Chapitre 2

Lois discrètes

Ce chapitre traite de fonctions de probabilités et de lois de probabilité discrètes : Loi binomiale, loi hypergéométrique, loi géométrique, loi binomiale négative, loi de Poisson, loi multinomiale. Rappel en annexe des notions d'espérance, de variance et de covariance; de distributions conjointes, marginales et conditionnelles; de covariance et de coefficient de corrélation; d'espérance et de variance conditionnelles.

Les variables *aléatoires*, et plus particulièrement les variables aléatoires dites *discrètes*[1], partagent plusieurs des attributs des variables *empiriques* traitées au chapitre 1 : elles génèrent un ensemble de valeurs et elles sont définies par une *distribution*. La différence est dans le sens des fréquences dans le cas d'une variable aléatoire: ce sont des fréquences *attendues*[2]. De ce fait, elles sont appelées *probabilités*. Ce qui ne change rien au développement mathématique, puisque les probabilités jouissent des mêmes propriétés que les fréquences : elles sont toutes comprises entre 0 et 1, et leur somme donne 1.

Considérons, par exemple, la distribution du nombre de pièces dans une population de 50 appartements :

Nombre de pièces	*Effectif*	*Fréquence*
2	10	0,20
3	20	0,40
4	13	0,26
5	7	0,14
Total	50	1

Les fréquences 0,20; 0,40; 0,26; et 0,14 sont proviennent d'un recensement de la population. C'est la distribution d'une variable empirique. Mais le même tableau peut provenir d'une variable aléatoire. Si on tire un appartement au hasard dans la population, et X désigne le nombre de pièce dans l'appartement obtenu, alors ce tableau est la *fonction de probabilité* de X.

Une *fonction de probabilité* p fait correspondre à chaque valeur d'une variable aléatoire X la probabilité $p(x) = P(X = x)$. Une *loi de probabilité* est une famille de fonctions de probabilités qui sont structurellement de même forme, mais diffèrent l'une de l'autre par la valeur de certains paramètres.

Dans le dernier exemple, on ne connaît les probabilités que dans la mesure où l'on a dénombré les effectifs, mais dans la plupart des cas, les probabilités découlent d'une théorie concernant le contexte expérimental. Si, par exemple, on lance 5 pièces de monnaie équilibrées et on compte le nombre de FACE, il est possible de déterminer la probabilité d'avoir 0,1, … , 5 FACE. Le processus menant d'un contexte donné à une fonction de probabilité est illustré dans le tableau suivant :

[1] Les valeurs d'une variable aléatoire ne sont pas nécessairement entières. Mais l'ensemble de leurs valeurs est fini ou de cardinalité égale à celle de l'ensemble des entiers naturels.

[2] Ces fréquences sont *attendues* dans le sens que ce sont les fréquences qu'on s'attend à observer si on répète l'expérience un grand nombre de fois.

	Étape 1		Étape 2		Étape 3		Étape 4
Contexte général	Une expérience aléatoire	⇒	On définit une variable aléatoire X	⇒	On établit les valeurs x de X	⇒	On détermine les valeurs $p(x)$ de la fonction de probabilité
Exemple	On lance une pièce de monnaie 5 fois	⇒	X = le nombre de faces	⇒	Valeurs possibles de X : $x = 0, 1, 2, 3, 4, 5$.	⇒	On calcule les probabilités $p(0), p(1), p(2), p(3), p(4), p(5)$

Les calculs de effectués à partir d'une distribution empirique peuvent aussi bien être effectués à partir d'une fonction de probabilité. Le calcul qui pour une variable empirique donnent la moyenne et la variance donne, dans le cas d'une variable aléatoire X, l'*espérance mathématique* de X et la *variance* de X. Un rappel de ces notions est présenté à la fin de ce chapitre.

Nous définirons dans ce chapitre quatre des lois les plus courantes (dont les variables X, Y, Z, et W décrites ci-dessous sont des exemples) : la loi *binomiale*, la loi *hypergéométrique*, la loi *binomiale négative* et, enfin, la loi *de Poisson*. La loi *géométrique*, discutée aussi, est un cas particulier de la loi binomiale négative. Nous introduirons aussi la loi *multinomiale*, une généralisation de la loi binomiale dont nous nous servirons plus loin.

Les quatre situations suivantes peuvent servir de paradigmes :

1. *Loi binomiale* On tire au hasard un échantillon de n pièces fabriquées ; la probabilité qu'une pièce tirée soit défectueuse est p ; X est le nombre de pièces défectueuses dans l'échantillon.
2. *Loi hypergéométrique* On forme au hasard un comité de n personnes choisies parmi les N étudiants d'une classe ; Y est le nombre de filles dans le comité.
3. *Loi binomiale négative* Dans le cadre d'une étude de marketing, on doit constituer un échantillon de n familles immigrantes. Pour ce faire, on signale des numéros de téléphone au hasard jusqu'au moment où l'on obtient n familles immigrantes ; Z est le nombre d'appels qu'il faudra faire pour constituer l'échantillon.
4. *Loi de Poisson* On s'installe dans un coin de rue de 14:00 à 16:00 ; le nombre *moyen* de taxis qui passent normalement à cette intersection entre 14:00 et 16:00 est un certain nombre λ ; W est le nombre de taxis qui passeront cette fois-ci dans cet intervalle de temps.

Nous allons obtenir une fonction de probabilité distincte pour chacune des variables X, Y, Z et W.

Les variables traitées ici ne prennent que des valeurs entières et à ce titre sont appelées *lois discrètes*. Nous remettons au chapitre 3, l'étude des lois *continues*.

Notation

$p(x)$ ou $p_X(x)$	La fonction de probabilité d'une variable aléatoire X; $p(x) = P(X = x)$
$F(x)$	La fonction de répartition de X; $F(x) = P(X \le x)$.
$E(X)$ ou μ ou μ_X	L'espérance mathématique de X; $E(X) = \Sigma_x x p(x)$.
$\mathrm{Var}(X)$ ou σ^2 ou σ_X^2	La variance de X; $\mathrm{Var}(X) = \Sigma_x (x-\mu)^2 p(x)$.
$M_X(t)$	La fonction génératrice des moments de X, $M_X(t) = E(e^{tX})$.

2.1 Loi binomiale

Soi X une variable aléatoire qui représente le *nombre de succès* en une expérience aléatoire qui satis-fait les conditions suivantes :

i) Elle est composée d'une suite de n épreuves *indépendantes*.

ii) Chaque épreuve peut donner lieu à deux résultats, « succès » et « échec ».

iii) La probabilité p de succès à chaque épreuve reste fixe.

Si X est le nombre de succès obtenus au cours d'une telle expérience, alors X est de *loi binomiale de paramètres n* et p.

***Exemples* 2.1.1** *Quelques variables de loi binomiale*

1. On lance un dé 20 fois ; X est le nombre de « 6 » obtenus. Alors $X \sim \mathcal{B}(20\,;\,1/6)$.
2. On tire un échantillon de 15 pièces dans un grand lot de pièces fabriquées ; X est le nombre de pièces défectueuses dans l'échantillon. Alors $X \sim \mathcal{B}(15\,;\,p)$, où p est la proportion de pièces défectueuses *dans le lot*.
3. Dans un sondage d'opinion, on interroge 500 personnes choisies au hasard dans une population ; X est le nombre de ceux qui répondent « oui » à la question « Êtes-vous en faveur d'une loi interdisant le port d'un masque dans une manifestation ? ». Alors $X \sim \mathcal{B}(500\,;\,p)$, où p est la proportion de personnes dans la population qui voudraient qu'on interdisent le port du masque dans une manifestation.
4. On observe 25 naissances dans un hôpital ; X est le nombre de garçons parmi les nouveau-nés. Alors $X \sim \mathcal{B}(25\,;\,p)$, où p est la probabilité qu'un nouveau-né soit un garçon, soit à peu près 51 %.
5. On teste une nouvelle pilule auprès de 65 personnes souffrant de migraines ; X est le nombre de ceux qui ont trouvé la pilule efficace. Alors $X \sim \mathcal{B}(65\,;\,p)$, où p est la probabilité qu'un sujet détecte un effet (et donc, éventuellement, le pourcentage de personnes qui trouveraient la pilule efficace).

On écrit $X \sim \mathcal{B}(n;\,p)$ pour signifier « X suit une loi binomiale de paramètres n et p ».

Si $X \sim \mathcal{B}(n;\,p)$ alors la fonction de probabilité de X est donnée par

$$p(x) = \mathrm{P}(X = x) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

où

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad (2.1.1)$$

Démonstration Chaque résultat de l'espace échantillon est une succession de « S » et de « *E* » représentant les succès et les échecs, respectivement. L'événement « x succès » comprend tous les éléments de l'espace échantillon constitués de x « S » et n-x « E », comme, par exemple, celui-ci:

$$\underbrace{SS\ldots S}_{x \text{ succès}}\underbrace{EE\ldots E}_{n-x \text{ échecs}}$$

Le nombre de tels résultats est $\binom{n}{x}$, chacun de probabilité $p^x(1\text{-}p)^{n\text{-}x}$. Par conséquent, la probabilité de l'événement $\{X = x\}$ est $\binom{n}{x} \times p^x(1\text{-}p)^{n\text{-}x}$. La fonction $\binom{n}{x} p^x q^{n\text{-}x}$ est bien une fonction de probabilité puisque, par le binôme de Newton, $\sum_{x=0}^{n}\binom{n}{x} p^{x}q^{n-x} = (p+q)^n = 1$.

Théorème 2.1.1 *Si $X \sim \mathcal{B}(n;\,p)$, alors*

$$\mathrm{E}[X] = np, \quad \mathrm{Var}[X] = np(1-p) = npq$$

***Exemple* 2.1.2** *Tirages avec remise*

Supposons que 100 transactions dans une « population » de N = 100 000 transactions sont erronées.

a) Quelle est la probabilité que dans un échantillon de n = 30 transactions tirées de cette population avec remise, on trouve 2 transactions erronées ?

b) Quelle est la probabilité de trouver au plus 2 transactions erronées ?

c) Quelle est la probabilité de trouver plus de 2 transactions erronées ?

Solution Soit X le nombre de transactions erronées dans l'échantillon.

Alors $X \sim \mathcal{B}(30; 0{,}001)$.

a) $P(X=2) = \binom{30}{2}(0{,}001)^2(0{,}999)^{28} = 0{,}000423$;

b) $P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$= \binom{30}{0}(0{,}001)^0(0{,}999)^{30} + \binom{30}{1}(0{,}001)^1(0{,}999)^{29} + \binom{30}{2}(0{,}001)^2(0{,}999)^{28}$

$= 0{,}970431 + 0{,}029142 + 0{,}000423 = 0{,}999996$;

c) P[plus de 2 transactions erronées] = 1 - P[au plus 2 transactions erronées]

= 1 - 0,999996 = 0,000004.

Remarque *Si les tirages se font sans remise, le nombre X de transactions erronées (de « succès ») dans l'échantillon, ne suit pas exactement une loi binomiale, car la condition d'indépendance des épreuves n'est pas satisfaite. En effet, le résultat d'un tirage dépend des résultats obtenus aux tirages précédents. En pratique, cependant, on se permet d'utiliser quand même la loi binomiale lorsque la population est très grande comparée à l'échantillon. Car dans ces cas, les tirages successifs ne modifient pas sensiblement la population, et sont donc « presque » indépendants.*

***Exemple* 2.1.3** *Détermination de n*

Les 25 employés d'un certain bureau organisent une loterie. Ils sont numérotés de 1 à 25 et chaque semaine un numéro est tiré au hasard parmi les nombres de 1 à 25. L'employé qui porte ce numéro gagne un prix de 25 \$. Jean se demande combien de semaines la loterie doit durer pour qu'il ait au moins 70 % des chances de gagner le prix au moins une fois durant cette période.

Réponse Soit n le nombre de semaines où les employés feront cette loterie. Durant cette période, Jean peut gagner 0, 1, …, ou n fois. Nous devons déterminer la valeur de n pour laquelle la probabilité que Jean *ne gagne jamais* soit inférieure à 30 %. La probabilité qu'il ne gagne jamais est

$$\binom{n}{0}\left(\frac{1}{25}\right)^0\left(\frac{24}{25}\right)^n = \left(\frac{24}{25}\right)^n$$

Il faut donc que $\left(\frac{24}{25}\right)^n \le 0{,}30 \Leftrightarrow n\ln(24/25) \le \ln(0{,}30) \Leftrightarrow n \ge \frac{\ln(0{,}30)}{\ln(24/25)} = 29{,}49$.

