STT1000

Chapitre 2 Lois discrètes

<u>Exercices</u>

Loi binomiale

2.1 On sait par expérience que 70% des personnes atteintes d'un certain type de cancer finissent par en mourir. Dans une petite ville, 5 habitants sont atteints de la maladie.

a) Quelle est la probabilité qu'exactement deux des cinq finissent par en mourir? 0,1323

b) Quelle est la probabilité qu'au moins une des cinq finisse par en mourir? 0,99757

2.2 Dix pour-cent des pièces produites par une machine sont défectueuses. Quelle est la probabilité pour que dans un échantillon de 100 pièces tirées au hasard on trouve 10 pièces défectueuses? 0,1318653

2.3 On estime à 13 % le pourcentage de gauchers dans la population. Quelle est la probabilité que dans une classe de 20 enfants il y ait a) exactement 2 gauchers ? b) moins de deux gauchers ? c) plus de 2 gauchers ?

2.4 Un examen de 8 questions objectives donne pour chaque question un choix de quatre réponses dont une seule est correcte. Pour réussir l'examen il faut avoir au moins 5 bonnes réponses. Quelle est la probabilité qu'un étudiant qui ne connaît pas sa matière et répond au hasard réussisse? 0,02729797

Loi hypergéométrique

2.5 On tire au hasard un échantillon de 5 personnes d'une classe de 12 personnes dont 4 sont des fumeurs. Soit X le nombre de fumeurs observés dans l'échantillon. Déterminer la distribution complète (c'est-à-dire, les probabilités $P(X = x)$ pour $x = 0, 1, \ldots, 4$). Comparer les résultats avec l'approximation par la loi binomiale.

2.6 À la fin d'une fête à laquelle participent 20 personnes, dont 12 femmes, on tire au hasard (sans remise) les noms de 7 personnes qui se voient offrir un cadeau.

a) Quelle est la probabilité qu'exactement deux des gagnants soient des femmes? 0,0476780186

b) Quelle est la probabilité qu'il y ait au plus 3 femmes parmi les gagnants? 0,250774

Loi géométrique

2.7 Un joueur à la roulette mise toujours sur le noir, avec l'intention de s'arrêter au premier gain. [On suppose que la probabilité d'avoir noir à la roulette est 18/38].

a) Quelle est la probabilité qu'il doive jouer exactement 8 fois?

b) Quelle est la probabilité qu'il doive jouer plus de 8 fois ? 0,005888046

c) Quelle est la probabilité qu'il doive jouer 8 fois ou moins ? 0,994112

2.8 Un couple décide d'avoir des enfants jusqu'à ce qu'il ait au moins un enfant de chaque sexe.

a) Quelle est la probabilité qu'il ait 4 enfants ? 1/8.

b) Quelles sont l'espérance et la variance du nombre d'enfants qu'il aura?

2.9 On lance un dé jusqu'à ce qu'apparaisse la face «6». Quelle est la probabilité que le dé soit lancé exactement 8 fois 0,04651361 ? 8 fois ou plus ? 0,2790816.

Loi binomiale négative

2.10 On lance un dé jusqu'à ce que la face « 6 » soit obtenue pour la 10^e fois. Soit X le nombre de lancers effectués au moment du 10^e « 6 ».

a) Déterminer les probabilités suivantes: i) $P(X = 30)$; ii) $P(X > 30)$; iii) $P(X \le 30)$; iv) $P(25 \le X \le 30)$.

b) Déterminer l'espérance mathématique et la variance du nombre de lancers requis.

2.11 Afin de constituer un échantillon aléatoire de 20 ménages francophones, on tire au hasard des numéros de téléphones dans une population dans laquelle 60 % des ménages sont francophones. Soit X le nombre de ménages qu'il faudra tirer pour atteindre cette cible.

a) Déterminer les probabilités suivantes: i) $P(X = 30)$; ii) $P(X > 30)$; iii) $P(X \le 30)$; iv) $P(25 \le X \le 30)$.

b) Déterminer l'espérance mathématique et la variance du nombre d'appels effectués.

Loi de Poisson

2.12 Supposons que le nombre d'erreurs typographiques dans les pages d'un livre est une variable de loi de Poisson de paramètre $\lambda = 0{,}8$. On tire une page au hasard, et on n'y trouve aucune erreur.
   a) Calculer les probabilités suivantes: i) $P(X = 0)$; ii) $P(X > 3)$; iii) $P(X \leq 3)$.
   b) Supposons que λ n'est pas connu. Quelle est la plus petite valeur de λ pour laquelle $P(X = 0) \leq 0{,}05$?

2.13 Soit X une variable de loi binomiale de paramètres $n = 8$ et $p = 0{,}06$. Déterminer la distribution complète (c'est-à-dire, les probabilités $P(X = x)$ pour $x = 0, 1,\ldots,10$). Comparer les résultats avec ceux que vous auriez obtenus en utilisant la loi de Poisson (calculer le pourcentage d'erreur pour chaque x, et commenter).

2.14 Le taux de naissance au Canada est d'environ 43 naissances par heure. Quelle est la probabilité que durant les 5 prochaines minutes il y ait 3 naissances ou plus? 0,6942913 Quelle est la probabilité que 10 minutes s'écoulent sans aucune naissance ?

2.15 Dans un pétrin contenant la quantité de pâte nécessaire pour faire 100 gâteaux, on ajoute 1000 raisins secs. Quelle est la probabilité qu'un gâteau tiré de ce pétrin contienne exactement 10 raisins secs?

