Chapitre 2

Lois discrètes

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.

Nous définirons dans ce chapitre quatre des lois les plus c0urantes (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 binomial* 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*[1]. 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; $F(x) = P(X \leq x)$.
$E(X)$ 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 satisfait 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.

---

[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.

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

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

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:

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

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) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

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

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

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 $\left(\frac{24}{25}\right)^n \leq 0{,}30 \Leftrightarrow n \ln(24/25) \leq \ln(0{,}30) \Leftrightarrow n \geq \frac{\ln(0{,}30)}{\ln(24/25)} = 29{,}49.$

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 $P(X = 0 \mid p = 0{,}10)$, où la barre verticale « | » signifie « étant donné que ». On trouve $P(X = 0 \mid p = 0{,}10) = \binom{35}{0}(0{,}10)^0(0{,}90)^{35} = 0{,}025$. 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 considèrerait « 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 $P(X = 0 \mid p) \geq 0{,}05$. Donc p est « vraisemblable » si $(1\text{-}p)^{35} \geq 0{,}05 \Leftrightarrow 1\text{-}p \geq (0{,}05)^{1/35}$ $\Leftrightarrow p \leq 1\text{-}(0{,}05)^{1/35} = 0{,}082$.

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?

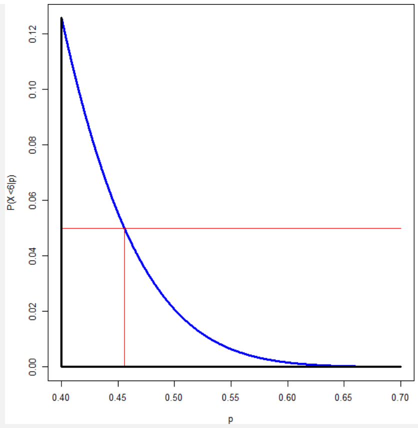
Réponse Si $p = 0{,}6$, on s'attend à trouver 12 croyants dans l'échantillon : $E(X) = np = 12$. Le nombre observé ($x = 5$), nettement inférieur à 12, peut-il être attribué au seul hasard? On calcule la probabilité d'un nombre aussi petit que 5, soit $P(X \leq 5 \mid p = 0{,}6)$, où la barre verticale « | » signifie « étant donné que » : $P(X \leq 5 \mid p = 0{,}6) = \sum_{i=0}^{5}\binom{20}{i}(0{,}6)^i(0{,}4)^{20-i} =$ 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 $P(X \leq 5 \mid p = 0{,}4) = 0{,}1256$. 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 $P(X \leq 5 \mid p)$ et exclure toute valeur de p pour laquelle cette probabilité est *trop faible*. Définissons cette notion : la probabilité $P(X \leq 5 \mid p)$ 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 $P(X \leq 5 \mid p) < 0{,}05$ si $p \geq 0{,}4555$. On conclut donc que p doit se situer quelque part à gauche de 0,4555. Le graphique suivant présente la probabilité $P(X \leq 5 \mid p)$ 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 choisit 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 pourrait 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.*

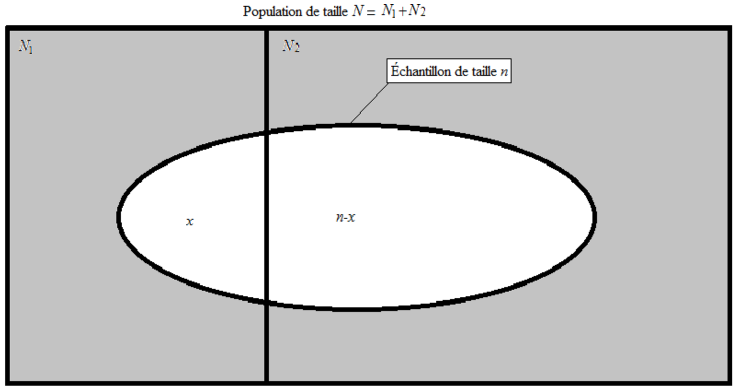
2.2 Loi hypergéométrique

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 ?

Population de taille $N = N_1 + N_2$



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

$$p = \frac{N_1}{N} \tag{2.2.1}$$

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:

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

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

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

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

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

b) Si nous supposons que $X \sim \mathcal{B}(5;\, 5/11)$, nous avons $P(X = 3) = \binom{5}{3}\left(\dfrac{5}{11}\right)^3\left(\dfrac{6}{11}\right)^2 = 0{,}2794$. 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 $P(X = x)$ 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 $P[X = 3]$ 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) $\binom{30}{3}\binom{42}{2}\Big/\binom{72}{5} = 0{,}24984$.