La loterie devra donc durer au moins 30 semaines pour que Jean ait au moins 70 % de chances de gagner une fois ou plus.

***Exemple* 2.1.4** *Test d'hypothèse*

Un détaillant reçoit un grand lot de noix mixtes. Selon le contrat signé avec le fournisseur, le lot doit contenir 10 % (ou plus) de noix de cajou. Le détaillant inspecte un échantillon de 35 noix et n'y trouve aucune noix de cajou.

a) Est-ce assez pour conclure que le pourcentage de noix de cajou dans le lot est inférieur à 10 % ?

Réponse Soit X le nombre de noix de cajou dans un échantillon de taille 35. $X \sim \mathcal{B}(35\,;p)$ et selon le fournisseur, p = 0,10. On calculera la probabilité P(X = 0 | p = 0,10), où la barre verticale « | » signifie « étant donné que ». On trouve P(X = 0 | p = 0,10) = $\binom{35}{0}(0{,}10)^0(0{,}90)^{35}$ = 0,025. Cette probabilité étant plutôt faible, nous considérons la valeur p = 0,10 peut vraisemblable. Le détaillant aurait raison de rejeter le lot.

b) Compte tenu de la valeur observée X = 0, quelles sont les valeurs de p qu'on considérerait « vraisemblables » ?

Réponse La notion de vraisemblable est plutôt subjective. On la définira par un seuil de probabilité, c'est-à-dire, nous considérerons « vraisemblable » toute valeur de p pour laquelle P(X = 0 | p) ≥ 0,05. Donc p est « vraisemblable » si $(1-p)^{35} \geq 0{,}05 \Leftrightarrow 1-p \geq (0{,}05)^{1/35} \Leftrightarrow p \leq 1-(0{,}05)^{1/35} = 0{,}082$.

***Exemple* 2.1.5** *Estimation de p*

Soit X le nombre d'étudiants qui, dans un échantillon de 20, se déclarent croyants. X suit donc une loi $\mathcal{B}(20\,;p)$. Supposons que 5 des 20 étudiants sélectionnés se disent croyants.

a) Pourquoi la valeur p = 0,6 est-elle peu crédible?

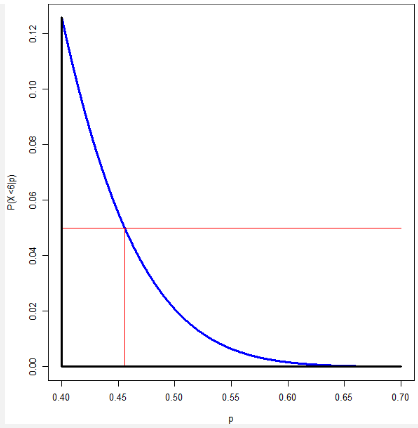
Réponse Si p = 0,6, on s'attend à trouver 12 croyants dans l'échantillon : E(X) = np = 12. Le nombre observé (x = 5), nettement inférieur à 12, peut-il être attribué au seul hasard? On calcule la probabilité d'un nombre aussi petit que 5, soit P(X ≤ 5 | p = 0,6), où la barre verticale « | » signifie « étant donné que » : P(X ≤ 5 | p = 0,6) = $\sum_{i=0}^{5}\binom{20}{i}(0{,}6)^i(0{,}4)^{20-i}$ = 0,0016. Cette probabilité étant très faible, on écarte l'hypothèse d'un accident du hasard et conclut que $p \neq 0{,}6$ (en fait que $p < 0{,}6$).

b) La valeur p = 0,4 est-elle crédible?

Réponse On calcule P(X ≤ 5 | p = 0,4) = 0,1256. Cette probabilité n'étant pas trop faible, l'hypothèse p = 0,4 n'est pas invraisemblable : l'événement observé (X = 5) peut bien s'être produit par hasard sous cette hypothèse.

c) L'approche suivie en a) et b) consiste à calculer P(X ≤ 5 | p) et exclure toute valeur de p pour laquelle cette probabilité est *trop faible*. Définissons cette notion : la probabilité P(X ≤ 5 | p) sera dite *trop faible* si elle est inférieure ou égale à 0,05—une définition arbitraire, un choix. Quelles sont les valeurs *plausibles* selon ce critère?

Réponse Le calcul ici est laborieux et ne peut se faire sans l'aide d'un logiciel. Ces calculs montrent que P(X ≤ 5 | p) < 0,05 si p ≥ 0,4555. On conclut donc que p doit se situer quelque part à gauche de 0,4555. Le graphique suivant présente la probabilité P(X ≤ 5 | p) en fonction de p :



Théorème 2.1.2 *Additivité de la loi binomiale*

Soit X_1, X_2, …, X_k des variables aléatoires indépendantes de loi binomiale de paramètres n_1, n_2, …, n_k, respectivement et p (commun à tous), alors

$$X = X_1 + X_2 + \ldots + X_k \sim \mathcal{B}(n; p),\ n = n_1 + n_2 + \ldots + n_k.$$

Cet énoncé n'exige pas de démonstration mathématique car il tombe sous le sens : X est le nombre total de succès au cours de n épreuves dont chacune a une probabilité de succès p.

Remarque

On risque d'appliquer la loi binomiale à tort si on utilise des données prélevées par d'autres sans s'enquérir de la façon dont elles ont été prélevées. Supposons, par exemple, que X est le nombre de maisons unifamiliales qui ont un garage double dans un échantillon de 500 *maisons. Si cette information provient d'une source extérieure, il se peut que X ne soit pas de loi binomiale, et ce pour plusieurs raisons.*

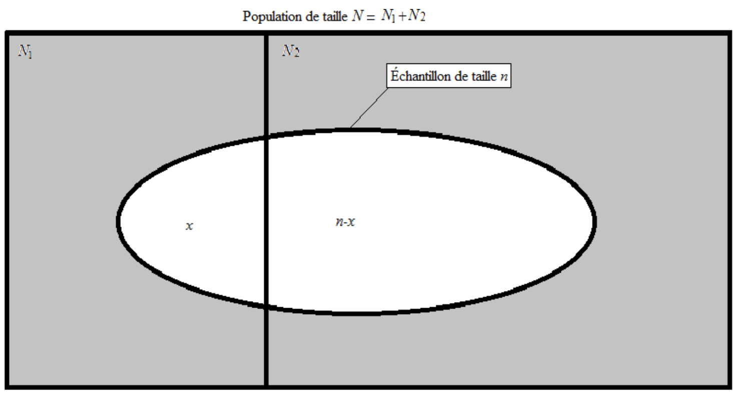
1) *L'échantillonnage a pu avoir été stratifié, c'est-à-dire, effectué de la façon suivante : On sépare la ville en* 5 *quartiers; dans chacun des quartiers, on tire au hasard* 100 *maisons. Dans ce cas, la probabilité de succès est fixe pour tous les tirages d'un même quartier, mais varient sûrement d'un quartier à un autre.*
2) *La stratification pourrait avoir d'autres conséquences qui rendrait l'emploi de la loi binomiale encore plus problématique. Rappelons que nous supposons l'indépendance des tirages en dépit du fait qu'ils sont effectués sans remise—une supposition justifiée par la présumée grande taille de la population. Mais si on avait plutôt créé* 100 *strates (et tiré 5 maisons dans chacune), ces strates pourraient trop petites pour qu'une approximation binomiale soit valable.*
3) *L'hypothèse d'indépendance entre les tirages peut être violée encore plus sérieusement. Supposons qu'on ait répertorié toutes les rues, pour ensuite tirer un échantillon de rues (et puis un échantillon de maisons dans les rues sélectionnées). Il y aurait une dépendance non négligeable entre les résultats des tirages effectués dans une même rue.*

2.2 Loi hypergéométrique

L'une des applications de la loi binomiale concerne des tirages effectués dans une population contenant deux sortes d'unités, comme, par exemple un lot d'objets manufacturés dont certains sont défectueux, d'autres pas. Le nombre d'objets défectueux (le nombre de « succès ») dans un échantillon de taille n est une variable de loi binomiale seulement si les tirages sont indépendants, c'est-à-dire, si les tirages sont effectués *avec remise*. Mais dans un sondage, les individus ne sont pas habituellement tirés avec remise — ce qui signifierait qu'un même individu risque d'être interrogé plus d'une fois. Par conséquent, le nombre de « succès » X n'est pas de loi binomiale. Précisons le contexte et fixons la notation :

- Une population de taille N est constituée de N_1 unités appartenant à une certaine classe, la classe $\mathcal{C}$, disons ; et $N_2 = N - N_1$ éléments n'appartenant pas à la classe $\mathcal{C}$;
- On tire *sans remise* un échantillon de n éléments ;
- X est le nombre d'éléments de la classe $\mathcal{C}$ qui se trouvent dans l'échantillon (on appellera X *le nombre de succès*).

Quelle est la fonction de probabilité de X ?



Faisons remarquer, pour commencer, que, tout comme avec la loi binomiale, la probabilité d'obtenir un « succès » à un tirage donné (le 1^e, le 2^e, le 3^e, …, etc.) reste toujours p, où

$$p = \frac{N_1}{N} \qquad (2.2.1)$$

quels que soient les résultats des tirages précédents. Ce qui change, c'est le fait que la probabilité *conditionnelle* de succès en un tirage dépend des résultats obtenus aux tirages précédents.

Si la population totale est très grande (pratiquement infinie), les résultats des n tirages sont *pratiquement* indépendants et dans ce cas on peut considérer X comme étant approximativement de loi $\mathcal{B}(n;\, p)$.

Si la taille de la population N n'est pas incomparablement supérieure à n, la loi binomiale est inapplicable. La loi exacte est nommée loi *hypergéométrique*. On écrit $X \sim \mathcal{H}(n;\, N_1;\, N_2)$ pour signifier que X est de loi hypergéométrique de paramètres n (le nombre de tirages), et N_1 , N_2 (les tailles des deux sous-populations, N_1 étant le nombre de «succès»).
On obtient la fonction de probabilité d'une variable de loi $\mathcal{H}(n;\, N_1, N_2)$ par un raisonnement combinatoire évident:

$$p(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}\,;\;\; 0 \le x \le N_1,\;\; 0 \le n - x \le N_2 \qquad (2.2.2)$$

Théorème 2.2.1 *Espérance et variance d'une loi hypergéométrique*

Si $X \sim \mathcal{H}(n;\, N_1;\, N_2)$, alors

$$\mathrm{E}(X) = np = n\frac{N_1}{N} \quad \text{et } \mathrm{Var}(X) = npq\,\frac{N-n}{N-1}\,,\; q = 1\text{-}p \qquad (2.2.3)$$

***Exemple* 2.2.1** *Loi hypergéométrique*

On forme au hasard un comité de 5 étudiants choisis parmi les 5 filles et les 6 garçons d'une classe. Quelle est la probabilité que 3 membres du comité soient des filles ? Faire le calcul a) exact, à l'aide de la loi hypergéométrique, et b) approximatif, par la loi binomiale.

Solution

a) La population, de taille $N = 11$, est composée de $N_1 = 5$ filles et de $N_2 = 6$ garçons. L'échantillon est de taille $n = 5$. Soit X le nombre de filles dans le comité.

Alors $\mathrm{P}(X = 3) = \dfrac{\binom{5}{3}\binom{6}{2}}{\binom{11}{5}} = 10(15)/462 = 0{,}3247$.

b) Si nous supposons que $X \sim \mathcal{B}(5;\ 5/11)$, nous avons $\mathrm{P}(X = 3) = \binom{5}{3}\left(\frac{5}{11}\right)^3\left(\frac{6}{11}\right)^2 = 0{,}2794$. La différence entre les deux résultats, 0,3247 et 0,2794, n'est pas négligeable: la loi binomiale ne fournit pas une bonne approximation de la loi hypergéométrique lorsque la taille de la population est petite.