2.16 Un typographe fait en moyenne une erreur par page. Quelle est la probabilité que dans un livre de 20 pages, il y ait au moins 10 pages sans erreur? (Supposez que le nombre d'erreurs dans une page suit une loi de Poisson).

Loi multinomiale

2.17 Supposons que lors d'une élection, trois candidats, A, B, et C sont en lice, et que dans la population, les pourcentages d'électeurs qui voteront pour A, B et C sont 45 %; 25 %; et 20 %, respectivement, tandis que 10 % de la population est indécise. Dans un échantillon de 30 personnes,
   a) Quelle est la probabilité de trouver12 personnes en faveur de A; 8 pour B; 6 pour C et 4 indécis? 0,005351913
   b) Quelle est la probabilité de trouver 8 indécis? 0,00576379
   c) Quelle est la probabilité de trouver 20 personnes en faveur de A *ou* B? 0,1415617

2.18 Des vis spécialisées sont vendues en paquets de 3. Le taux de défectuosité de ces vis est de 15 %. On tire un échantillon de 40 paquets. Soit X_0, X_1, X_2 et X_3 le nombre de paquets avec, respectivement, 0, 1, 2, et 3 vis défectueuses. Déterminer la probabilité $P(X_0 = 18;\ X_1 = 10;\ X_2 = 2;\ X_3 = 0)$. 0,03829103.

Exercices divers

2.19 On suppose que dans une certaine ville, il se produit en moyenne 1,5 décès par jour. Calculez la probabilité que, la semaine prochaine (7 jours) il y ait
   a) exactement 8 décès; 0,1009.
   b) exactement 2 jours sans décès $P(Y = 2) = 0{,}29585$
   c) au moins un décès chaque jour. 0,1708

2.20 Un laboratoire qui effectue sur une grande échelle des tests pour détecter un certain anticorps peut épargner de l'effort en faisant un seul test sur plusieurs spécimens à la fois. Lorsque l'anticorps n'est pas présent dans l'ensemble des spécimens, c'est parce qu'il n'est présent dans aucun. On déclare alors un résultat négatif pour tous les patients sans plus de tests. Si le résultat est positif, cependant, on analyse chaque spécimen séparément.
   a) Si on utilise cette approche avec 10 spécimens d'une population dont une certaine proportion p ont l'anticorps en question (sont «positifs»), quelle est l'espérance du nombre de tests qu'il faudra effectuer (i) si $p = 0{,}10$, et (ii) si $p = 0{,}25$.
   b) Pour quelles valeurs de p l'approche décrite ici est-elle préférable à l'approche usuelle (tester les 10 spécimens séparément)?
   c) Si n est le nombre de spécimens qu'on groupe, montrer que l'approche décrite ici est préférable à l'approche usuelle si et seulement si $p < 1 - (1/n)^{1/n}$.

2.21 Une fabrique produit des bouteilles dont 20 % sont défectueuses. Une bouteille est défectueuse si la probabilité qu'elle se casse lorsqu'on la laisse tomber d'une hauteur de 1 mètre est 0,7. Pour une bouteille non défectueuse, la probabilité qu'elle se casse n'est que de 0,1.
   a) Vous tirez au hasard une bouteille. Quelle est la probabilité qu'elle se casse lorsque vous la laissez tomber d'une hauteur de 1 m ?
   b) Vous tirez au hasard 4 bouteilles et vous les laissez toutes (une à une) tomber d'une hauteur de 1m. Quelle est la probabilité que 2 d'entre elles se cassent ?

c) Vous tirez au hasard 4 bouteilles et vous les laissez tomber d'une hauteur de 1 m. Quelle est la probabilité que l'événement « toutes les défectueuses et seules les défectueuses se cassent »?

d) Vous tirez au hasard 4 bouteilles et vous les laissez tomber d'une hauteur de 1m. Deux des 4 bouteilles se sont cassées.
 (i) Quelle est la probabilité que les 4 bouteilles aient été défectueuses ?
 (ii) Quelle est la probabilité que 3 des 4 bouteilles de l'échantillon aient été défectueuses ?

e) Vous tirez une bouteille au hasard, la laissez tomber 4 fois. Elle ne se casse pas. Quelle est la probabilité qu'elle soit défectueuse ? [Faites les suppositions d'indépendance qu'il faut: une bouteille n'est pas « affaiblie » par le choc]

2.22 Deux sortes de véhicules arrivent régulièrement à un poste de péage: les véhicules privés et les véhicules commerciaux. Entre 10 heures et 15 heures, le nombre moyen de véhicules privés est de 2 à la minute, alors que le nombre moyen de véhicules commerciaux est de 3 à la minute.

a) Quelle est la probabilité qu'en une minute 2 véhicules privés et 2 véhicules publics arrivent au poste de pé age?

b) Quelle est la probabilité qu'en une minute, 4 véhicules arrivent au poste de péage?
0,17554674

c) Quelle est la probabilité conditionnelle qu'en une minute 2 véhicules privés et 2 véhicules publics arrivent au poste de péage étant donné que 4 véhicules sont arrivés au poste de péage ?

2.23 Au coin de la rue, il passe en moyenne un taxi à chaque 3 minutes mais 40% seulement de ces taxis sont inoccupés. J'ai besoin d'un taxi.

a) Quelle est la probabilité que les 3 premiers taxis à passer soient occupés? 0,216

b) Quelle est l'espérance du nombre de taxis occupés qui précéderont l'arrivée du premier taxi libre ? 1,5

c) Quelle est la probabilité qu'aucun taxi libre n'arrive durant les 20 premières minutes ?

d) Montrer, plus généralement, que le nombre de taxis libres qui arrivent en 20 minutes est de loi de Poisson de paramètre 0,4λ.