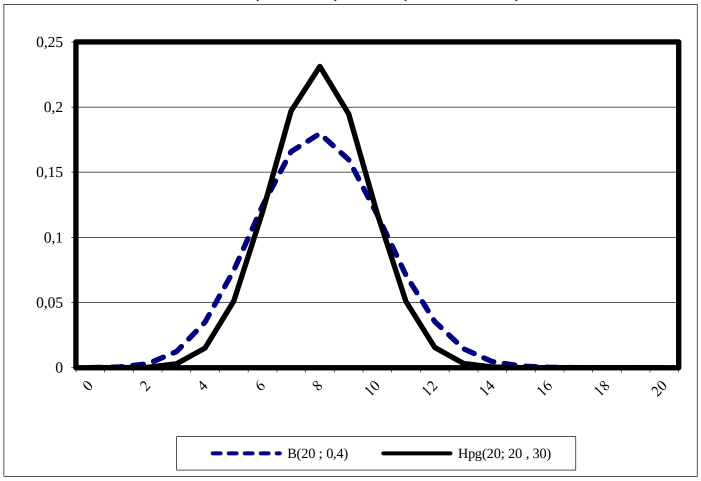
Si nous supposons que $X \sim \mathcal{B}(5;\, 5/12)$, on a $P[X = 3] = \binom{5}{3}(5/12)^3(7/12)^2 = 0{,}24615$. Le tableau suivant résume les résultats:

Loi	$P(X = 3)$
$\mathcal{H}(5;\, 5,\, 7)$	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

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

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

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

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 $x = 1, \ldots, 6$, et calculer $\mathrm{P}(X \le 6)$. Déterminer $\mathrm{E}(X)$ et $\mathrm{Var}(X)$.

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

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

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

$= 0,737856$

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

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 $x = 1, \ldots, 6$, et calculer $\mathrm{P}(X > 6)$. Déterminer $\mathrm{E}(X)$ et $\mathrm{Var}(X)$.

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

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

$P(X > 6) = q^6 = (0,8)^6 = 0,262144$. $E(X) = 1/p = 5$; $Var(X) = q/p^2 = 20$.

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 4^e^ 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 2^e^ tirage. Ou jusqu'au 3^e^, 4^e^, …, ou n^e tirage.

Les valeurs de X sont $\{n, n+1, \dots\}$. 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)$\times$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 ?

b) Quelle est la probabilité de devoir tirer 300 transaction pour en trouver 2 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) $\mathrm{E}(X) = n/p = 3/0{,}01 = 300$.

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

c) Nous devons calculer la probabilité $\mathrm{P}(X > 300)$, ce qui, à première vue peut sembler une tâche colossale. $\mathrm{P}(X > 300) = \sum_{x=301}^{\infty} p(x) = 1 - \sum_{x=3}^{300} p(x)$. 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

$\mathrm{P}(Y < 3) = \mathrm{P}(Y = 0) + \mathrm{P}(Y = 1)$

$= \binom{300}{0}(0{,}01)^0(0{,}99)^{300} + \binom{300}{1}(0{,}01)(0{,}99)^{299} + \binom{300}{2}(0{,}01)^2(0{,}99)^{298}$

$= 0{,}049 + 0{,}149 + 0{,}224 = 0{,}422$.

Théorème 2.4.2 *Additivité de la loi binomiale négative*

Soit $X_1, X_2, \ldots, X_k$ sont des variables aléatoires indépendantes de loi binomiale négative de paramètres $n_1, n_2, \ldots\ n_k$, 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 *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*

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

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

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

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) $\mathrm{P}(X = 4)$ b) $\mathrm{P}(X \le 3)$

Solution

a) $\mathrm{P}(X = 4) = \dfrac{e^{-\lambda}\lambda^4}{4!}$. Mais que vaut λ ? Sachant que $\lambda = \mathrm{E}(X)$ 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 $\mathrm{P}(X = 4)$ $= \dfrac{e^{-6{,}82}(6{,}821918)^4}{4!} = 0{,}0984$.

b) $\mathrm{P}(X \le 3)$

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

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

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 $P(X_n = x) = \binom{n}{x} p_n^{\,x}(1-p_n)^{n-x}$.