***Exemple* 2.2.2** *Approximation d'une hypergéométrique par une binomiale*

Si $X \sim \mathcal{H}(n;\ N_1;\ N_2)$, on peut approcher les probabilités $\mathrm{P}(X = x)$ par la fonction de probabilité d'une variable X de loi $\mathcal{B}(n;\ p)$, où $p = N_1/N$, à condition que N soit grand comparé à n. Soit $n = 5$ et $p = 5/12$. Calculer $\mathrm{P}[X = 3]$ lorsque a) $N = 12$; b) $N = 24$; c) $N = 48$, d) $N = 72$, et comparer ces probabilités à celle calculée au moyen d'une $\mathcal{B}(5;\ 5/12)$.

Solution

a) $\binom{5}{3}\binom{7}{2}\Big/\binom{12}{5} = 0{,}26515$; b) $\binom{10}{3}\binom{14}{2}\Big/\binom{24}{10} = 0{,}25692$;

c) $\binom{20}{3}\binom{28}{2}\Big/\binom{48}{5} = 0{,}25166$; d) $\binom{30}{3}\binom{42}{2}\Big/\binom{72}{5} = 0{,}24984$.

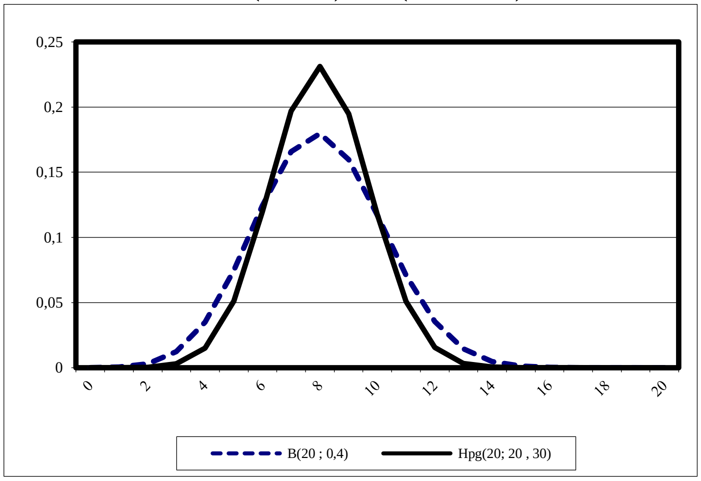
Si nous supposons que $X \sim \mathcal{B}(5;\ 5/12)$, on a $\mathrm{P}[X = 3] = \binom{5}{3}(5/12)^3(7/12)^2 = 0{,}24615$. Le tableau suivant résume les résultats:

Loi	$\mathrm{P}(X = 3)$
$\mathcal{H}(5;\ 5,\ 7)$	0,26515
$\mathcal{H}(5;\ 10,\ 14)$	0,25692
$\mathcal{H}(5;\ 20,\ 28)$	0,25166
$\mathcal{H}(5;\ 30,\ 42)$	0,24984
$\mathcal{B}(5;\ 5/12)$	0,24615

Dans tous les cas ci-dessus, $n = 5$ et $p = 5/12$. Nous voyons que les probabilités hypergéométriques s'approchent de plus en plus de la probabilité binomiale à mesure que N_1 et N_2 augmentent, de telle sorte que le rapport $N_1/N = 5/12$ reste constant.

Figure 2.2.1

Lois $\mathcal{B}(20; 0{,}4)$ *et* $\mathcal{H}(20; 20, 30)$



***Exemple* 2.2.3** *Test d'hypothèse*

Un bloc d'habitation, composé de 20 logements, est mis en vente. Le propriétaire prétend que dans 4 de ces logements — pas plus — il y a un animal domestique. Un client éventuel, voulant vérifier cette affirmation, fait inspecter 5 logements tirés au hasard. Il trouve un animal domestique dans 3 d'entre eux. Le propriétaire a-t-il (vraisemblablement) menti ?

Solution Le nombre X de logements avec un animal domestique, parmi les 5 logements visités, suit une loi $\mathcal{H}(5\,;N_1\,,N_2)$. Si le propriétaire a dit vrai, $N_1 = 4$ (et $N_2 = 16$) et la valeur attendue de X sous l'hypothèse que $N_1 = 4$ est $E(X) = np = 5(4/20) = 1$. La valeur observée $X = 3$, trois fois supérieure à la valeur attendue, semble excessive. Nous voulons savoir si une valeur *aussi élevée* lorsque $N_1 = 4$ est probable. On calculera donc $P(X \geq 3 \mid N_1 = 4)$, où la barre verticale « | » signifie « étant donné que ». La formule nous donne $P(X \geq 3 \mid N_1 = 4) = P(X = 3 \mid N_1 = 4) + P(X = 4 \mid N_1 = 4) = 480/15504 + 16/15504 = 0{,}0320$. Si on admet que la probabilité 0,0320 est « petite », on conclut, compte tenu de l'observation $X = 3$, que l'hypothèse $N_1 = 4$ est *peu vraisemblable* : il serait trop peu probable d'avoir autant de logements abritant un animal domestique si N_1 n'était que de 4. On dira que la valeur observée est *peu conforme* à l'hypothèse que $N_1 = 4$.

2.3 Loi géométrique

Les variables binomiale et hypergéométrique ont ceci en commun qu'elles représentent le nombre de succès en un nombre *fixe* d'épreuves. Il arrive, cependant, qu'on ne veuille pas fixer le nombre d'épreuves au départ, comme dans le prochain exemple.

***Exemple* 2.3.1** *Nombre d'essais aléatoire*

Une des applications des lois binomiale et hypergéométrique survient lorsqu'on cherche à estimer une proportion p à partir d'un échantillon aléatoire. L'exemple 4.1.2 illustre une telle application. On tire un échantillon de 15 pièces d'un grand lot afin d'estimer la proportion de pièces défectueuses dans le lot. On trouve 4 défectueuses dans l'échantillon. On estime alors que la proportion p de pièces défectueuses dans le lot et de $4/15 \approx 26{,}7$ %. Cette façon classique de procéder peut poser un problème lorsque p est très petit, comme, par exemple, lorsque p est le pourcentage de personnes souffrant d'une maladie très rare, ou la proportion de transactions comptables erronées ou frauduleuses. Ce qui peut arriver alors, c'est qu'un échantillon de taille fixée d'avance ne comprenne *aucun* succès. L'estimation qui en découlerait (p = 0 %) n'aurait pas de sens (le fait que toutes les unités sont conformes dans l'échantillon ne signifie pas qu'elles le sont tours dans la population.) En prévision de cette éventualité, il convient de ne pas fixer d'avance le nombre de tirages mais de se préparer plutôt à poursuivre les tirages jusqu'au moment où l'on obtient un succès. Le nombre d'essais nécessaire pour réaliser ce premier succès est un indice de la valeur de p. Si le nombre d'essais est élevé, c'est que p est petit ; s'il est faible, c'est que p est grand.

Considérons donc une expérience composée d'épreuves indépendantes, avec probabilité de succès p à chaque épreuve. Supposons qu'on continue à effectuer des essais jusqu'à ce qu'on réalise un premier succès. Le nombre de succès ici n'est donc pas aléatoire, il est toujours égal à 1; ce qui est aléatoire, et qu'on désigne par X, c'est le *nombre d'essais* nécessaire pour obtenir un premier succès. On dit alors que X est de *loi géométrique* de paramètre p et on écrit : $X \sim \mathcal{G}(p)$.
Il vient aisément que

$$p(x) = pq^{x-1} \text{ pour } x = 1, 2, 3, \ldots \qquad (2.3.1)$$

En effet si le premier succès apparaît au x^e essai, les $x - 1$ premiers essais doivent avoir conduit à des échecs, d'où le résultat

$$p(x) = \mathrm{P}(\underbrace{\mathrm{EE...E}}_{x-1 \text{ echecs}}\mathrm{S}) = q^{x-1}p$$

Théorème 2.3.1

Si $X \sim \mathcal{G}(p)$, alors

$$\mathrm{E}(X) = \frac{1}{p}\,;\ \mathrm{Var}(X) = \frac{q}{p^2}\,;\ \mathrm{P}\{X > x\} = q^x \qquad (2.3.2)$$

Le troisième énoncé est évident : l'événement $\{X > x\}$ se produit si et seulement si les x premiers essais sont des échecs. La probabilité de cet événement est évidemment q^x.

***Exemple* 2.3.1** *Loi géométrique*

Perdu dans une ville étrangère, vous avez l'intention de demander à un passant comment vous rendre au musée. On suppose qu'une personne sur 5 saurait vous guider. Si on désigne par X le nombre de personnes que vous allez devoir approcher pour enfin obtenir une réponse, déterminer $p(x)$ pour x = 1, ... , 6, et calculer $\mathrm{P}(X \le 6)$. Déterminer $\mathrm{E}(X)$ et $\mathrm{Var}(X)$.

Solution $X \sim \mathcal{G}(1/5)$; $p(x) = (1/5)(4/5)^{x-1} = (0{,}2)(0{,}8)^{x-1}$.

$p(1) = 0{,}2$; $p(2) = 0{,}16$; $p(3) = 0{,}128$, $p(4) = 0{,}1024$, $p(5)= 0{,}08192$, $p(6) = 0{,}065536$.

$\mathrm{P}(X \le 6) = 0{,}2 + 0{,}16 + 0{,}128 + 0{,}1024 + 0{,}08192 + 0{,}065536$

$= 0{,}737856$

$\mathrm{E}(X) = 1/p = 5$; $\mathrm{Var}(X) = q/p^2 = 20$.

Tout comme avec la loi binomiale, la loi géométrique repose sur des hypothèses qui risquent fort de ne pas être vérifiées. En voici un exemple:

***Exemple* 2.3.2** *Application douteuse de la loi géométrique*

Une compagnie pétrolière effectue des forages. On suppose que chaque puits creusé a une chance sur 5 de donner du pétrole. Supposons que X, le nombre de puits qui doivent être creusés pour trouver du pétrole, est de loi géométrique. Déterminer $p(x)$ pour $x = 1, ..., 6$, et calculer $P(X > 6)$. Déterminer $E(X)$ et $Var(X)$.

Solution $X \sim \mathcal{G}(1/5)$; $p(x) = (1/5)(4/5)^{x-1} = (0,2)(0,8)^{x-1}$.
$p(1) = 0,2$, $p(2) = 0,16$; $p(3) = 0,128$; $p(4) = 0,1024$; $p(5)= 0,08192$, $p(6) = 0,065536$.
$P(X > 6) = q^6 = (0,8)^6 = 0,262144$. $E(X) = 1/p = 5$; $Var(X) = q/p^2 = 20$.

Remarque *Pour que le modèle géométrique puisse s'appliquer au dernier exemple, il faudrait que les lieux de forage soient suffisamment éloignés les uns des autres pour qu'on puisse raisonnablement supposer les résultats indépendants. Si tous les forages sont effectués au même site, l'indépendance de leurs résultats est douteuse et la loi géométrique ne s'applique plus. Une autre hypothèse difficile à justifier : peut-on vraiment supposer que la probabilité de succès est la même à chaque essai ?*

Une particularité commode de la loi géométrique: il existe une expression très simple de la fonction de répartition $F(x)$. La voici :

Fonction de répartition d'une variable $X \sim \mathcal{G}(p)$:

$$F(x) = P(X \le x) = 1 - q^x. \tag{2.3.3}$$

La preuve est simple: $F(x) = P(X \le x) = 1 - P(X > x)$. Or $P(X > x) = q^x$ car l'événement $\{X > x\}$ se produit si et seulement si les x premier essais ne donnent aucun succès.

Donc $F(x) = 1 - P(X > x) = 1 - q^x$.

***Exemple* 2.3.3** *Fonction de répartition*

Dans le contexte de l'exemple 2.3.1, calculez la probabilité d'avoir une réponse avant la 4e tentative.

Solution On cherche $P(X < 4) = P(X \le 3) = 1 - P(X > 3) = 1 - (0,2)^3 = 0,992$.

2.4 Loi binomiale négative

La loi géométrique est un cas particulier d'une loi plus générale, la loi *binomiale négative*. Le contexte est le même : des épreuves indépendantes, probabilité de succès p, nombre d'épreuves aléatoire. Ce qui change, c'est la règle d'arrêt : on s'arrête dès qu'on obtient n succès, où n est un entier positif fixé d'avance. La variable aléatoire X est le nombre d'essais nécessaire pour obtenir le n^e succès. On dit alors que X est de loi *binomiale négative* et on écrit $X \sim \mathcal{B}^-(n ; p)$.