2.24 Un certain défaut dans la fabrication de plaques d'émail se présente sous la forme de minuscules taches blanches sur la surface de l'émail. Admettons que le nombre de taches X sur une plaque suit une loi de Poisson de paramètre λ. Si $\lambda=1$, calculer

a) la probabilité qu'une plaque d'émail contienne 2 taches; 0,1839397

b) la probabilité qu'une plaque d'émail contienne deux taches ou plus; 0,2642411

c) la probabilité que 3 plaques d'émail contiennent en tout 3 taches;
0,2240418.

d) la probabilité que 3 plaques d'émail contiennent chacune une tache. 0,04978707 ($\lambda = 1$).

e) la probabilité que 3 plaques d'émail contiennent chacune une tache, étant donné que les 3 plaques contiennent 3 taches en tout. 0,2222222.

f) Montrez que la probabilité que k plaques contiennent chacune une tache, étant donné qu'elles en contiennent en tout k est $k!/k^k$. Voir g).

g) Montrez que votre réponse en f) demeure vraie même si $\lambda \neq 1$.

2.25 Dans un casino il y a trois machines dans lesquelles les probabilités de gagner sont 1/3, 1/2 et 2/3, respectivement. Vous avez choisi une machine au hasard et gagné 4 fois en 4 essais. Quelle est la probabilité de gagner la prochaine fois (avec la même machine)?
0,613314

2.26 Deux urnes contiennent chacune une proportion p de boules rouges. On tire avec remise n_1 boules de l'urne 1 et n_2 boules de l'urne 2. Sachant que m des $n_1 + n_2$ boules tirées sont rouges, quelle est la probabilité que le nombre de boules rouges tirées de l'urne 1 soit k?

Tests d'hypothèses

2.27 Afin de contrôler la qualité des pièces produites par une machine, on prélève de temps en temps un échantillon de 10 pièces. On arrête la production pour inspecter la machine si dans l'échantillon on trouve plus de deux pièces défectueuses ; autrement, on laisse la machine fonctionner. Supposons qu'en fait 20% de la production est défectueuse. Quelle est la probabilité qu'après un échantillonnage de 10, on la laisse fonctionner ?

2.28 Un célèbre magicien qui prétendait avoir des pouvoirs de perception extra-sensorielle a accepté de se livrer à une expérience dans laquelle il se proposait de deviner le résultat du lancer d'un dé. En 12 essais, il a réussi à deviner le résultat 10 fois. Vérifier que la probabilité d'un nombre de succès aussi grand que 10 (c'est-à-dire supérieur ou égal à 10) est excessivement petite pour quelqu'un qui répond au hasard; et expliquer à quelle conclusion ce fait a tendance à mener.

2.29 Un certain test psychologique consiste à lire un paragraphe, et puis répondre à 20 questions portant sur le texte lu. Un choix de 5 réponses est donné pour chaque question. Un évaluateur, tentant de démontrer que le test ne mesure pas l'aptitude à la lecture, répond aux 20 questions sans avoir lu le texte. Il choisit la bonne réponse à 8 des questions. Calculer la probabilité d'avoir 8 succès ou plus, et discuter les implications sur la qualité du test.

2.30 Une compagnie se fait accuser de discrimination pour avoir engagé 6 hommes et une femme pour 7 postes identiques alors que des 17 candidats qui s'étaient présentés, 9 étaient des femmes. Calculer la probabilité d'avoir si peu de femmes (c'est-à-dire, une ou moins) en supposant un choix au hasard. Est-ce que ce calcul de probabilité peut contribuer au débat?

2.31 Il existe des conjectures selon lesquelles certaines personnes sont capables (dans une certaine mesure) de surseoir à leur mort afin de pouvoir une dernière fois vivre un des bons moments de la vie. Définissant un anniversaire de naissance comme un de ces bons moments, des chercheurs ont prélevé les dates de naissance et de mort dans un échantillon de 500 décès. Ils ont constaté que sur ces 500 décès, 5 sont survenus le jour même de l'anniversaire du décédé. Ce nombre est supérieur à la normale, mais l'est-il assez pour confirmer les conjectures ?

2.32 On suppose que dans une certaine région, la proportion des gens qui sont en faveur du mariage entre personnes de même sexe est $p = 40\%$. Lors d'un sondage auprès de 15 personnes, on trouve $X = 11$ personnes en faveur du libre-échange.

a) Calculer la probabilité d'un écart absolu, $|X-\mathrm{E}(X)|$, aussi grand que (c'est-à-dire, supérieur ou égal à) l'écart observé de 5.

b) Étant donné la probabilité calculée en a), y a-t-il lieu de retoucher l'hypothèse que $p = 0{,}4$?

2.33 Dans un village où ont été entreposés des déchets chimiques, on constate que 8 personnes ont été atteintes d'une certaine sorte de cancer dans une période de 5 ans. Étant donné que la population du village n'est que de 8 000, ce nombre semble excessif. Une commission chargée de déterminer si les déchets chimiques ont contribué à hausser le taux prélève des données sur les populations de plusieurs villages de taille et situation comparables. La commission découvre que durant la même période, il y a eu 588 cas dans un bassin de population de 2 350 000 habitants. Considérer ce taux comme un taux normal (et connu sans erreur) pour calculer la probabilité d'avoir 8 cas ou plus dans une population de 8 000. Expliquer ce que ce calcul peut contribuer à la question posée par la commission.

