Documents autorisés : cours, TD, notes manuscrites, calculatrice. Barème indicatif sur 30 : 5+5+10+5+5 Durée : 1h 30.

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

Exercice 1

Loi conjointe

Dans cet exercice, les résultats sont donnés sous la forme de fractions.

On lance deux fois un dé tétraédrique parfait numéroté de 1 à 4. X est le premier numéro lu sur la base, Y le deuxième et M le plus grand des deux.

1. Construire la loi conjointe de X et M, en précisant les lois marginales de X et de M :

$X\backslash M$	1	2	3	4	$P(X=i)$
1					
2					
3					
4					
$P(M=j)$					

2. En déduire la loi de probabilité de M :

j	1	2	3	4
$P(M=j)$				

   Représenter graphiquement.

3. X et M sont-elles indépendantes ? Expliquer.
4. Calculer $E(M)$ et $V(M)$. Reproduire et compléter le tableau suivant.

$E(M)$	$V(M)$

Exercice 2

Calcul intégral et loi exponentielle

On modélise le temps d'attente en minutes entre deux clients à un guichet comme une variable aléatoire X suivant une loi exponentielle de paramètre $a = \frac{1}{5}$: $X \sim \mathcal{E}\left(\frac{1}{5}\right)$.

1. Préciser le temps moyen d'attente $E(X)$.
   Indication : si $X \sim \mathcal{E}(a)$ alors $E(X) = \frac{1}{a}$.
2. Préciser la fonction de répartition $F(x) = P\left(X \leq x\right) = \int_0^x ae^{-at}$ pour $x \geq 0$.
3. Quelle est la probabilité d'attendre plus de 5 minutes ? plus de 10 minutes ?
4. Sachant qu'on a déjà attendu au moins 10 minutes, quelle est la probabilité d'attendre encore au moins 5 minutes : $P\left((X \geq 15)/(X \geq 10)\right) = P_{X\geq 10}\left(X \geq 15\right)$?

Exercice 3

Intervalle de fluctuation, loi binomiale et loi normale
Un village isolé situé près d'une usine chimique a vu naître 70 enfants, parmi lesquels 43 garçons. Intrigué par ce pourcentage, le maire se demande si, pour son village, la probabilité p qu'un nouveau-né soit un garçon vaut 0,52 (proportion en France). 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 garçons par une détermination directe, puis par une approximation en utilisant une loi normale. On prendra $\alpha = 0,05$.
On note X_i $(1 \leq i \leq 70)$ la variable aléatoire prenant la valeur 1 pour un garçon et la valeur 0 pour une fille. On suppose $X_i \sim \mathcal{B}(1;\ 0,52)$ et que les X_i sont indépendantes $(1 \leq i \leq 70)$.
On pose $X = \sum_{i=1}^{70} X_i$ (nombre de garçons).

1. **Loi binomiale**
   X suit la loi binomiale $\mathcal{B}(70;\ 0,52)$: $X \sim \mathcal{B}(70;\ 0,52)$.
   (a) 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.
   (b) Est-ce que $43 \in [\![n_1, n_2]\!]$?
   (c) Calculer $P(n_1 \leq X \leq n_2)$.
2. **Loi normale**
   On envisage à présent d'approcher la loi binomiale $\mathcal{B}(70;\ 0,52)$ par une loi normale $\mathcal{N}(m, \sigma)$ pour déterminer l'intervalle de fluctuation.
   (a) Vérifier que la loi binomiale peut être approchée par une loi normale.
   Indication : d'après le théorème central limite, l'approximation de la loi binomiale $\mathcal{B}(n,p)$ par la loi normale $\mathcal{N}(m,\sigma)$ avec $m = np$ et $\sigma = \sqrt{np(1-p)}$ est envisageable pour n assez grand $(n \geq 30)$, $np \geq 5$ et $n(1-p) \geq 5$.
   (b) Préciser les paramètres m et σ de cette loi normale.
   (c) On suppose donc à présent $X \sim \mathcal{N}(m, \sigma)$.
   Pour construire l'intervalle de fluctuation au seuil de risque $\alpha = 0,05$ (ou de niveau $1 - \alpha = 0,95$), on détermine d'abord l'intervalle $[x_1,\ x_2]$ défini par
   $P\left(\left|\frac{X-m}{\sigma}\right| \leq u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$
      i. Déterminer le réel $u_{1-\frac{\alpha}{2}} = u_{0,975}$.
      Indication : $u_{0,975}$ est le réel vérifiant $P(T \leq u_{0,975}) = 0,975$ avec $T \sim \mathcal{N}(0,1)$.
      ii. En déduire l'intervalle $[x_1,\ x_2]$ défini par $P\left(\left|\frac{X-m}{\sigma}\right| \leq u_{0,975}\right) = 0,95$.
   (d) En déduire l'intervalle de fluctuation $[\![n'_1,\ n'_2]\!]$ au seuil de risque $\alpha = 0,05$.
   Indication : n'_1 et n'_2 sont les arrondis de x_1 et x_2 à l'entier le plus proche.
   (e) Comparer $[\![n'_1,\ n'_2]\!]$ et $[\![n_1,\ n_2]\!]$.

3. Que peut conclure le maire au seuil de risque $\alpha = 0,05$?

$\boxed{\textbf{Exercice 4}}$

Loi de Student et loi du χ^2

Voulant évaluer rapidement les résultats obtenus par ses 200 étudiants lors d'un partiel, un professeur décide de corriger n copies tirées au hasard. Il admet par ailleurs que les notes de ses étudiants suivent une loi normale de moyenne μ et d'écart-type σ inconnus.