Remarque *Si* $n = 1$, *la loi binomiale négative est en fait la loi géométrique.*

***Exemple* 2.4.1** *Nombre d'essais aléatoire*

Reprenons la discussion de l'exemple 2.3.1. On sait que le nombre d'essais nécessaire pour avoir un succès donne une idée de la proportion p qu'on tente d'estimer: Si le premier succès arrive vite (X petit), on a une bonne raison de croire que p est grand. Mais cette déduction est risquée, car il n'est pas improbable que le premier succès arrive assez vite même si p est petit. On peut se prémunir contre ce risque en poursuivant les tirages après le premier, jusqu'au 2e tirage. Ou jusqu'au 3e, 4e, …, ou n^e tirage.

Les valeurs de X sont $\{n, n+1, … \}$. Pour tout x dans cet ensemble, l'événement $\{X = x\}$ se produit si les x-1 premiers essais donnent n-1 succès et le x^e donne un succès. Donc

P($X = x$) = P(les x-1 premiers essais donnent n-1 succès)×P(le x^e essai donne un succès)

$$= \binom{x-1}{n-1} p^{n-1} q^{(x-1)-(n-1)} \times p$$

Nous avons donc le résultat suivant :

La fonction de probabilité d'une variable X de loi $\mathcal{B}^{-}(n;\ p)$ est

$$p(x) = \binom{x-1}{n-1} p^n q^{x-n},\ x = n,\ n+1,\ \ldots \qquad (2.4.1)$$

Théorème 2.4.1 *Espérance et variance d'une loi binomiale négative*

Si X : $\mathcal{B}^{-}(n;\ p)$, alors

$$\mathrm{E}(X) = \frac{n}{p}\ ;\ \mathrm{Var}(X) = \frac{nq}{p^2} \qquad (2.4.2)$$

***Exemple* 2.4.1** *Loi binomiale négative*

On vérifie, l'une après l'autre, les transactions d'un fichier contenant des dizaines de milliers de transactions. Supposons que 1% des comptes du fichier contiennent des erreurs.

a) En moyenne, combien devra-t-on vérifier de transactions pour en trouver 3 qui contiennent des erreurs ?

avec des erreurs ?

c) Quelle est la probabilité qu'il faille en prélever plus de 300 ?

Solution X = nombre de transactions vérifiées au moment où la 3^{e} erreur est détectée.

a) E(X) = n/p = 3/0,01 = 300.

b) P(X = 300) = $\binom{299}{2}(0,01)^3(0,99)^{297}$ = 0,00225

c) Nous devons calculer la probabilité P(X > 300), ce qui, à première vue peut sembler une tâche colossale. P(X > 300) = $\sum_{x=301}^{\infty} p(x)$ = $1-\sum_{x=3}^{300} p(x)$. Mais en fait cette probabilité peut être calculée en sommant quelques termes d'une fonction de probabilité binomiale. L'événement {X > 300} est équivalent à {les 300 premiers essais donnent moins de 3 succès}. Si Y est le nombre de succès au cours des 300 premiers essais, Y est de loi $\mathcal{B}(300\ ;\ 0{,}01)$, la probabilité de cet événement est P(Y < 3) = P(Y = 0) + P(Y = 1) = $\binom{300}{0}(0,01)^0(0,99)^{300}$ + $\binom{300}{1}(0,01)(0,99)^{299}$ + $\binom{300}{2}(0,01)^2(0,99)^{298}$ = 0,049 + 0,149 + 0,224 = 0,422.

Théorème 2.4.2 *Additivité de la loi binomiale négative*

Soit X_1, X_2,…, X_k sont des variables aléatoires indépendantes de loi binomiale négative de paramètres n_1, n_2, … n_k, respectivement, et p (commun à tous), alors la

$$X = X_1+X_2+\ldots+X_k \sim \mathcal{B}^{-}(n;\ p)$$

où $n = n_1+n_2+\ldots+ n_k$ *et p*.

Cet énoncé tombe sous le sens : X est le nombre d'essais nécessaire pour obtenir n succès lorsque chaque épreuve a une probabilité de succès p.

Cas particulier : une somme de k variables géométriques indépendantes de même paramètre p est de loi binomiale négative de paramètre k et p.

2.5 Loi de Poisson

La loi de Poisson, comme la loi binomiale, représente le nombre de fois où un certain événement se produit. Mais l'expérience ne consiste pas à répéter une même épreuve plusieurs fois ; elle consiste plutôt à compter le nombre de fois où un certain événement se produit dans un intervalle de temps; ou le nombre de points sur une surface ou dans une région de l'espace. La variable pourrait représenter le nombre d'appels reçus par une réceptionniste entre 14 heures et 16 heures; le nombre d'arrivées à un poste de péage dans l'espace d'une heure; le nombre d'émissions de certaines particules sous-atomiques en une seconde. Elle peut aussi représenter le nombre de points dans un espace, comme par exemple le nombre de petits défauts sur une plaque d'émail. Finalement, la loi de Poisson peut servir à approcher la distribution d'une variable de loi binomiale, comme, par exemple, le nombre de pages sans faute de frappe dans un livre.

Les conditions sous lesquelles la loi de Poisson s'applique sont plus délicates à définir. Nous remettons donc cette question à plus tard et commençons par une définition formelle :

Définition *Loi de Poisson*

Une variable aléatoire est dite de loi de Poisson *de paramètre* λ [$X \sim \mathcal{P}(\lambda)$] *si sa fonction de probabilité est donnée par*

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, ..., (\lambda > 0) \qquad (2.5.1)$$

Remarquez que le nombre de valeurs est infini. Il sera utile de savoir, au moment d'appliquer cette loi, que le paramètre λ est la moyenne de X. On note également le fait particulier que la variance est égale à la moyenne:

Théorème 2.5.1 *Espérance et variance d'une loi de Poisson*

Si $X \sim \mathcal{P}(\lambda)$, alors

$$E(X) = \lambda;\ \mathrm{Var}(X) = \lambda \qquad (2.5.2)$$

***Exemple* 2.5.1** *Nombre de décès*

Admettons que X, le nombre de décès à Laval au cours des prochaines 24 heures, suit une loi de Poisson. Calculer

a) $P(X = 4)$ b) $P(X \leq 3)$

Solution

a) $P(X = 4) = \frac{e^{-\lambda}\lambda^4}{4!}$. Mais que vaut λ ? Sachant que $\lambda = E(X)$ est le nombre moyen de décès en un jour à Laval, nous pouvons estimer λ à partir de données statistiques. Nous savons qu'il y a eu 2490 décès à Laval en 2007, ce qui fait environ 6,821918 par jour. C'est une estimation raisonnable de λ. On prendra donc $\lambda = 6{,}821918$ et alors $P(X = 4) = \frac{e^{-6,82}(6{,}821918)^4}{4!} = 0{,}0984$.

b) $P(X \leq 3)$

$$= \frac{e^{-6,821918}(6{,}821918)^0}{0!} + \frac{e^{-6,821918}(6{,}821918)^1}{1!} + \frac{e^{-6,821918}(6{,}821918)^2}{2!} + \frac{e^{-6,821918}(6{,}821918)^3}{3!}$$

$= 0{,}001089629+0{,}007433361+0{,}025354889+0{,}057656322 = 0{,}0915342.$

Lien avec la loi binomiale

On établit ici un lien entre la loi de Poisson et la loi binomiale, un lien qui montre que la loi de Poisson peut servir d'approximation à la loi binomiale. Ce lien servira à justifier la fonction de probabilité.

Considérons une suite de variables aléatoires $X_n \sim \mathcal{B}(n; p_n)$ telle que lorsque $n \to \infty$, $p_n \to 0$ de telle sorte que le produit np_n converge vers une constante positive λ. Alors la loi limite de X_n est une loi de Poisson de paramètre λ. Voici l'énoncé formel :

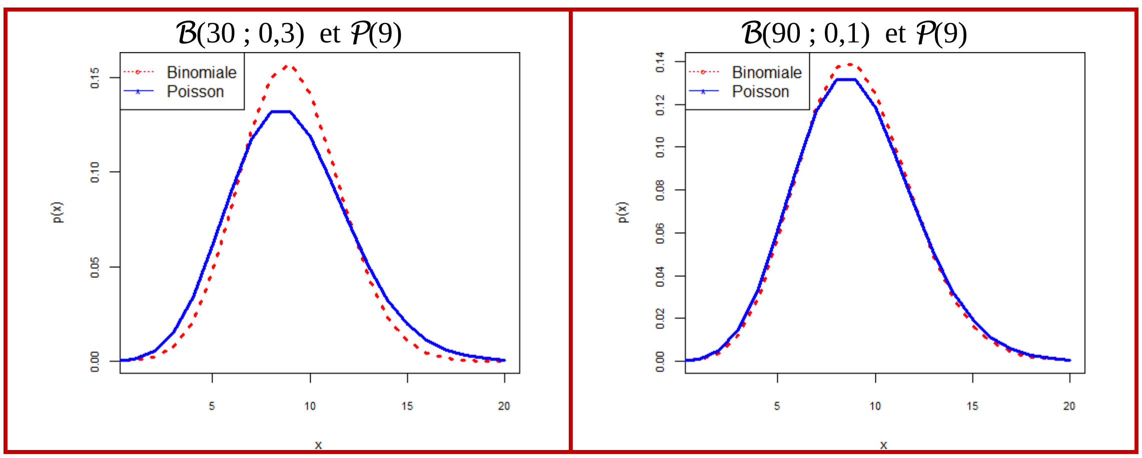
Théorème 2.5.2 *La loi binomiale tend vers une loi de Poisson*

Soit $X_1, X_2, \ldots, X_n, \ldots$une suite de variables aléatoires, $X_n \sim \mathcal{B}(n; p_n)$, où p_n *est une suite* telle que $\lim\limits_{n\to\infty} np_n = \lambda$. Alors

X_n tend en loi vers une loi $\mathcal{P}(\lambda)$.

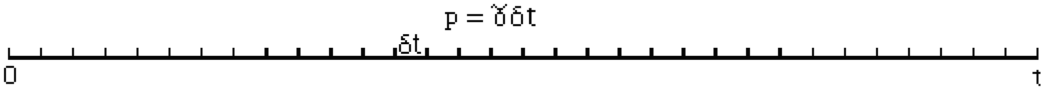
Précisons la portée du théorème : pour un x fixe, on sait que $P(X_n = x) = \binom{n}{x} p_n^{\,x}(1-p_n)^{n-x}$. Ce qu'affirme le théorème, c'est que lorsque $n \to \infty$, et $p_n \to 0$ de telle sorte que $np_n \to \lambda$, alors $\binom{n}{x} p_n^{\,x}(1-p_n)^{n-x} \to \dfrac{e^{-\lambda}\lambda^x}{x!}$. Concrètement, cela signifie que lorsque n est grand et p petit, $\binom{n}{x} p_n^{\,x}(1-p_n)^{n-x} \approx \dfrac{e^{-np}(np)^x}{x!}$: la distribution $\mathcal{B}(n;\ p)$ peut être approchée par une loi $\mathcal{P}(\lambda)$, avec $\lambda = np$: $P(X = x) \approx \dfrac{e^{-np}(np)^x}{x!}$.

En principe, l'approximation est bonne lorsque n est grand et p petit. Mais, comme le montre le graphique suivant, elle peut être acceptable même lorsque ces conditions sont relâchées :



Dans quelles conditions la loi de Poisson s'applique-t-elle ? La loi de Poisson ne sert pas qu'à approcher la loi binomiale, loin de là : Plusieurs phénomènes dans la nature peuvent être modélisés par cette loi. Le problème, c'est qu'il n'est pas toujours aisé de les reconnaître en pratique. Le théorème 2.5.2 peut aider à le faire. Typiquement, une variable X de loi de Poisson représente le nombre de fois où un certain « événement » se produit dans un intervalle de temps donné $[0; t]$. L'événement en question pourrait être, par exemple, un accident d'auto, et X le *nombre* d'accidents au courant d'une journée donnée à Montréal. Supposons qu'on découpe l'intervalle $[0\,;\,t]$ en un grand nombre n de sous-intervalles de longueur δt $(= t/n)$, et admettons qu'on puisse faire les suppositions suivantes concernant le phénomène des accidents d'auto à Montréal :

1. a) La probabilité p qu'un accident ait lieu exactement une fois dans un sous-intervalle de durée δt est à peu près proportionnelle à δt lorsque δt est petit: $p \approx \gamma\delta t$, où γ est un paramètre positif qui mesure l'intensité ou la densité du phénomène, le taux d'accidents.



   b) La probabilité de plus d'un accident dans un sous-intervalle de durée δt, est négligeable lorsque δt est petit.