2.34 Une compagnie reçoit un lot de pièces électroniques qu'elle entend tester avant de décider si elle l'accepte ou le rejette. On considère acceptable une proportion de pièces défectueuses de 5 %, mais on ne peut tester toutes les pièces du lot. On décide donc de tester un échantillon de 25 pièces, et de suivre la règle suivante:

S'il y a 3 pièces défectueuses *ou plus* dans l'échantillon,
on *rejette* le lot. S'il y en a moins de 3, on l'accepte.

a) Quelle est la probabilité de rejeter le lot si en fait il est juste acceptable ? 0,127106.

b) Quelle est la probabilité de rejeter le lot si en fait il est mieux qu'acceptable, dans le sens que seulement 3% de ses pièces sont défectueuses ? 0,03796.

c) Quelle est la probabilité d'accepter le lot si en fait il n'est *pas* acceptable, dans le sens que 10 % de ses pièces sont défectueuses? 0,5371.

d) Comment doit-on changer la règle ci-dessus si on tient à ce que la probabilité de rejeter le lot lorsque celui-ci est acceptable (c'est-à-dire, lorsqu'il a exactement 5 % de pièces défectueuses) soit $\leq 1\%$?

e) Avec la règle développée en d), quelle est la probabilité d'accepter le lot si en fait il n'est pas acceptable, dans le sens que 10 % de ses pièces sont défectueuses ? 0,9020064.

f) La règle développée en d) est différente de celle énoncée au début de cet exercice. Expliquez brièvement les avantages et désavantages des deux règles.

2.35 Afin d'estimer le nombre N de truites dans un lac on réalise l'expérience suivante: on prélève 100 truites du lac puis, après les avoir marquées, on les remet dans l'eau. Plus tard on repêche 200 truites du lac et on observe le nombre X de truites marquées dans ce second prélèvement. Si $X = 5$, quelle valeur de N vous paraît la plus vraisemblable ?

2.36 Supposons que le nombre d'erreurs typographiques dans les pages d'un livre est une variable de loi de Poisson de paramètre λ. On tire une page au hasard, et on n'y trouve aucune erreur.

a) Calculer $\mathrm{P}\{X = 0\}$ en supposant que $\lambda = 5$. La valeur $\lambda = 5$ est-elle plausible?

b) Calculer $\mathrm{P}\{X = 0\}$ en supposant que $\lambda = 1$. La valeur $\lambda = 1$ est-elle plausible?

c) Convenons d'appeler « plausible » toute valeur de λ pour laquelle $P\{X = 0\} \geq 0{,}05$. Quel est l'ensemble des valeurs plausibles de λ?
$\lambda \leq 2{,}9957$

2.37 Une compagnie reçoit régulièrement des lots de plaques d'émail. On considère acceptable un lot dont le nombre moyen α de défauts (des taches blanches dans l'émail) par plaque est $\leq 0{,}10$. On décide qu'à chaque réception d'un lot on examinera un échantillon de 25 plaques afin de compter le nombre total X de taches dans les 25 plaques. On convient de la règle suivante: on rejettera le lot si $X \geq 4$.

a) Quelle est la probabilité de rejeter un lot pour lequel α est précisément égal à 0,10? 0,242424

b) Quelle est la probabilité de rejeter un lot qui est en fait mieux qu'acceptable dans le sens que $\alpha = 0{,}05$? ,03827

c) Quelle est la probabilité de rejeter un lot qui est en fait inacceptable, dans le sens que $\alpha = 0{,}15$? 0,5162

d) Quelle est la probabilité d'accepter un lot qui est inacceptable dans le sens que $\alpha = 0{,}2$? 0,26503

e) Comment doit-on changer la règle ci-dessus si on tient à ce que la probabilité de rejeter un lot acceptable avec ($\alpha = 0{,}10$) ne soit pas supérieure à 5%? 0,0420.

f) Avec la règle développée en e), quelle est la probabilité d'accepter un lot qui est inacceptable, dans le sens que $\alpha = 0{,}2$? 0,61596

g) Avec la règle développée en e), quelle est la probabilité de rejeter un lot pour lequel α est précisément égal à 0,10? 0,04202

h) Commenter les réponses en d) et en f); laquelle des deux procédures est-elle préférable selon cette comparaison?

i) Commenter les réponses en a) et en g); laquelle des deux procédures est-elle préférable selon cette comparaison?

j) Faire une synthèse des réponses en h) et en i).

Exercices théoriques

2.38 Soit $\hat{p}$ la proportion des personnes en faveur du libre-échange dans un échantillon de taille $n = 20$ d'une population dont 60% sont en faveur du libre-échange.

a) Déterminer $E[\hat{p}]$ et $Var[\hat{p}]$.

b) Déterminer $P[|\hat{p} - 0{,}6| > 0{,}3]$

2.39 Si X est de loi $\mathcal{B}(n; p)$ alors X peut s'exprimer comme une somme $X = X_1 + X_2 + \ldots + X_n$ où X_i est le nombre de succès au i^e tirage.

Utilisez ce fait pour montrer que $E(X) = np$ et $Var(X) = npq$.

