantes et de loi normale $\mathcal{N}(\mu, \sigma)$, μ et σ étant inconnus.
On réalise $n = 41$ simulations.
Les estimateurs des paramètres la moyenne μ et la variance σ^2 sont respectivement
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ avec $n = 41$.
On obtient une moyenne empirique de $\overline{x} = 55$ minutes et une variance empirique de $s^2 = 100$ minutes^2.

Donner un intervalle de confiance I pour μ de niveau de confiance 95 %. Expliquer.

Indications :

1. $\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathcal{T}_{n-1}$, loi de Student à $n - 1 = 40$ degrés de liberté.
2. L'intervalle de confiance pour μ de niveau de confiance $1 - \alpha = 0,95$ est défini par $P\left(\left|\frac{\overline{X} - \mu}{S/\sqrt{n}}\right| \leq t_{n-1;1-\alpha/2}\right) = 1 - \alpha$.

Exercice 5

Test et loi normale
On souhaite étudier les effets d'une campagne publicitaire sur la vente d'un journal hebdomadaire. En dehors de toute campagne publicitaire, on admet que le produit X de la vente hebdomadaire de ce journal suit, en milliers d'euros, une loi normale $\mathcal{N}(\mu, \sigma)$, où $\mu = 22$ et $\sigma^2 = 12$. Après la mise en place de la campagne publicitaire, on observe le produit de la vente du journal sur $n = 8$ semaines. On obtient alors les valeurs x_i $(1 \leq i \leq 8)$ suivantes (en milliers d'euros) :

$$18,7\,;\ 25,3\,;\ 23,1\,;\ 27,7\,;\ 26,2\,;\ 19,2\,;\ 21,9\,;\ 19,9.$$

On veut savoir si la campagne de publicité a été efficace ou non. Pour ce faire, on veut tester l'hypothèse nulle $H_0 : \mu = 22$ contre l'hypothèse alternative $H_1 : \mu > 22$ au niveau $\alpha = 5$ % (test unilatéral à droite).
On admet que

- la variance σ^2 du produit de la vente ne varie pas et on suppose donc $\sigma^2 = 12$.
- $X_i \sim \mathcal{N}\left(22, \sqrt{12}\right)$ et les X_i indépendantes, en notant X_i le produit de la semaine i $(1 \leq i \leq 8)$.

1. Lorsqu'on teste $H_0 : \mu = 22$ contre $H_1 : \mu > 22$, se place-t-on du point de vue de l'éditeur du journal ou du concepteur de la campagne publicitaire ?
   Justifier la réponse.
2. Calculer $\overline{x} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} x_i$.
3. Préciser le nombre u vérifiant $P(T \leq u) = 0,95$ avec $T \sim \mathcal{N}(0,1)$.
   Indication : utiliser une table de probabilité.
4. En déduire $P(T \geq u)$, puis la région de rejet $\mathcal{R}$.
   Indication : en posant $\mu_0 = 22$ et $\sigma = \sqrt{12}$, sous l'hypothèse nulle $H_0 : \mu = \mu_0$,
   $T = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,\, 1)$ avec $\overline{X} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$.
   La région de rejet $\mathcal{R}$ est définie par $P\left(\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \geq u\right) = 0,05$.
5. Est-ce que $\overline{x}$ appartient à la région de rejet ?
6. Quelle est la conclusion du test au niveau $\alpha = 5$ % : la campagne publicitaire a-t-elle été efficace ?
7. Préciser la p-value $P_c(\overline{x}) = P_{H_0}\left(\overline{X} \geq \overline{x}\right)$.
   Indication : utiliser une table de probabilité.
8. Si le test est significatif, quel est le degré de signification du test (test *significatif*, *très significatif*, *hautement significatif*).

Exercice 6

Test du χ^2

Dans une agence de location de voitures, le patron veut savoir quelles sont les voitures qui n'ont roulé qu'en ville pour pouvoir les revendre immédiatement. Pour cela, il y a dans chaque voiture une boîte noire qui enregistre le nombre d'heures pendant lesquelles la voiture est restée au point mort, au premier rapport, au deuxième rapport, ..., au cinquième rapport. On sait qu'une voiture qui ne roule qu'en ville passe en moyenne 10 % de son temps au point mort, 5 % en première, 30 % en seconde, 30 % en troisième, 20 % en quatrième, et 5 % en cinquième.