2. Les événements A_i = {un accident a lieu dans l'intervalle i}, $i = 1, 2, \ldots, n$, sont mutuellement indépendants.

Compte tenu de ces hypothèses, et considérant chaque sous-intervalle comme une épreuve, on voit que X est à peu près de loi $\mathcal{B}(n;\, p_n)$, où $p_n = \gamma\delta t = \gamma\frac{t}{n}$. Les hypothèses sont d'autant plus crédibles que n est grand. Alors on fait tendre n vers l'infini (et donc $p_n = \gamma\frac{t}{n}$ vers 0), alors que $np_n = n\gamma\frac{t}{n} \to \gamma t$ (en fait ici $n\gamma\frac{t}{n} = \gamma t$). Ce sont précisément les conditions dans lesquelles la loi binomiale tend vers une loi de Poisson de paramètre $\lambda = \gamma t$.

Remarque *Interprétation de* λ.

- *Le paramètre* $\lambda = \gamma t$ *est le nombre moyen de réalisations dans l'intervalle de temps considéré, alors que* γ *en est la* fréquence, *c'est-à-dire, le nombre moyen de réalisations* par unité de temps. *Par exemple, si le taux de naissance au Canada est de* 43 *par heure, alors* $\gamma = 43$, *le taux unitaire. Si X représente le nombre de naissances au courant des prochaines* 5 *heures, alors X suit une loi de Poisson de paramètre* $\lambda = 5(43) = 215$.
- *Les hypothèses décrites ci-dessus peuvent être exprimées en termes d'espace lorsqu'il s'agit de la distribution de points sur une surface : l'intervalle* $[0\,;\, t]$ *peut être remplacé par une surface S et les sous-intervalles par des petits carrés.*

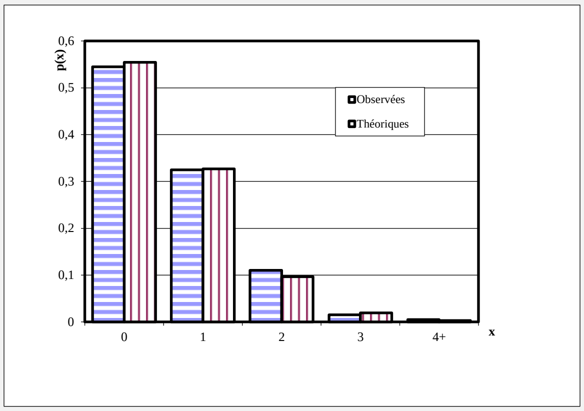
Remarque *La loi de Poisson diffère des lois binomiale et hypergéométrique, géométrique et binomiale négative en ce que les conditions qui justifient son utilisation sont propres au phénomène observé et échappent au contrôle de l'observateur. Lorsqu'on utilise une loi binomiale, c'est qu'on a effectué nos tirages (ou autres épreuves) de telle sorte qu'ils soient indépendants — en tirant avec remise, par exemple; et de telle sorte que la probabilité de succès demeure fixe — en effectuant nos tirages toujours dans la même population. On n'a pas cette assurance avec la loi de Poisson : l'indépendance supposée est-elle vraiment réaliste ? La probabilité de succès dans un sous-intervalle est-elle partout la même ? Le fait : on ne le sait pas toujours, car cela dépend du phénomène étudié et non de la méthodologie de l'expérimentateur.*

***Exemple* 2.5.3** *Application classique* [von Bortkiewicz, L. (1898). *Das Gesetz der Kleinen Zahlen*. Leipzig: Teubner. Cité dans le site de Michael Friendly, http://www.math.yorku.ca/SCS/friendly.html]

L'une des premières applications de la loi de Poisson porte sur le nombre X de décès dus à des coups de sabots de cheval ou de mules dans un corps de l'armée prussienne pendant un an. On disposait de données sur 10 corps militaires sur une période de 20 ans. On avait donc 200 réalisations de la variable X. La première colonne du tableau suivant donne les valeurs observées de X; la deuxième donne l'effectif observé (par exemple, il y a eu 109 jours-corps sans décès; 65 avec un décès, etc.); et la troisième donne la fréquence observée. Si la loi de Poisson est un modèle adéquat, les fréquences des valeurs i = 0, 1, 2, et 3 devraient être proches des probabilités $p(0)$, $p(1)$, $p(2)$, $p(3)$, respectivement, et la fréquence de 4+ devrait être proche de $P(X \geq 4)$, les probabilités étant calculées à l'aide de la loi de Poisson. Ces probabilités figurent dans la quatrième colonne et ont été calculées en prenant pour λ la moyenne des 200 observations, soit λ = 0,59. C'est une estimation raisonnable de λ, puisque λ est l'espérance de X.

x	Effectifs observés	Fréquences observées	Fréquences théoriques	
0	109	0,545	$e^{-0,59}(0,59)^0/0!$	0,5543
1	65	0,325	$e^{-0,59}(0,59)^1/1!$	0,3271
2	22	0,11	$e^{-0,59}(0,59)^2/2!$	0,0965
3	3	0,015	$e^{-0,59}(0,59)^3/3!$	0,0190
4+	1	0,005	$1-\sum_{i=0}^{3} e^{-0,59}(0,59)^i/i!$	0,0032
	200	1	1	1

Graphiquement, la concordance entre les fréquences observées et les fréquences théoriques est frappante :



On a donc de bonnes raisons de croire que X suit, du moins approximativement, une loi de Poisson.

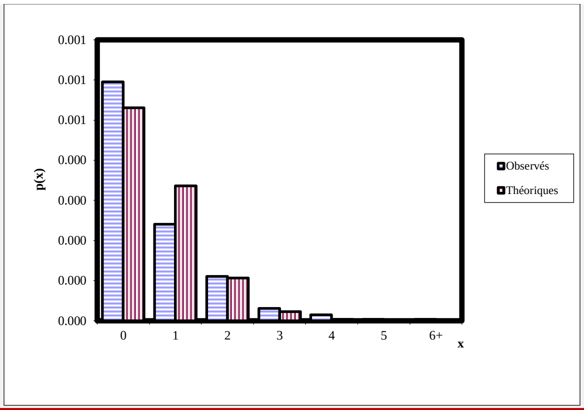
La situation est rarement aussi claire. Dans le prochain exemple, une comparaison des fréquences observées avec les fréquences théoriques révèle des différences plus importantes que dans le dernier exemple.

***Exemple* 2.5.4** *Un ajustement douteux : les articles de Madison* [Mosteller, F., and Wallace, D. L. (1984). *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. New York: Springer-Verlag]

Les données suivantes ont été recueillies dans le but d'établir la paternité de certains textes, mais c'est à d'autres fins que nous les reproduisons ici. Les 262 valeurs présentées dans le tableau ci-dessous représentent le nombre d'occurrences du mot (anglais) « *may* » dans des blocs de textes d'environ 200 mots. Est-ce que la variable X (nombre de d'occurrence de « *may* » en un bloc de 200 mots ») est de loi de Poisson ? Les calculs effectués dans la dernière colonne du tableau sont semblables à ceux effectués dans le dernier exemple, sauf que la moyenne ici est 0,6336, et c'est ce que nous prenons pour λ.

x	Observés	Observés	Théoriques
0	156	0,595	0,531
1	63	0,240	0,336
2	29	0,111	0,107
3	8	0,031	0,022
4	4	0,015	0,004
5	1	0,004	0,000
6+	1	0,004	0,000
	262	1	1

Comme le montre le graphique ci-dessous, l'ajustement ici est nettement moins bon et on pourrait douter de la loi de Poisson comme modèle pour ce phénomène :



Remarque *Le dernier exemple, contrairement au précédent, est un cas où la conclusion n'est pas évidente visuellement : les écarts entre les fréquences théoriques et les probabilités sont plus importants, certes, mais le sont-ils trop ? Est-ce qu'on ne pourrait pas quand même accepter l'hypothèse d'une loi de Poisson en se disant que le hasard seul suffit à expliquer les écarts ? L'intuition n'est pas un bon guide dans des cas comme celui-ci, sans extrêmes : Visuellement, l'ajustement n'est ni très bon, comme dans l'exemple* 2.5.3, *ni très mauvais. Ces considérations sont importantes, et nous allons développer plus loin des techniques permettant de les traiter de façon plus rigoureuse. On pourra, en fait, montrer formellement que l'hypothèse d'une loi de Poisson dans ce cas-ci est très peu vraisemblable. Il aurait été difficile, d'ailleurs, de la justifier a priori. Il aurait fallu supposer que chaque mot d'un bloc de* 200 *mots peut être* « may » *ou pas, toujours avec probabilités fixes p et* (1-*p*)*. Et il aurait fallu accepter une hypothèse d'indépendance qui n'est manifestement pas vraie : le mot* « may » *ne peut pas paraître deux fois de suite, par exemple*.

Additivité de la loi de Poisson

Il arrive qu'on ait affaire à plus d'une variable aléatoire suivant une loi de Poisson, et qu'on s'intéresse à leur somme. Si ces variables sont indépendantes, leur somme est aussi de loi de Poisson, avec pour paramètre la somme des paramètres. Voici l'énoncé précis de cette propriété :

Théorème 2.5.3 Si $X_i \sim \mathcal{P}(\lambda_i)$, $i = 1, 2, \ldots, n$, sont des variables aléatoires indépendantes, alors

$$X = \sum_{i=1}^{n} X_i \ \sim \mathcal{P}\left(\sum_{i=1}^{n} \lambda_i\right).$$

***Exemple* 2.5.5** *Somme de variables de loi de Poisson*

Supposons que le nombre de verres de boisson demandés par un voyageur aérien de première classe est de loi $\mathcal{P}(1{,}2)$. Quelle est la probabilité que les 6 voyageurs de première classe en un voyage donné demandent 10 verres ou moins ? Y a-t-il une raison pour laquelle la loi de Poisson pourrait ne pas être applicable.

Solution Si X_1, X_2,... , X_6 sont les nombres de verres demandés par les 6 voyageurs, respectivement, alors la consommation totale $X = X_1 + ... + X_6$ suit une loi $\mathcal{P}(7{,}2)$. Donc $P(X \leq 10) = 0{,}8867$. Ce calcul exige que la variable X suive une loi de Poisson, ce qui suppose que les variables X_1, ... , X_6 sont indépendantes. Le sont-elles ? Si les six voyageurs sont en fait trois couples, il y aura des dépendance qui invalident l'application du théorème.

2.6 Loi multinomiale

La loi multinomiale est une généralisation de la loi binomiale. Nous avons dit que dans une expérience binomiale, chaque épreuve n'a que deux résultats possibles, {*Succès* ; *Échec*} : un bébé peut être une fille ou un garçon ; une pièce fabriquée peut être défectueuse ou conforme. Cette dualité n'est pas nécessairement intrinsèque au résultat. Elle peut être le résultat d'une classification en deux catégories d'une multiplicité de résultats. Une question concernant les droits de scolarité à l'université peut offrir au répondant d'un sondage le choix suivant de réponses :

Gratuité	Gel	Indexation au coût de la vie	Majoration supérieure à l'augmentation du coût de la vie

La loi binomiale n'est pas nécessairement exclue, car on peut être disposé à grouper ces réponses en deux catégories, comme, par exemple, ceux qui ne tolèrent aucune augmentation (les deux premières réponses) et les autres.