2.40 Soit $X \sim \mathcal{B}(n; p)$. Montrer que

$$P(X = k) = \begin{cases} (1-p)^n & k = 0 \\ \dfrac{n-k+1}{k}\,\dfrac{p}{1-p}\,P(X = k-1) & k \geq 1 \end{cases}$$

Cette formule est particulièrement utile lorsque n est grand et k petit.

2.41 Soit n épreuves indépendantes avec probabilité p de succès. Déterminer l'espérance, la variance et la fonction de probabilité de la variable Y = nombre de succès moins nombre d'échecs. [Y est une fonction simple du nombre de succès]

2.42 Montrez que la fonction de probabilité d'une variable de loi binomiale de paramètres n et p est non décroissante pour $x \leq (n+1)p$.

2.43 Soit X une variable de loi $\mathcal{B}(n; p)$. Montrez que pour x fixe, $F(x) = P(X \leq x)$ est fonction décroissante de p [*Suggestion*: *commencez par supposer que* $x \leq np$ *et démontrer le résultat en dérivant* $F(x)$ *par rapport à* p; si $x > np$, *montrez que* $1-F(x)$ *est croissante en* p]

2.44 Soit X_n une variable aléatoire de loi $\mathcal{B}(n; p)$, $F_n(x) = P(X_n \leq x)$ et $p_n(x) = P(X_n = x)$.

a) Montrer par un argument probabiliste que $p_n(x) = p_{n-1}(x-1) \times p + p_{n-1}(x) \times q$ et que $F_n(x) = F_{n-1}(x) - p\,p_{n-1}(x)$.

b) Montrer que $F_n(x) = (n-x)\binom{n}{x}\int_0^q t^{n-x-1}(1-t)^x\,dt$.

c) Montrer que $\sum_{\nu=0}^{x} p_{n_1}(\nu)\, p_{n_2}(x-\nu) = p_{n_1+n_2}(x)$ et donnez-en une interprétation probabiliste.

2.45 Soit X une variable de loi $\mathcal{B}(n; p)$, et soit $\hat{p} = X/n$. Montrer que $\mathrm{E}[\hat{p}(1-\hat{p})] = \dfrac{(n-1)p(1-p)}{n}$.

2.46 Soit a et b deux entiers, $a \le b$ et $X \sim \mathcal{G}(p)$. Montrer que $\mathrm{P}\{X > b \mid X > a\} = \mathrm{P}\{X > b - a\}$.

2.47 Soit X une variable de loi géométrique de paramètre p.
a) Montrer que $\mathrm{E}(1/X) = -p\,(\ln p)/(1-p)$;
b) Considérer une famille qui a l'intention d'avoir autant d'enfants que nécessaire pour avoir un garçon (et de s'arrêter dès qu'ils en ont un). Déterminer l'espérance de la *proportion* de garçons dans leur ménage.

2.48 Si $X \sim \bar{\mathcal{B}}(n; p)$, démontrer que

$$\sum_{x=n}^{\infty}\binom{x-1}{n-1}p^n(1-p)^{x-n} = 1.$$

(*Se servir du développement de Taylor suivant* $(1-t)^{-n} = \sum_{y=0}^{\infty}\binom{n-1+y}{n-1}t^y$).

2.49 Soit X une variable de loi $\bar{\mathcal{B}}(n; p_n)$. Montrer que si $n \to \infty$ et $q_n \to 0$ de telle sorte que $nq_n \to \lambda$, alors $P(X = n+u) \to \dfrac{e^{-\lambda}\lambda^u}{u!}$.

2.50 Il est utile d'évaluer la fonction de probabilité hypergéométrique à l'aide de la formule récursive suivante

$$P(X = x+1) = \frac{(N_1 - x)(n-x)}{(x+1)(N_2 - n + x + 1)} \times P(X = x)$$

Démontrer la validité de cette formule, et s'en servir pour évaluer la fonction de probabilité $\mathcal{H}(5; 8, 10)$.

x	0	1	2	3	4	5
$p(x)$	0,0294	0,1961	0,3922	0,2941	0,0817	0,0065

2.51 Soit X_1 et X_2 deux variables aléatoires indépendantes, X_1 de loi $\mathcal{B}(N_1; p)$ et X_2 de loi $\mathcal{B}(N_2; p)$. Soit $Y = X_1 + X_2$. Montrer que la distribution conditionnelle de X_1 étant donné que $Y = n$ est $\mathcal{H}(n; N_1, N_2)$.

Rappel: $\sum_{u=0}^{n}\binom{N_1}{u}\binom{N_2}{n-u} = \binom{N_1+N_2}{n}$, et $X_1 + X_2 \sim \mathcal{B}(N_1 + N_2; p)$.

2.52 Soit X une variable de loi hypergéométrique de paramètres n, N_1 et N_2. Montrer que lorsque $N_1 \to \infty$, $N_2 \to \infty$ de telle sorte que $N_1/N = p$ ($N = N_1 + N_2$), alors la fonction de probabilité de X tend vers celle d'une variable $\mathcal{B}(n; p)$. *Suggestion*: Montrer que la fonction de probabilité d'une variable de loi hypergéométrique peut s'écrire de la façon suivante:

$$\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} = \binom{n}{x}\frac{N_1(N_1-1)\ldots(N_1-(x-1))\; N_2(N_2-1)\ldots(N_2-(n-x-1))}{N(N-1)\ldots(N-(n-1))}$$

$$= \frac{\binom{n}{x}}{(1-1/N)\ldots(1-(n-1)/N)} \times \frac{N_1}{N}\left(\frac{N_1}{N}-\frac{1}{N}\right)\ldots\left(\frac{N_1}{N}-\frac{x-1}{N}\right)\times\frac{N_2}{N}\left(\frac{N_2}{N}-\frac{1}{N}\right)\ldots\left(\frac{N_2}{N}-\frac{n-x+1}{N}\right).$$

2.53 Soit X et Y deux variables aléatoires indépendantes, de loi de Poisson de paramètres λ_1 et λ_2, respectivement. Montrer que $Z = X + Y$ est de loi de Poisson de paramètre $\lambda_1 + \lambda_2$.