1. Le professeur corrige un échantillon de $n = 7$ copies et il obtient les notes suivantes : 12 ; 10 ; 15 ; 9 ; 13 ; 8 ; 10.
2. Préciser une estimation ponctuelle $\overline{x}$ de μ, puis une estimation ponctuelle s^2 de σ^2.
   Indications : Les estimateurs de la moyenne μ et de la variance σ^2 sont respectivement $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.
3. **Intervalle de confiance de μ avec σ inconnu**
   La variance n'étant pas supposée connue, elle doit être estimée. On utilise l'estimation s^2 trouvée ci-dessus.
   On a alors $\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim \mathcal{T}_{n-1}$, loi de Student à $n - 1 = 7 - 1 = 6$ degrés de liberté.
   Donner un intervalle de confiance I pour μ de niveau de confiance 95 %.
   Indication : $P\left(\left|\frac{\overline{X} - \mu}{S/\sqrt{n}}\right| \leq t_{n-1;1-\alpha/2}\right) = 1 - \alpha$ $(\alpha = 0,05)$.
4. Combien de copies au minimum le professeur doit-il corriger s'il veut situer la moyenne générale de ses étudiants dans un intervalle de confiance d'amplitude inférieur à 2, avec un niveau de confiance de $1 - \alpha = 95$ % ?
5. **Intervalle de confiance de σ avec μ inconnue**
   La moyenne μ n'étant pas connue, $\frac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1}$, loi du χ^2 à $n - 1 = 7 - 1 = 6$ degrés de liberté.
   Donner un intervalle de confiance I' pour σ de niveau de confiance 95 %.
   Indication : $P\left(\chi_{n-1;\alpha/2} \leq (n-1)\frac{S^2}{\sigma^2} \leq \chi_{n-1;1-\alpha/2}\right) = 1 - \alpha$.

Exercice 5

Ajustement à une loi

Pour vérifier si un dé est pipé, on compare une distribution observée et une distribution théorique (ici uniforme de paramètre $\pi = \frac{1}{6}$).

En posant $n = 120$ et N_i/n les fréquences observées, ainsi que $\pi_i = \pi$ les probabilités attendues, on associe les effectifs observés n_i et les effectifs théoriques attendus $n\pi_i = n\pi = 20$ $(1 \leq i \leq 6)$:

Chiffre	1	2	3	4	5	6	Total
Effectif observé n_i	13	24	27	15	14	27	120
Effectif théorique $n\pi_i$	20	20	20	20	20	20	120

On pose :

1. H_0 : le chiffre obtenu suit la loi uniforme de paramètre $\frac{1}{6}$,
2. H_1 : le chiffre obtenu ne suit pas la loi uniforme de paramètre $\frac{1}{6}$.

On teste H_0 au niveau 5 %.

La statistique de test (dite du χ^2) utilisée est $D = \sum_{i=1}^{6} \frac{(N_i - n\pi)^2}{n\pi}$.

$n \geq 30$ et pour tout i $(1 \leq i \leq 6)$, $n\pi_i = 120 \times \frac{1}{6} = 20 \geq 5$.

On en déduit que, sous H_0, D suit une loi du χ^2 à $\nu = 6 - 1 = 5$ de degrés de libertés.

1. Calculer

$$d = \sum_{i=1}^{6} \frac{(n_i - n\pi)^2}{n\pi} = \frac{(13-20)^2}{20} + \frac{(24-20)^2}{20} + \frac{(27-20)^2}{20} + \frac{(15-20)^2}{20} + \frac{(14-20)^2}{20} + \frac{(27-20)^2}{20}.$$

2. Préciser δ tel que $P_{H_0}(D \leq \delta) = 0,95$.
3. En déduire la région de rejet $\mathcal{R} = [\delta; +\infty[$ au niveau 5 %.
4. Peut-on rejeter l'hypothèse H_0 d'uniformité au niveau de 5 % ?
5. Préciser un encadrement de la p-value $P_c(d) = P_{H_0}(D \geq d)$.
   Indication : utiliser une table de probabilité.
6. En déduire le degré de signification du test (test *significatif*, *très significatif*, *hautement significatif*).

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

$$F(i) = P(X \leq i) = \sum_{k=0}^{i} C_n^k p^k (1-p)^{n-k} = \sum_{k=0}^{i} \left(\begin{array}{c} n \\ k \end{array}\right) p^k (1-p)^{n-k} \ \mathbf{(n = 70\,;\ p = 0,52)}$$

i	$P(X \leq i)$
22	0,0004
23	0,0010
24	0,0021
25	0,0044
26	0,0088
27	0,0165
28	0,0293
29	0,0493
30	0,0790
31	0,1206
32	0,1754
33	0,2438
34	0,3244
35	0,4143
36	0,5089
37	0,6031
38	0,6917
39	0,7705
40	0,8366
41	0.8891
42	0.9283
43	0.9560
44	0.9743
45	0.9859
46	0.9926
47	0.9964
48	0.9983
49	0.9993
50	0.9997
51	0.9999
52	1.0000
53	1.0000
54	1.0000
55	1.0000
56	1.0000
57	1.0000
58	1.0000
59	1.0000
60	1.0000
61	1.0000
62	1.0000
63	1.0000
64	1.0000
65	1.0000
66	1.0000
67	1.0000
68	1.0000
69	1.0000
70	1.0000