

Examen réparti 2

Durée : 2h00

*Seuls documents autorisés : deux feuille A4 recto-verso et calculatrice
– Barème indicatif –*

Exercice 1 (5 points) – Des souris et des binomiales

La loi binomiale $\mathcal{B}(n, p)$ décrit le comportement d'une expérience où l'on répète n fois une même épreuve comportant 2 issues : 0 (échec) et 1 (succès) ; et dont la probabilité de succès est p . Soit X une variable suivant une telle loi $\mathcal{B}(n, p)$, X peut alors prendre une valeur parmi $\{0, \dots, n\}$ (il y a de 0 à n succès possibles pour n épreuves) et :

$$\forall k \in \{0, \dots, n\}, p(X = k) = \mathbb{C}_n^k \cdot p^k \cdot (1 - p)^{n-k}$$

Q 1.1 *Espérance et variance d'une loi binomiale*

En remarquant qu'une variable X suivant une loi binomiale peut s'écrire comme la somme de variables X_i indépendantes suivant toutes une même loi de Bernoulli, calculer l'espérance et la variance de X .

Q 1.2 *Loi binomiale et souris*

Dans un laboratoire, on reçoit par lot de 4 des souris, blanches ou grises. On observe sur un envoi de 200 lots que :

- 13 lots n'ont pas de souris blanche,
- 65 lots contiennent une unique souris blanche,
- 72 lots contiennent exactement 2 souris blanches,
- 35 lots contiennent exactement 3 souris blanches,
- 15 lots contiennent exactement 4 souris blanches,

Les biologistes du laboratoire se demandent si cette répartition est compatible avec la théorie génétique selon laquelle il y a autant de souris blanches que de souris grises ?

Q 1.2.1 Quel type de test doit-on donc effectuer ?

Q 1.2.2 Exprimer H_0 . Soit X la variable représentant le nombre de souris blanches dans un lot, quelle loi doit-elle suivre sous H_0 ?

Q 1.2.3 Quels seraient alors les effectifs attendus pour les lots comprenant de 0 à 4 souris blanches ?

Q 1.2.4 Vérifier si H_0 est acceptable pour un seuil de signification $\alpha = 0.05$.

Exercice 2 (5pts) – Cigarettes & famille

On appelle "comportement de fumeur" le fait qu'une personne fume ou non. On étudie l'impact du genre et de l'environnement familial comme facteurs du comportement de fumeur. Pour cela, on a recueilli les données suivantes parmi des étudiants :

	Homme	Femme	Total	Père F	Père F	Père NF	Père NF
				Mère F	Mère NF	Mère F	Mère NF
F	24	41	65	13	16	7	29
NF	23	35	58	5	24	6	23
Total	47	76	123	18	40	13	52

F=Fumeur, NF=Non fumeur

Ainsi, on peut, par exemple, lire dans ce tableau que, sur les 123 individus de l'échantillon, 65 sont fumeurs, dont 41 femmes, 29 sans aucun parent fumeur, etc.”.

On considérera des tests au seuil de 95%.

Q 2.1 Le comportement de fumeur est-il indépendant du genre ?

Q 2.2 Le comportement de fumeur est-il indépendant des comportement de fumeur de ses parents ?

Q 2.3 Les comportements de fumeur des conjoints sont-ils indépendants ?

Q 2.4 Le tabagisme se développe-t-il chez les femmes ? Chez les hommes ?

Exercice 3 (9pts) – Explorateur markovien



1		7	8	9
2	4	6		10
3		5		11

Le labyrinthe ci-contre est constitué de 11 pièces. L'explorateur démarrera toujours sur la case 1 et il peut passer d'une pièce à l'autre seulement si elle sont adjacentes (les mouvements en diagonale ne sont pas possibles).

Par exemple, de la salle 1, l'explorateur peut choisir de rester sur place ou d'aller en salle 2 : il n'y a pas d'autre possibilité.

Par ailleurs, à chaque pas de temps, l'explorateur effectue un mouvement qu'il choisit en suivant une distribution uniforme sur l'ensemble des déplacements possibles depuis la case courante.

Q 3.1 Peut-on représenter le déplacement de l'explorateur comme une chaîne de Markov ? Justifier en une phrase.

Q 3.2 Donner la distribution initiale π_0 , le graphe de transition G et la matrice de transition M . Le problème étant fastidieux et répétitif pour M , limitez-vous aux 6 premières lignes. Rappeler les dimensions de Π et M .

Q 3.3 Démontrer le fait que le système est ergodique. Quelle est la conclusion pour la chaîne de Markov ? Quelle méthode de résolution peut-on utiliser pour calculer EXACTEMENT la distribution stationnaire ?

Q 3.4 En supposant que vous ayez une fonction `solve` qui, à partir d'une matrice A et d'un vecteur B , calcule le vecteur $X = \text{solve}(A, B)$, tel que $A \cdot X = B$, donner un algorithme qui trouve la distribution stationnaire π^* .

Q 3.5 On note π^* la distribution stationnaire de l'explorateur. On s'intéresse à la limite des (M^n) (la suite des puissance de M) ? On note M^* cette limite. Quelle est la relation entre M^* et π^* ?

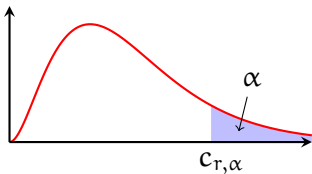
Q 3.6 Proposer un algorithme pour calculer expérimentalement (pas EXACTEMENT) le temps moyen pour que l'explorateur atteigne le trésor (case 11) ?

Q 3.7 Afin d'aider notre explorateur, le maître du jeu indique avant la quête que le trésor se trouve au sud-est (cf boussole sur le schéma). Comment modifier les paramètres de transition A pour prendre en compte cette information afin que l'explorateur trouve le

trésor plus vite ? La modification optimale est-elle triviale –expliquer en une phrase– ? Dans la négative, proposer un protocole pour trouver empiriquement les paramètres de transition qui permettent de trouver le trésor le plus vite possible.

Extrait de la table du χ^2

Tableau des $c_{r,\alpha}$ tels que $P(Z \geq c_{r,\alpha}) = \alpha$ avec $Z \sim \chi^2_{(r)}$



r \ α	0,995	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,005
1	0,0000393	0,000157	0,000982	0,00393	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,103	0,211	4,61	5,99	7,38	9,21	10,6
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,3	12,8
4	0,207	0,297	0,484	0,711	1,06	7,78	9,49	11,1	13,3	14,9
5	0,412	0,554	0,831	1,15	1,61	9,24	11,1	12,8	15,1	16,7
6	0,676	0,872	1,24	1,64	2,20	10,6	12,6	14,4	16,8	18,5
7	0,989	1,24	1,69	2,17	2,83	12,0	14,1	16,0	18,5	20,3
8	1,34	1,65	2,18	2,73	3,49	13,4	15,5	17,5	20,1	22,0
9	1,73	2,09	2,70	3,33	4,17	14,7	16,9	19,0	21,7	23,6
10	2,16	2,56	3,25	3,94	4,87	16,0	18,3	20,5	23,2	25,2
11	2,60	3,05	3,82	4,57	5,58	17,3	19,7	21,9	24,7	26,8
12	3,07	3,57	4,40	5,23	6,30	18,5	21,0	23,3	26,2	28,3
13	3,57	4,11	5,01	5,89	7,04	19,8	22,4	24,7	27,7	29,8
14	4,07	4,66	5,63	6,57	7,79	21,1	23,7	26,1	29,1	31,3
15	4,60	5,23	6,26	7,26	8,55	22,3	25,0	27,5	30,6	32,8