2.54 Soit X et Y deux variables aléatoires indépendantes, de loi de Poisson de paramètres λ_1 et λ_2, respectivement. Montrer que la distribution conditionnelle de X étant donné $X + Y = n$ est $\mathcal{B}\left(n\,;\frac{\lambda_1}{\lambda_1+\lambda_2}\right)$.

2.55 Un gâteau est censé contenir des raisins blancs et des raisins rouges. Le nombre de raisins blancs est de loi de Poisson de paramètre λ_1 et le nombre de raisins rouges suit une loi de Poisson de paramètre λ_2. Soit X le nombre de raisins blancs et Y le nombre de raisins rouges, et soit $Z = X + Y$ le nombre total de raisins.

a) Déterminez une expression pour $P(X = x$ et $Y = y)$, c'est-à-dire pour la probabilité qu'il y ait x raisins blancs et y raisins rouges.

b) Montrez que Z suit une loi de Poisson. Quel est le paramètre de cette loi ?

c) Déterminez $P(Y = y \mid Z = z)$, c'est-à-dire la probabilité qu'il y ait y raisins rouges sachant qu'il y a en tout z raisins. (Montrez qu'il s'agit d'une loi binomiale et identifiez-en les paramètres).

2.56 Montrer que $\sum_{i=0}^{\infty}\frac{e^{-\lambda}\lambda^i}{i!}=\frac{1}{x!}\int_{\lambda}^{\infty}e^{-t}t^x\,dt$.

2.57 Soit $\boldsymbol{U} = [X; Y; Z]$ un vecteur de loi multinomiale de paramètres n, p_1, p_2, et p_3.

a) Montrer que $X + Y$ est de loi $\mathcal{B}(n; p_1 + p_2)$;

b) Montrer que la distribution conditionnelle de X étant donné $X + Y = m$ est $\mathcal{B}(m; p_1/(p_1+p_2))$.

2.58 n personnes participent à une loterie dans laquelle il y a r gagnants, où r croît avec le nombre de participants de telle sorte que $\lambda = r/n$, le nombre de prix par participant reste fixe. Montrez que lorsque le nombre de participants (et donc de prix) tend vers l'infini, alors la probabilité qu'un participant donné gagne x prix tend vers $e^{-1}\lambda^x/x!$

Exercices à faire à l'ordinateur

2.59 Calculez les valeurs de la fonction de probabilité d'une variable $X \sim \mathcal{B}(15; 0{,}03)$, et vérifiez numériquement les propriétés $E(X) = np$ et $Var(X) = npq$.

2.60 Calculez les valeurs de la fonction de probabilité d'une variable $X \sim \mathcal{H}(15; 20, 30)$, et vérifiez numériquement les propriétés $E(X) = np$ et $Var(X) = npq(N-n)/(N-1)$.

2.61 Calculez les valeurs de la fonction de probabilité d'une variable de loi $\mathcal{P}(0{,}3)$ pour $x = 0, 1, \dots, 20$, et vérifiez approximativement les propriétés $E(X) = \lambda$ et $Var(X) = \lambda$.

2.62 Comparez graphiquement la fonction de probabilité $\mathcal{B}(30; 0{,}01)$ à celle d'une variable de loi $\mathcal{P}(0{,}3)$. Répétez avec la fonction de répartition.

Lois, variance, indépendance

L'application de la statistique à un problème particulier débute toujours par un processus de modélisation, un processus qui consiste à décrire le contexte le plus fidèlement possible en termes mathématiques. Ce qui exige normalement certains jugements concernant l'indépendance de variables aléatoires, leur variance, leur loi. Les problèmes qui suivent ont pour but d'encourager une réflexion sur ces questions.

2.63 Pour chacune des variables aléatoires X suivantes, dites quelle loi pourrait s'appliquer et déterminez $E(X)$ et $Var(X)$. Dites si d'autres lois pourraient également servir approximativement. Justifiez votre ou vos choix, en précisant les hypothèses qu'on doit supposer vraies pour appliquer la loi. Il se peut qu'aucune des lois que vous connaissez ne s'applique.

a) Dans un usine on fait fonctionner 5 machines identiques. La probabilité qu'une de ces machines tombe en panne en un jour donné est 0,15.
X = le nombre de machines qui tombent en panne.

b) On tire un échantillon de 10 lentilles cornéennes dans un lot de 10 000 lentilles dont 25 sont défectueuses.