Mais cette condensation entraîne la perte d'une information qu'on tient parfois à conserver; et les nouvelles catégories peuvent être dénuées de sens. Par exemple, dans le cadre d'une enquête sur la promotion de la santé faite par Statistique Canada, on demande aux répondants d'indiquer leur activité principale au cours des douze derniers mois. Les réponses possibles sont : {Travailleur ; À la recherche d'un emploi ; Étudiant ; Retraité ; Ménagère ; Autre}. Il y a donc 6 résultats possibles et on souhaite préserver l'identité de chacune. On peut désigner par p_1 la probabilité qu'un individu réponde qu'il est « travailleur » ; par p_2 la probabilité qu'il soit à la recherche d'un emploi ; … ; et par p_6 la probabilité qu'il ait une activité autre que ces dernières. On suppose qu'un individu donné ne peut choisir qu'une seule occupation. Si on interroge n personnes, on peut résumer l'information relevée sous la forme d'une suite de 6 nombres, $(x_1 ; x_2 ; … ; x_6)$, où x_1 est le nombre de personnes qui travaillent, x_2 est le nombre de personnes qui cherchent un emploi, etc. Résumons :

Résultat (i)	Probabilité du résultat (i) en un tirage	Nombre de fois où le résultat (i) est observé en n tirages
Travailleur	p_1	X_1
À la recherche d'un emploi	p_2	X_2
Étudiant	p_3	X_3
Retraité	p_4	X_4
Ménagère	p_5	X_5
Autre	p_6	X_6
Total	1	n

Cet exemple illustre une situation que nous décrivons maintenant formellement en termes généraux. Une expérience aléatoire est composée de n épreuves indépendantes, où chaque épreuve peut donner lieu à k résultats possibles, $\{R_1 ; R_2 ; … ; R_k\}$, le résultat R_i avec probabilité p_i, $i = 1, ... , k$. Considérons les k variables aléatoires $X_1, ... , X_k$, où X_i est le nombre de fois (parmi n) où le résultats R_i se produit ($i = 1, ... , k$). Ensemble, ces k variables constituent ce qu'on appelle un *vecteur aléatoire* $\boldsymbol{X}$: $\boldsymbol{X} = [X_1 ; ... ; X_k]$.

Remarque *Les probabilités p_1, p_2, … , p_k et les variables X_1 , X_2 , … , X_k satisfont les conditions*

$$p_1 + p_2 + \ldots + p_k = 1 \text{ et } X_1 + X_2 + \ldots + X_k = n.$$

La fonction de probabilité d'un *vecteur* aléatoire $\boldsymbol{X} = [X_1 ; \ldots ; X_k]$ fait correspondre une probabilité à chaque vecteur $\boldsymbol{x} = [x_1 ; \ldots ; x_k]$ tel que $\Sigma x_i = n$. Elle est définie par

$$p(x_1 ; \ldots ; x_k) = \mathrm{P}(X_1 = x_1 \text{ et } X_2 = x_2, \ldots , \text{ et } X_k = x_k).$$

La fonction de probabilité de la loi multinomiale fait intervenir le coefficient multinomial défini de la façon suivante :

$$\binom{n}{x_1;\ldots;x_k} = \frac{n!}{x_1!\ldots x_k!} \tag{2.6.1}$$

Notons que

$$\binom{n}{x;\, n-x} = \frac{n!}{x!(n-x)!} = \binom{n}{x}$$

de sorte que le coefficient binomial n'est autre qu'un coefficient multinomial avec $k = 2$.

Définition *Considérons une expérience constituée de n épreuves indépendantes, chacune donnant lieu à l'un de k résultats R_1; …; R_k mutuellement exclusifs, avec probabilités p_1, …, p_k, respectivement. Le vecteur aléatoire $\boldsymbol{X} = [X_1; \ldots; X_k]$, où X_i (i = 1,…, k) est le nombre d'occurrences du résultat R_i suit une loi appelée* ***loi multinomiale*** *de paramètres n* et $p_1; p_2; \ldots; p_k$. *Chacune des composantes de* $\boldsymbol{X}$ *prend sa valeur dans l'ensemble* $\{0; 1; \ldots; n\}$. Notons que $x_1 + x_2 + \ldots + x_k = n$. *De plus, on a*

$$\mathrm{P}\left(\boldsymbol{X} = [x_1; x_2; \ldots; x_k]\right) = \binom{n}{x_1; x_2; \ldots; x_k} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \tag{2.6.2}$$

On écrit $\boldsymbol{X} \sim \mathcal{M}(n; p_1; p_2; \ldots; p_k)$ pour signifier que le vecteur $\boldsymbol{X}$ est de loi multinomiale de paramètres n et p_1, p_2 , … , p_k.

Propriétés de la loi multinomiale

- Si $k = 2$, $\boldsymbol{X} = [X_1 ; X_2]$, alors X_1 est de loi $\mathcal{B}(n ; p_1)$;
- Si $\boldsymbol{X} = [X_1 ; X_2 ; \ldots ; X_k]$ est de loi $\mathcal{M}(n; p_1 ; p_2 ;\ldots ; p_k)$, alors chaque composante X_i est de loi $\mathcal{B}(n ; p_i)$;
- La covariance entre X_i et X_j ($i \neq j$) est donnée par $\mathrm{Cov}(X_i ; X_j) = -np_i p_j$,
- Plus généralement, si $\boldsymbol{X} = [X_1; \ldots ; X_j ; X_{j+1} ; \ldots ; X_k] \sim \mathcal{M}(n; p_1 ;\ldots; p_j ; p_{j+1} ;\ldots ; p_k)$, alors conditionnellement étant donné $X_1 + \ldots + X_\ell = m$ la loi conjointe de $[X_1; \ldots ; X_j]$ et $[X_{j+1} ; \ldots ; X_k]$ celle d'un produit de deux vecteurs indépendants, de lois $\mathcal{M}\left(m; \frac{p_1}{p_1+\ldots+p_j};\ldots;\frac{p_j}{p_1+\ldots+p_j}\right)$ et $\mathcal{M}\left(n-m; \frac{p_{j+1}}{p_{j+1}+\ldots+p_k};\ldots;\frac{p_k}{p_{j+1}+\ldots+p_k}\right)$, respectivement.

***Exemple* 2.6.1** *Loi multinomiale*

Lors d'une élection politique, trois partis sont en lice. Si dans la population 50 % des gens favorisent le parti A, 40% le parti B, et 10% le parti C, quelle est la probabilité que sur 6 personnes choisies au hasard dans cette province, 3 favorisent A, 1 favorise B et 2 favorisent C ?

Ici, $\boldsymbol{X} \sim \mathcal{M}(6; 0{,}5; 0{,}4; 0{,}1)$. Par conséquent,

$$\mathrm{P}(\boldsymbol{X} = 3;\, 1;\, 2]) = \binom{n}{3;\ 1;\ 2}(0,5)^3(0,4)^1(0,1)^2 = \frac{6!}{3!1!2!}\times 0,125\times 0,4\times 0,01 = 0,03\,.$$

Remarque *Il découle du sens même d'un vecteur* $[X_1;\, X_2;\, \ldots; X_n]$ *de loi multinomiale que chacune des variables* X_i *est de loi* $\mathcal{B}(n\,;\,p_i)$. *On peut même dire plus: toute* somme *de k des n variables (k < n) est de loi binomiale. Si* $[X_{i_1}; X_{i_2}; \ldots; X_{i_k}]$ *est un sous-ensemble de k variables, alors*

$$X_{i_1} + X_{i_2} + \ldots + X_{i_k} \sim \mathcal{B}(n;\, p_{i_1} + p_{i_2} + \ldots + p_{i_k})\,.$$

2.7 Quelques démonstrations

Théorème 2.1.1

À montrer que si $X \sim \mathcal{B}(n;\, p)$, alors i) $\mathrm{E}[X] = np$; ii) $\mathrm{Var}[X] = np(1-p) = npq$; iii) $M_X(t) = (q+pe^t)^n$.

i) $$\mathrm{E}(X) = \sum_{x=0}^{n} x\binom{n}{x}p^x q^{n-x} = \sum_{x=1}^{n} x\frac{n!}{x!(n-x)!}p^x q^{n-x} = \sum_{x=1}^{n}\frac{n(n-1)!}{(x-1)!(n-x)!}p^x q^{n-x}$$

$$= np\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!(n-x)!}p^{x-1}q^{n-1-(x-1)} = np\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}p^y q^{m-y}$$

$$= np\sum_{y=0}^{m}\binom{m}{y}p^y q^{m-y}\text{, où } y = x\text{-}1 \text{ et } m = n\text{-}1.$$

La dernière somme étant égale à $(p+q)^m = 1$ par le binôme de Newton, nous avons le résultat $\mathrm{E}[X] = np$.

ii) Pour déterminer la variance de X, notons que

$\mathrm{Var}[X] = \mathrm{E}[X^2] - [\mathrm{E}(X)]^2 = \mathrm{E}[X^2] - (np)^2 = \mathrm{E}[X^2] - \mathrm{E}(X) + np - (np)^2 = \mathrm{E}(X^2 - X) + np - (np)^2 = \mathrm{E}[X(X-1)] + np - (np)^2$.

Nous montrons maintenant que le premier terme, $\mathrm{E}[X(X-1)]$, est égal à $n(n-1)p^2$.

$$\mathrm{E}[X(X-1)] = \sum_{x=0}^{n} x(x-1)\binom{n}{x}p^x q^{n-x} = \sum_{x=2}^{n} x(x-1)\frac{n(n-1)(n-2)!}{x(x-1)(x-2)!(n-x)!}p^x q^{n-x} =$$

$$n(n-1)p^2\sum_{x=2}^{n}\frac{(n-2)!}{(x-2)!(n-x)!}p^{x-2}q^{n-x} = n(n-1)p^2\sum_{y=0}^{m}\frac{m!}{y!(m-y)!}p^y q^{m-y}\text{, où } y = x\text{-}2 \text{ et } m = n\text{-}2.$$

La dernière somme étant égale à $(p+q)^m = 1$ par le binôme de Newton, nous avons le résultat $\mathrm{E}[X(X-1)] = \mathrm{E}(X^2) - \mathrm{E}(X) = n(n-1)p^2 \Rightarrow \mathrm{E}(X^2) = \mathrm{E}(X) + n(n-1)p^2 = np + n(n-1)p^2 \Rightarrow \mathrm{Var}(X) = \mathrm{E}[X^2] - (np)^2 = np + n(n-1)p^2 - (np)^2 = npq$.

iii) $M_X(t) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x q^{n-x} = \sum_{x=0}^{n}\binom{n}{x}(e^t p)^x q^{n-x} = (q+pe^t)^n$ par le binôme de Newton.

Théorème 2.2.1

À montrer que si $X \sim \mathcal{H}(n;\, N_1,\, N_2)$, alors; i) $\mathrm{E}[X] = np = n\dfrac{N_1}{N}$; ii) $\mathrm{Var}[X] = np(1-p)\ \dfrac{N-n}{N-1}$.

Nous utiliserons la propriété d'analyse combinatoire suivante :

$\binom{N_1+N_2}{n} = \binom{N_1}{0}\binom{N_2}{n} + \binom{N_1}{1}\binom{N_2}{n-1} + \ldots + \binom{N_1}{n}\binom{N_2}{0}$, où, par convention, la combinaison $\binom{n}{k} = 0$ si $k > n$. (C'est par cette égalité qu'on vérifie le fait que la somme des probabilités est égale à 1).

On utilisera aussi le fait que $\dbinom{N}{n} = \dfrac{N}{n}\dbinom{N-1}{n-1} = \dfrac{N(N-1)}{n(n-1)}\dbinom{N-2}{n-2}$.

i) L'espérance de X est

$$E[X] = \sum_{x=0}^{n} x\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} = \sum_{x=1}^{n} x\frac{N_1}{x}\frac{\binom{N_1-1}{x-1}\binom{N_2}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{nN_1}{N}\sum_{x=1}^{n}\frac{\binom{N_1-1}{x-1}\binom{N_2}{n-x}}{\binom{N-1}{n-1}}$$

$= \frac{nN_1}{N}\sum_{y=0}^{m}\frac{\binom{N_1-1}{y}\binom{N_2}{m-y}}{\binom{N-1}{m}}$, où $m = n$-1 et $y = x$-1.