- Sur une première voiture, on constate sur $n = 2\,000$ heures de conduite : 210 h au point mort (rapport $i = 0$), 94 h en première (rapport $i = 1$), 564 h en seconde (rapport $i = 2$), 630 h en troisième (rapport $i = 3$), 390 h en quatrième (rapport $i = 4$), et 112 h en cinquième (rapport $i = 5$). On se demande si cette voiture n'a fait que rouler en ville.
- Avec une autre voiture, on obtient les données suivantes : 176 h au point mort, 70 h en première, 540 h en seconde, 594 h en troisième, 450 h en quatrième et 170 h en cinquième. On se pose la même question.

Pour répondre à la question pour chacune des voitures, on teste l'adéquation de l'échantillon à la loi discrète définie par : $\pi_0 = 0,1$; $\pi_1 = 0,05$; $\pi_2 = 0,30$; $\pi_3 = 0,30$; $\pi_4 = 0,20$; $\pi_5 = 0,05$ en utilisant un test du χ^2. On teste H_0 = *la voiture n'a roulé qu'en ville*, contre H_1 = *la voiture n'a pas roulé qu'en ville*, au niveau 5 % (test unilatéral à droite).

1. **Première voiture**

   (a) Pour la première voiture, on obtient le tableau suivant.

Rapport i	0	1	2	3	4	5	Total n
Effectif observé n_i	210	94	564	630	390	112	2 000
Effectif théorique $n\pi_i$	200	100	600	600	400	100	2 000

   Vérifier :

   i. $n \geq 30$,

   ii. pour tout i $(0 \leq i \leq 5)$, $n\pi_i \geq 5$.

   (b) La statistique de test (dite du χ^2) utilisée est $D = \sum_{i=0}^{5} \frac{(N_i - n\pi_i)^2}{n\pi_i}$.
   Sous H_0, la variable D suit une loi du χ^2 à $\nu = 6 - 1 = 5$ de degrés de libertés.
   Préciser $\delta = x_{5;0,95}$ tel que $P(D \leq \delta) = 0,95$.
   Indication : utiliser une table de probabilité.

   (c) En déduire la région de rejet $\mathcal{R}$ au niveau 5 %.

   (d) Calculer $d = \sum_{i=0}^{5} \frac{(n_i - n\pi_i)^2}{n\pi_i} = \frac{(210-200)^2}{200} + \ldots.$

   (e) Peut-on rejeter l'hypothèse H_0 au niveau 5 % ?

   (f) Préciser un encadrement de la p-value $P_c(d) = P_{H_0}(D \geq d)$.
   Indication : utiliser une table de probabilité.

   (g) Si le test est significatif, quel est le degré de signification du test (test *significatif*, *très significatif*, *hautement significatif*).

Deuxième voiture

2. Pour la deuxième voiture, on obtient le tableau suivant.

Rapport i	0	1	2	3	4	5	Total n
Effectif observé n_i	176	70	540	594	450	170	2 000
Effectif théorique $n\pi_i$	200	100	600	600	400	100	2 000

   (a) Calculer $d = \sum_{i=0}^{5} \frac{(n_i - n\pi_i)^2}{n\pi_i} = \frac{(176-200)^2}{200} + ...$

   (b) Peut-on rejeter l'hypothèse H_0 au niveau 5 % ?

   (c) Préciser un encadrement de la p-value $P_c(d) = P_{H_0}(D \geq d)$.
   Indication : utiliser une table de probabilité.

   (d) Si le test est significatif, quel est le degré de signification du test (test *significatif*, *très significatif*, *hautement significatif*).

Annexe : Loi binomiale $\mathcal{B}(n,\, p)$

$$F(i) = P(X \leq i) = \sum_{k=0}^{i} C_n^k p^k (1-p)^{n-k} = \sum_{k=0}^{i} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} \ \mathbf{(n = 80\,;\ p = 0,90)}$$

i	$P(X \leq i)$
59	0.0000
60	0.0001
61	0.0003
62	0.0008
63	0.0021
64	0.0053
65	0.0123
66	0.0267
67	0.0538
68	0.1004
69	0.1734
70	0.2766
71	0.4073
72	0.5544
73	0.6995
74	0.8231
75	0.912
76	0.9647
77	0.9893
78	0.9978
79	0.9998
80	1.0000