X = le nombre de lentilles défectueuses dans l'échantillon.

c) On fait un relevé de tous les dons offerts à une certaine université en 1999. En moyenne, l'université en question reçoit un don de plus de 10 000 $ par année.
X = le nombre de dons de plus de 10 000 $.

d) On observe l'état des 500 autobus qui seront en service demain dans la ville. Les données montrent que sur les 2 000 dernières journées-autobus il y a eu 2 pannes.
X = le nombre d'autobus qui tomberont en panne demain.

e) On observe les appels reçus dans un bureau entre 13h00 et 16h00 un certain après-midi. Normalement les appels arrivent dans cet intervalle de temps au rythme d'un appel par deux minutes.
X = le nombre d'appels reçus entre 13h00 et 16h00 cet après-midi.

f) On observe 25 appels effectués par un vendeur au téléphone. Dans le passé, ce vendeur a réussi en moyenne une vente à chaque 100 appels.
X = le nombre de ventes.

g) On observe une vendeuse au téléphone pendant une journée. Dans le passé, cette vendeuse a réalisé en moyenne 5 ventes par jour.
X = le nombre de ventes.

h) Vingt-cinq personnes (8 femmes 17 hommes) posent leur candidature à 10 postes identiques.
X = le nombre de femmes parmi les dix personnes engagées.

i) Un service d'approvisionnements distribue au hasard 14 ordinateurs, 5 au service A et 9 au service B. Huit des 14 ordinateurs sont neufs.
X = le nombre d'ordinateurs neufs reçus par A.

j) Un vendeur décide qu'il terminera sa journée dès qu'il aura réussi 4 ventes. La probabilité d'une vente est de 25 % à chaque essai.
X = nombre d'appels qu'il fera pendant la journée.

k) On tire un échantillon de 100 vis dans un lot de 100 000 vis dont 250 sont défectueuses.
X = le nombre de vis défectueuses dans l'échantillon.

ℓ) On tire un échantillon de 100 adultes d'une population de 100 000 adultes dont 30 000 sont au chômage.
X = le nombre d'adultes au chômage dans l'échantillon.

m) Un interviewer engagé par une maison de sondages frappe à des portes jusqu'à ce qu'il atteigne son quota de 100 ménages francophones. Supposons que 10 % des ménages sont francophones.
X = le nombre de portes auxquelles il aura frappé. 9 000.

n) On observe les arrivées au service d'urgence d'un hôpital un lundi après-midi. On sait que les lundis après-midi il y a en moyenne 25 arrivées durant la période de 13h00 à 17h00.
X = le nombre d'arrivées un lundi de 13h00 à 13h30. 3,125.

o) On tire un échantillon de 5 oranges dans une caisse de 590 oranges dont 8 sont gâtées.
X = le nombre d'oranges gâtées dans l'échantillon.
0,06642.

p) Un vendeur décide qu'il terminera sa journée dès qu'il aura réussi une vente. La probabilité d'une vente est de 25 % à chaque essai.
X = le nombre de jours parmi les 5 jours de la semaine prochaine où il fera plus de 10 essais.

q) Le nombre d'erreurs dans une page est une variable de loi de Poisson (à peu près) de moyenne $\mu = 0{,}5$.
X = le nombre de pages sans erreur parmi les 100 premières pages.

r) Le nombre d'erreurs dans une page est une variable de loi de Poisson de moyenne $\lambda = 0{,}5$. On vérifie les pages une à une en commençant par la page 1.
X = le numéro de la première page dans laquelle on trouve une erreur.

2.64 Pour chacune des paires de variables aléatoires X et Y, dites si d'après vous $\sigma_X > \sigma_Y$ ou si $\sigma_X < \sigma_Y$:

a) X: La température le 1e janvier prochain à Montréal;
Y: La température le 1e janvier prochain à Nairobi. [*Réponse* : X]

b) X: Le poids d'une personne choisie au hasard dans une école de garçons;
Y: Le poids d'une personne choisie au hasard dans une école mixte. [*Réponse* : Y]

c) X: Le temps que vous mettez à vous rendre à l'université à pied;
Y: Le temps que vous mettez à vous rendre à l'université en métro. [*Réponse* : Y]

d) X: Le temps d'attente dans une file où il n'y a qu'une personne devant vous;
Y: Le temps d'attente dans une file où il y a 2 personnes devant vous. -------------------------------- [*Réponse* : Y]

e) X: La proportion d'objets défectueux dans un échantillon de 10 objets tirés d'une certaine population;
Y: La proportion d'objets défectueux dans un échantillon de 100 objets tirés de la même population. [*Réponse* : X]

f) X: .. Le nombre d'objets défectueux dans un échantillon de 10 objets tirés sans remise d'une certaine population;
Y: Le nombre d'objets défectueux dans un échantillon de 10 objets tirés avec remise de la même population; [*Réponse* : Y]

g) X: Le revenu moyen de 10 familles choisies au hasard dans une population;
Y: Le revenu moyen de 100 familles choisies au hasard dans la même population. [*Réponse* : X]

h) X: Le Q.I. moyen de deux personnes choisies indépendamment dans une population.
Y: Le Q.I. moyen de deux jumeaux identiques choisis dans une population de jumeaux. [*Réponse* : Y]

i) Le prix d'une action de A et le prix d'une action de B ont chacun la même espérance mathématique et la même variance. Les prix des deux actions sont des variables indépendantes.
X : La valeur totale de 100 actions de la compagnie A;
Y : La valeur totale d'un portefeuille constitué de 50 actions de chaque compagnie. [*Réponse* : X]

j) On mesure la longueur du fémur (F) et la longueur du tibia (T).
X : $F + T$ lorsque le fémur et le tibia appartiennent à la même personne ;
Y : $F + T$ lorsque le fémur et le tibia appartiennent à deux personnes. [*Réponse* : X]

k) La quantité d'ingrédient actif d'une certaine pilule a un écart-type de 0,01 mg, quelle que soit la grosseur de la pilule.
X : La quantité d'ingrédient actif dans une pilule de 10 mg ;
Y : La quantité totale d'ingrédient actif dans deux pilules de 5 mg chacune. [*Réponse* : Y]