Mais $\sum_{y=0}^{m}\frac{\binom{N_1-1}{y}\binom{N_2}{m-y}}{\binom{N-1}{m}} = 1$ car c'est la somme des probabilités des valeurs d'une variable aléatoire Y de loi $\mathcal{H}(m\,;N_1\text{-}1;N_2)$. Nous avons donc $E[X] = \frac{nN_1}{N} = np$.

ii) Pour déterminer la variance, commençons par montrer que $E[X(X\text{-}1)] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)}$.

$$E[X(X\text{-}1)] = \sum_{x=0}^{n} x(x-1)\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} = \sum_{x=2}^{n} x(x-1)\frac{N_1(N_1-1)}{x(x-1)}\frac{\binom{N_1-2}{x-2}\binom{N_2}{n-x}}{\frac{N(N-1)}{n(n-1)}\binom{N-2}{n-2}} =$$

$\frac{n(n-1)N_1(N_1-1)}{N(N-1)}\sum_{x=2}^{n}\frac{\binom{N_1-2}{x-2}\binom{N_2}{n-x}}{\binom{N-2}{n-2}} = \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\sum_{y=0}^{m}\frac{\binom{M_1}{m}\binom{N_2}{m-y}}{\binom{M}{m}}$, $M_1 = N_1$-2, $M = N$-2, $m = n$-2, $y = x$-2.

Mais $\sum_{y=0}^{m}\frac{\binom{M_1}{m}\binom{N_2}{m-y}}{\binom{M}{m}} = 1$ car c'est la somme des probabilités des valeurs d'une variable aléatoire Y de loi $\mathcal{H}(m; M_1;N_2)$. Donc $E[X(X\text{-}1)] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)}$.

Alors $\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = [E(X^2) - E(X)] + E(X) - [E(X)]^2 = E[X(X\text{-}1)] + E(X) - [E(X)]^2 =$

$$\frac{n(n-1)N_1(N_1-1)}{N(N-1)} + \frac{nN_1}{N} - \left(\frac{nN_1}{N}\right)^2 = n\frac{N_1}{N}\frac{N_2}{N}\frac{N-n}{N-1} = npq\frac{N-n}{N-1}.$$

Théorème 2.3.1

À montrer que si $X \sim \mathcal{G}(p)$, alors

i) $E[X] = \frac{1}{p}$; ii) $\mathrm{Var}[X] = \frac{q}{p^2}$; iii) $M_X(t) = \frac{pe^t}{1-qe^t}$.

Nous savons que lorsque $|q| < 1$, la somme $\sum_{x=0}^{\infty} q^x = 1+q+q^2+\ldots$ converge et $\sum_{x=0}^{\infty} q^x = \frac{1}{1-q}$. Ce qui nous permet de vérifier que la somme des probabilités d'une variable de loi $\mathcal{G}(p)$ est bien égale à 1, puisque

$$\sum_{x=1}^{\infty} q^{x-1}p = p(1+q+q^2+\ldots) = \frac{p}{1-q} = \frac{p}{p} = 1.$$

i) L'espérance de X est $E(X) = \sum_{x=1}^{\infty} xq^{x-1}p = p\left(\sum_{x=1}^{\infty}\frac{d}{dq}q^x\right) = p\frac{d}{dq}\left(\sum_{x=0}^{\infty}q^x - 1\right) =$

$$p\frac{d}{dq}\left(\sum_{x=0}^{\infty}q^x\right) = p\frac{d}{dq}\left(\frac{1}{1-q}\right) = p\frac{1}{(1-q)^2} = p\frac{1}{p^2} = \frac{1}{p}.$$

ii) Pour déterminer la variance de X nous commençons par montrer que $E[X(X\text{-}1)] = \frac{2}{p^3}$:

$$\mathrm{E}[X(X\text{-}1)] = \sum_{x=1}^{\infty} x(x-1)q^{x-1}p = pq\sum_{x=2}^{\infty} x(x-1)q^{x-2} = pq\left(\sum_{x=2}^{\infty}\frac{d^2}{dq^2}q^x\right) = pq\left(\frac{d^2}{dq^2}\sum_{x=2}^{\infty}q^x\right) =$$

$$pq\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right) = pq\frac{d}{dq}\left(\frac{2q-q^2}{(1-q)^2}\right) = pq\left(\frac{(1-q)^2(2-2q)+2(1-q)(2q-q^2)}{(1-q)^4}\right) = \frac{2}{p^3}.$$

D'où, $\mathrm{Var}(X) = \mathrm{E}[X(X\text{-}1)] + \mathrm{E}(X)\text{-}[\mathrm{E}(X)]^2 = \dfrac{2}{p^3}+\dfrac{1}{p}-\dfrac{1}{p^2} = \dfrac{q}{p^2}$.

iii) $M_X(t) = \sum_{x=1}^{\infty} e^{tx}pq^{x-1} = pe^t\sum_{x=1}^{\infty}e^{t(x-1)}q^{x-1} = pe^t\sum_{x=1}^{\infty}e^{t(x-1)}q^{x-1} = pe^t\sum_{y=0}^{\infty}(e^tq)^y = \dfrac{pe^t}{1-qe^t}$.

Théorème 2.4.1

À montrer que si $X \sim \mathcal{B}^-(n;p)$, alors i) $\mathrm{E}[X] = \dfrac{n}{p}$; ii) $\mathrm{Var}[X] = \dfrac{nq}{p^2}$; iii) $M_X(t) = \left(\dfrac{pe^t}{1-qe^t}\right)^n$.

Une variable de loi $\mathcal{B}^-(n;p)$ peut s'écrire comme une somme $X = X_1+X_2+\ldots+X_n$, où X_i est le nombre d'essais nécessaire, après le $(i\text{-}1)^e$ succès, pour obtenir un i^e succès. Chaque X_i est de loi $\mathcal{G}(p)$, donc de moyenne $1/p$, de variance q/p^2 et de fonction génératrice des moments $M_X(t) = \dfrac{pe^t}{1-qe^t}$.

i) $\mathrm{E}(X) = \mathrm{E}(X_1+X_2+\ldots+X_n) = n/p$;

ii) $\mathrm{Var}(X) = \mathrm{Var}(X_1+X_2+\ldots+X_n) = nq/p^2$;

iii) $M_X(t) = \prod_{i=1}^{n}\dfrac{pe^t}{1-qe^t} = \left(\dfrac{pe^t}{1-qe^t}\right)^n$

Théorème 2.5.1

On se sert du fait que $\sum_{i=0}^{\infty}\dfrac{\lambda^i}{i!} = e^{\lambda}$ (c'est ce qui permet de conclure que la somme des probabilités donne 1.)

i) $\mathrm{E}(X) = \sum_{x=0}^{\infty} x\dfrac{e^{-\lambda}\lambda^x}{x!} = \lambda\sum_{x=1}^{\infty}\dfrac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda\sum_{y=0}^{\infty}\dfrac{e^{-\lambda}\lambda^y}{y!} = \lambda$.

ii) Pour déterminer la variance, nous montrons que $\mathrm{E}[X(X\text{-}1)] = \lambda^2$.

$$\mathrm{E}[X(X\text{-}1)] = \sum_{x=0}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=2}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\lambda^2\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} =$$

$e^{-\lambda}\lambda^2\sum_{y=0}^{\infty}\dfrac{\lambda^y}{y!} = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2$. Alors $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = [\mathrm{E}(X^2) - \mathrm{E}(X)] + \mathrm{E}(X) - [\mathrm{E}(X)]^2 =$
$\mathrm{E}[X(X\text{-}1)] + \lambda - \lambda^2 = \lambda^2+\lambda-\lambda^2$.

iii) $M_X(t) = \sum_{x=0}^{\infty}e^{tx}\dfrac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\dfrac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{e^t\lambda} = e^{\lambda(e^t-1)}$

Théorème 2.5.2

À montrer que si $X_1, X_2, \ldots, X_n, \ldots$ est une suite de variables aléatoires, $X_n \sim \mathcal{B}(n;p_n)$, où p_n est une suite telle que $\lim_{n\to\infty} np_n = \lambda$, alors X_n tend en loi vers une loi $\mathcal{P}(\lambda)$. Nous montrons que

$\binom{n}{x}p^x(1-p)^{n-x}\xrightarrow[n\to\infty]{}\dfrac{e^{-\lambda}\lambda^x}{x!}$ et donc que $P(X=x)\xrightarrow[n\to\infty]{}\dfrac{e^{-\lambda}\lambda^x}{x!}$.

Nous utilisons le fait que $\left(1-\dfrac{\lambda}{n}\right)^n \to e^{-\lambda}$ lorsque $n\to\infty$.

$$\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} = \frac{n(n-1)\ldots[n-(x-1)]}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$
$$= \frac{\lambda^x}{x!}\frac{n(n-1)\ldots[n-(x-1)]}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$
$$= \underbrace{\frac{\lambda^x}{x!}}_{\to\lambda^x/x!}\underbrace{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)}_{\to 1}\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\to e^{-\lambda}}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-x}}_{\to 1}$$

Ce produit tend vers $e^{-\lambda}\lambda^x/x!$

Théorème 2.5.3

À montrer que si $X_i \sim \mathcal{P}(\lambda_i)$, $i = 1, 2, \ldots, n$, sont des variables aléatoires indépendantes, alors $X = \sum_{i=1}^n X_i \sim \mathcal{P}(\sum_{i=1}^n \lambda_i)$.

On le démontre pour $n = 2$. Donc soit $X \sim \mathcal{P}(\lambda_1)$ et $Y \sim \mathcal{P}(\lambda_2)$, X et Y indépendantes. Soit $Z = X+Y$. Nous montrons que $Z \sim \mathcal{P}(\lambda)$, où $\lambda = \lambda_1+\lambda_2$, c'est-à-dire, que pour $z = 0, 1, \ldots$, on a $P(Z = z) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}$.

$P(Z = z) = P(X+Y= z) = P(X=0;Y=z)+P(X=1;Y=z-1)+\ldots+P(X=z;Y=0) = \sum_{j=0}^{z}\frac{e^{-\lambda_1}\lambda_1^j}{j!}\frac{e^{-\lambda_2}\lambda_2^{z-j}}{(z-j)!} =$
$\frac{1}{z!}e^{-\lambda_1}e^{-\lambda_2}\sum_{j=0}^{z}\frac{z!}{j!(z-j)!}\lambda_1^j\lambda_2^{z-j} = \frac{1}{z!}e^{-(\lambda_1+\lambda_2)}\sum_{j=0}^{z}\binom{z}{j}\lambda_1^j\lambda_2^{z-j} = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}$, par le binôme de Newton.

Pour n quelconque, la technique des fonctions génératrices des moments offre une démonstration facile. La fonction génératrice de X est égale au produit des fonctions génératrices des moments.

Donc $M_X(t) = \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n e^{\lambda_i(e^t-1)} = e^{(\Sigma\lambda_i)(e^t-1)}$, ce qui montre que X est de loi de Poisson de paramètre $\Sigma\lambda_i$.

RÉSUMÉ

Distribution	*Modalités de X*	*Fonction de probabilité*	E(*X*)	Var(*X*)	$M_X(t)$
Binomiale $\mathcal{B}(n;\, p)$	$x \in \{0, 1, \ldots, n\}$	$p(x) = \binom{n}{x} p^x(1\text{-}p)^{n\text{-}x}$	np	$np(1\text{-}p)$	$(q+pe^t)^n$
Poisson $\mathcal{P}(\lambda)$	$x \in \{0, 1, 2, \ldots\}$	$p(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$	λ	λ	$e^{\lambda(e^t-1)}$
Hypergéométrique $\mathcal{H}(n;\, N_1\, ;\, N_2)$	$0 \le x \le m$ $0 \le n - x \le N\text{-}N_1$	$p(x) = \dfrac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}}$	$np,$ $p = \dfrac{N_1}{N}$	$npq\dfrac{N-n}{N-1},$ $q = 1\text{-}p$	
Binomiale négative $\mathcal{B}^{-}(p)$	$x \in \{n, n+1, \ldots\}$	$p(x) =$ $\binom{x-1}{n-1} p^n(1\text{-}p)^{x\text{-}n}$	$\dfrac{n}{p}$	$\dfrac{n(1-p)}{p^2}$	$\left(\dfrac{pe^t}{1-(1-p)e^t}\right)^n$

Exercices

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 ?

a) 0,2618; b) 0,2461; c) 0,4920.