2.65 Dans chacun des numéros suivants, on décrit une expérience et deux variables aléatoires, X et Y. Dites si X et Y sont indépendantes ou non et justifiez votre réponse.

a) Je tire au hasard deux personnes dans une salle de cours, avec remise.
X: Note au dernier examen de la première personne;
Y: Note au dernier examen de la deuxième personne.

b) Même contexte qu'en a), sauf qu'on tire sans remise.

c) Je tire au hasard une fille parmi toutes les filles du secondaire (tous les niveaux confondus)
X: Une mesure de sa force physique;
Y: Une mesure de son vocabulaire.

d) Je tire au hasard un logement dans une population de logements.
X : Nombre de chambres à coucher;
Y : Nombre de salles de bains.

e) Je tire au hasard une personne dans une grande population.
X : La longueur de son fémur;
Y : La longueur de son tibia.

f) Je tire au hasard un ménage dans une population de ménages.
X : Revenu familial;
Y : Superficie du logement.

g) Je tire au hasard deux pommes de mon pommier (qui contient un nombre quasi infini de pommes).
X: Le poids de la première pomme;
Y: Le poids de la deuxième pomme.

h) Même contexte qu'en *g*), sauf que j'ai deux pommiers de deux espèces différentes, A et B, et je choisis au hasard celui des deux dans lequel je cueillerai mes deux pommes.

i) Même contexte qu'en *g*) sauf que j'ai deux pommiers, A et B, et je cueille une pomme de chacun.
X: Le poids de la pomme cueillie dans le pommier A;
Y: Le poids de la pomme cueillie dans le pommier B.

j) Même contexte qu'en *i*) sauf que je choisis les pommiers dans un ordre aléatoire.
X: Le poids de la première pomme (qui peut avoir été cueillie dans le pommier A ou dans le pommier B);
Y: Le poids de la deuxième pomme (cueillie dans l'autre pommier).

k) Diane et Carole, deux femmes qui ne se connaissent pas, attendent un enfant :
X : Le poids du bébé de Diane;
Y : Le poids du bébé de Carole.

ℓ) Lundi prochain, je prendrai la pression artérielle de Pierre, un pensionnaire dans un foyer.
X: Sa pression systolique;
Y: Sa pression diastolique.

m) Lundi prochain, je prendrai la pression systolique de Pierre, un élève de ma classe ; et je recommencerai deux mois plus tard.
X: La pression systolique de Pierre lundi prochain;
Y: La pression systolique de Pierre deux mois plus tard.

n) Lundi prochain, je choisirai au hasard un élève de ma classe et je prendrai sa pression systolique le jour même et une deuxième fois deux mois plus tard.
X: La pression systolique lundi;
Y: La pression systolique deux mois plus tard.

o) Je tire au hasard un garçon parmi tous les garçons du secondaire II.
X: Une mesure de sa force physique;
Y: Une mesure de son vocabulaire.

p) Diane et Carole, deux sœurs, attendent un enfant :
X : Le poids du bébé de Diane;
Y : Le poids du bébé de Carole.

q) Au hasard, on choisit deux sœurs adultes ayant donné naissance à des bébés
X : Le poids du bébé de l'une d'elles;
Y : Le poids du bébé de l'autre.

Numéro supplémentaire

2.66 Dans chacun des cas suivants, on présente deux variables aléatoires, X et Y. Dites si $\sigma_x > \sigma_Y$, $\sigma_x < \sigma_Y$, $\sigma_x = \sigma_Y$, ou s'il est impossible d'y répondre sans information supplémentaire. Justifiez brièvement votre réponse (à l'aide de calculs s'il y a lieu).

a) On effectue des essais indépendants jusqu'au moment du premier succès. La probabilité de succès est p.
X : Nombre d'essais nécessaires lorsque p = 0,1
Y : Nombre d'essais nécessaires lorsque p = 0,5

b) On tire *avec* remise un échantillon de n personnes dans une salle de 40 personnes.
X : Nombre de fumeurs lorsque n = 5
Y : Nombre de fumeurs lorsque n = 20

c) On tire *avec* remise un échantillon de n personnes dans une salle de 40 personnes.
X : Proportion de fumeurs lorsque n = 5
Y : Proportion de fumeurs lorsque n = 20

d) On tire *sans* remise un échantillon de n personnes dans une salle contenant 30 personnes.
X : Nombre de fumeurs lorsque n = 5
Y : Nombre de fumeurs lorsque n = 20

e) On tire *sans* remise un échantillon de n personnes dans une salle contenant 30 personnes
X : Proportion de fumeurs lorsque n = 5
Y : Proportion de fumeurs lorsque n = 20

f) On tire *sans* remise un échantillon de 5 personnes dans une salle contenant N personnes dont une certaine proportion p sont des fumeurs.
X : Nombre de fumeurs dans l'échantillon lorsque N = 10
Y : Nombre de fumeurs dans l'échantillon lorsque N = 30

g) On tire *sans* remise un échantillon de 5 personnes dans une salle contenant N personnes dont une certaine proportion p sont des fumeurs.
X : Proportion de fumeurs dans l'échantillon lorsque N = 10
Y : Proportion de fumeurs dans l'échantillon lorsque N = 30