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

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

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? 0,005299241

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

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? 8 fois ou plus ? 0,04651361 ; 0,2790816

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) $P(X \leq 30)$; iv) $P(25 \leq X \leq 30)$. i) 0,004320298; ii) 0,9802937; iii) 0,9957873; iv) 0,008366395

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

$E(X) = 60$; $Var(X) = 300$.

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) $P(X \leq 30)$; iv) $P(25 \leq X \leq 30)$ i) 0,07679025; ii) 0,7085281; iii) 0,2914719. iv) 0,2780228

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

$E(X) =$ 33,3333 ; $Var(X) =$ 22,22222.

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) $P(X \leq 2)$ iii) $P(X > 3)$; iii) $P(X \leq 3)$.

i) 0,449329; ii) 0,9525774; iii) 0,0474226 iv) 0,9909201.

b) Supposons que λ n'est pas connu. Quelle est la plus petite valeur de λ pour laquelle $P(X = 0) \leq 0{,}05$? 2,995732

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).

Probabilités exactes (loi binomiale)

X	0	1	2	3	4	5	6	7	8
px	0,4304672	0,3826375	0,1488035	0,03306744	0,0045927	0,00040824	2,268e-05	7,2e-07	1e-08

Probabilités approximatives (loi de Poisson)

X	0	1	2	3	4	5	6	7	8
pois	0,449329	0,3594632	0,1437853	0,03834274	0,007668548	0,001226968	0,000163596	1,8697e-05	1,87e-06

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? Quelle est la probabilité que 10 minutes s'écoulent sans aucune naissance ? 0,6942913; 0,00077.

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?

0,12511 (utilisant la loi de Poisson, un modèle crédible dans les circonstances).

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). 0,1600727

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.

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 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 $p = 0{,}25$. $E(X) = 7{,}51$ si $p = 0{,}1$ et $E(X) = 10{,}44$ 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)? On épargne des tests si $p = 0{,}1$ mais pas si $p = 0{,}25$.

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 ? 0,22.
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? 0,17667936
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 »? 0,5470082
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 ? 0,002396205.
   (ii) Quelle est la probabilité que 3 des 4 bouteilles de l'échantillon aient été défectueuses ? 0,06024744486.
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] 0,003076923

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 péage? 0,06064152
(On suppose que nombre de véhicules privés et le nombre de véhicules commerciaux sont des variables indépendantes).
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 ? 0,3456

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 ? 0,06948345
d) Montrer, plus généralement, que le nombre de taxis libres qui arrivent en 20 minutes est de loi de Poisson de paramètre $0{,}4\lambda$.

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
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 ? 0,6778.

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, $|X - E(X)|$, aussi grand que (c'est-à-dire, supérieur ou égal à) l'écart observé de 5. 0,014520

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 ≤ 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 ?

La probabilité d'accepter le lot si p = 0,10 est 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 P$\{X = 0\}$ en supposant que $\lambda = 5$. La valeur $\lambda = 5$ est-elle plausible?

b) Calculer 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 ≤ 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 α = 0,05? 0,03827

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

d) Quelle est la probabilité d'accepter un lot qui est inacceptable dans le sens que α = 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 (α = 0,10) ne soit pas supérieure à 5%?

Voici les probabilités de rejeter un lot acceptable (α = 0,1) selon le critère de rejet :

	x			
On rejette si $X \geq x$	5	6	7	8
Probabilité de rejet	0,1088220	0,04202104	0,01418731	0,00424670

f) Avec la règle développée en e), quelle est la probabilité d'accepter un lot qui est inacceptable, dans le sens que α = 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 $E[\hat{p}(1-\hat{p})] = \dfrac{(n-1)\,p(1-p)}{n}$.

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

2.47 Soit X une variable de loi géométrique de paramètre p.

a) Montrer que $E(1/X) = -p\,(\ln p)/(1-p)$;

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 $X \sim \overline{\mathcal{B}}(n; p)$, 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 $\overline{\mathcal{B}}(n; p_n)$. 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 $Var(X) = npq(N\text{-}n)/(N\text{-}1)$.

2.61 Calculez les valeurs de la fonction de probabilité d'une variable de loi $\mathcal{P}(0,3)$ pour $x = 0, 1, \ldots, 20$, 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 une 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.

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.

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.

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 μ = 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 λ = 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.

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.

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.

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.

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. ---

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;

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.

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.

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.

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.

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.

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 $N = 30$

Annexe
Espérance, variance, covariance

Soit X une variable aléatoire discrète de fonction de probabilité $p(x)$.

Définition *Espérance d'une variable aléatoire*

L'espérance mathématique μ de X est définie par

$$\mu = E(X) = \sum_x xp(x) .$$

Plus généralement, l'espérance d'une fonction φ(X) est définie par:

$$E[\varphi(X)] = \sum_x \varphi(x)p(x) .$$

Définition *Variance d'une variable aléatoire*

La *variance* σ^2 de X est définie par

$$\mathrm{Var}(X) = \sum_x (x-\mu)^2 p(x) = E[(X-\mu)^2]$$

Propriété *Une autre expression de la variance*

$$\mathrm{Var}(X) = \sum_x x^2 p(x) - \mu^2 = E(X^2) - [E(X)]^2.$$

Écart-type *Écart-type*

L'écart-type σ de X est la racine carrée de sa variance :

$$\sigma = \sqrt{V(X)} .$$

Propriété *Fonctions affines*

Soit $Y = a + bX$, où a et b sont des constantes. Alors

$$E(Y) = E(a + bX) = a + bE(X) \text{ et } \mathrm{Var}(Y) = \mathrm{Var}(a + bX) = b^2\mathrm{Var}(X).$$

Vecteur aléatoire : distributions conjointe, marginale et conditionnelle

Définition *Fonction de probabilité conjointe*

La fonction de probabilité conjointe d'un vecteur aléatoire $\boldsymbol{X} = [X_1 ; X_2 ; \dots ; X_n]$ est une fonction $p_{\boldsymbol{X}}(x_1 ; x_2 ; \dots ; x_n)$ qui fait correspondre à chaque vecteur de nombres $\boldsymbol{x} = [x_1 ; x_2 ; \dots ; x_n]$ la probabilité $P(\boldsymbol{X} = \boldsymbol{x})$:

$$p_{\boldsymbol{X}}(x_1 ; x_2 ; \dots ; x_n) = P(X_1 = x_1 \text{ et } X_2 = x_2 \text{ et } \dots \text{ et } X_n = x_n).$$

Définition *Fonction de probabilité marginale (cas $n = 2$)*

Soit $p_{X;Y}(x ; y)$ la fonction de probabilité conjointe du vecteur aléatoire $[X ; Y]$.

La fonction de probabilité $p_X(x)$ de X est appelée fonction de probabilité *marginale* lorsqu'elle est déterminée à partir de la fonction de probabilité conjointe $p_{X;Y}(x ; y)$.

$$p_X(x) = \sum_y p_{X;Y}(x;y)$$

[De même, $p_Y(y) = \sum_x p_{X;Y}(x;y)$.]

Définition Distributions conditionnelles

La fonction de probabilité conditionnelle de Y étant donné $X = x$ est définie par

$$p_{Y|x}(y) = \frac{p_{X,Y}(x;y)}{p_X(x)} .$$

Fonctions scalaires de variables aléatoires

Si $Z = \varphi(X ; Y)$ est une fonction scalaire de $[X ; Y]$, l'espérance de Z est calculée par :

$$E[\varphi(X;Y)] = \sum_x \sum_y \varphi(x;y)\, p_{X;Y}(x;y) .$$

Propriété *Espérance d'une combinaison linéaire de variables aléatoires*

Soit $X_1, X_2, \dots, X_n$ n variables aléatoires et $a_1, a_2, \dots, a_n$ des constantes. Alors

$$E(a_1X_1 + a_2X_2 + \dots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \dots + a_nE(X_n).$$

Covariance

Définition *Covariance*

Soit X et Y deux variables aléatoires de moyennes $E(X) = \mu_x$ et $E(Y) = \mu_y$, et de fonction de probabilité conjointe $p_{X;Y}(x\ ;y)$. La *covariance* entre X et Y, désignée par $\mathrm{Cov}(X\ ;Y)$, est définie par

$$\mathrm{Cov}(X;Y) = E[(X-\mu_x)(Y-\mu_y)] = \sum_x \sum_y [(x-\mu_x)(y-\mu_y)]p_{X;Y}(x;y)\ .$$

Propriété *Une autre expression de la covariance*

$$\mathrm{Cov}(X;Y) = E(XY) - E(X)E(Y)\ .$$

Propriétés *Covariance*

Soit X et Y deux variables aléatoires, et a, b et c des constantes. Alors

(i) $\mathrm{Cov}(X\ ;b) = 0$

(ii) $\mathrm{Cov}(X + Y\ ;Z) = \mathrm{Cov}(X\ ;Z) + \mathrm{Cov}(Y\ ;Z)$

(iii) $\mathrm{Cov}(aX + c;\ bY) = ab\ \mathrm{Cov}(X\ ;Y)$

(iv) $\mathrm{Cov}(X\ ;X) = \mathrm{Var}(X)$

Propriété *Variance d'une combinaison linéaire*

Soit X et Y deux variables aléatoires indépendantes et a_1, a_2, … , a_n des constantes. Alors

$$\mathrm{Var}\left(a_1X_1 + a_2X_2 + \ \dots\ + a_kX_k\right) = \sum_{i=1}^k a_i^2 V(X_i) + \sum_i \sum_{j\neq i} a_i a_j \mathrm{Cov}(X_i;X_j)\ .$$

Cas particuliers

$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X\ ;Y)$;
$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X\ ;Y)$;
$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X\ ;Y)$.

Indépendance de variables aléatoires

Définition *Variables aléatoires indépendantes*

Les variables aléatoires X_1, X_2, …, X_n sont dites *indépendantes* si pour tout vecteur $[x_1;\ x_2\ ;\ \dots\ ;\ x_n]$

$$px(x_1\ ;x_2\ ;\ \dots\ ;x_n) = p_{X_1}(x_1)\times p_{X_2}(x_2)\times\ldots\times p_{X_n}(x_n)$$

Covariance et indépendance

Soit X_1, X_2, …, X_n n variables aléatoires indépendantes. Alors

$\mathrm{E}(X_1 X_2\dots X_n) = \mathrm{E}(X_1)\mathrm{E}(X_2)\dots\mathrm{E}(X_n)$
$\mathrm{Cov}(X_i\ ;X_j) = 0$ pour toute paire $(i\ ;j)$, $i = 1, 2, .., n$; $j = 1, 2, .., n$; $i \neq j$
$\mathrm{Var}(a_1X_1 + a_2X_2 + \dots + a_kX_k) = a_1^2\mathrm{Var}(X_1) + a_2^2\mathrm{Var}(X_2) + \dots + a_k^2\mathrm{Var}(X_k)\ .$

Corollaire : $\mathrm{Var}(X - Y) = \mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Espérance et variance conditionnelles

Soit X et Y deux variables aléatoires de fonction de probabilité conjointe $p_{X;Y}(x\ ;y)$ et fonction de probabilité conditionnelle $p_{Y|x}(y)$. Alors l'espérance et la variance conditionnelles sont définies par

$$\mathrm{E}(Y\mid x) = \sum_y y p_{Y|x}(y) \ \text{ et }\ \mathrm{Var}(Y\mid x) = \sum_y [y - E(Y\mid x)]^2 p_{Y|x}(y)\ .$$

Calcul d'une espérance, d'une variance et d'une covariance par conditionnement

Soit X, Y et Z trois variables aléatoires. Alors

$\mathrm{E}(Y) = \mathrm{E}[\mathrm{E}(Y\mid X)]$.
$\mathrm{Var}(Y) = \mathrm{Var}[E(Y\mid X)] + E[\mathrm{Var}(Y\mid X)]$.
$\mathrm{Cov}(X;Y) = \mathrm{Cov}[\mathrm{E}(X\mid Z);E(Y\mid Z)] + E[\mathrm{Cov}(X;Y\mid Z)]$

Coefficient de corrélation

Le *coefficient* de *corrélation* $\rho(X\,;\,Y)$ entre X et Y, également désigné par ρ ou par $\rho_{X\,Y}$, est défini par

$$\rho = \frac{\mathrm{Cov}(X\,;Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$