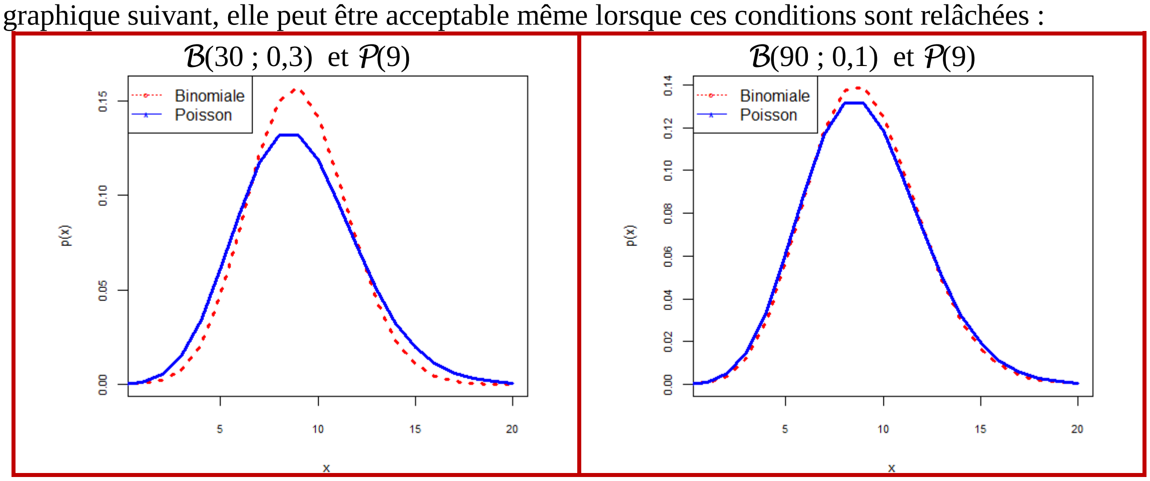
Ce qu'affirme le théorème, c'est que lorsque $n \to\infty$, et $p_n \to 0$ de telle sorte que $np_n \to\lambda$, alors $\binom{n}{x} p_n^{\,x}(1-p_n)^{n-x} \to \dfrac{e^{-\lambda}\lambda^x}{x!}$.

Concrètement, cela signifie que lorsque n est grand et p petit,

$$\binom{n}{x} p_n^{\,x}(1-p_n)^{n-x} \approx \frac{e^{-np}(np)^x}{x!}$$

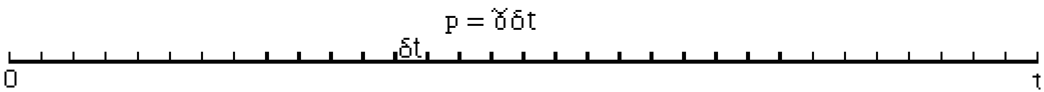
Donc que la distribution $\mathcal{B}(n;\ p)$ peut être approchée par une loi $\mathcal{P}(\lambda)$, avec $\lambda = np$: $P(X = x) \approx \dfrac{e^{-np}(np)^x}{x!}$.

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.*

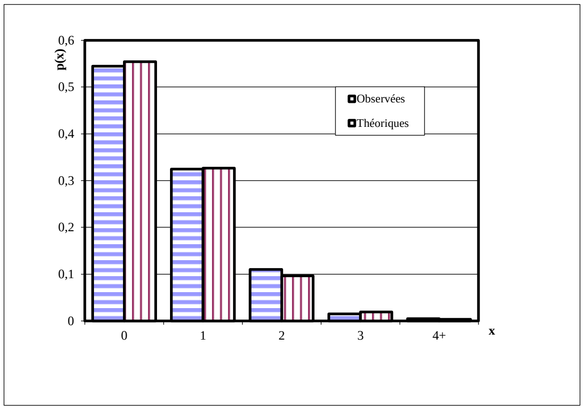
Remarque *La loi de Poisson diffère des lois binomiale et hypergéométrique, géométrique et binomiale négative en ce que les conditions qui justifient son utilisation sont propres au phénomène observé et échappent au contrôle de l'observateur. Lorsqu'on utilise une loi binomiale, c'est qu'on a effectué 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 $\lambda = 0{,}59$. 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.

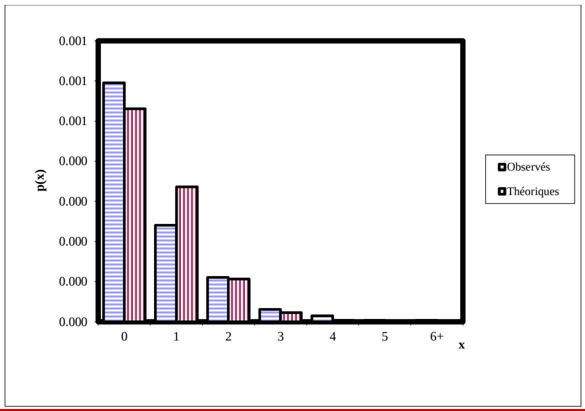
La situation est rarement aussi claire. Dans le prochain exemple, une comparaison des fréquences observées avec les fréquences théoriques révèle des différences plus importantes que dans le dernier exemple.

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 $X_1, X_2, \ldots, X_6$ sont les nombres de verres demandés par les 6 voyageurs, respectivement, alors la consommation totale $X = X_1 + \ldots + X_6$ suit une loi $\mathcal{P}(7,2)$. Donc $P(X \le 10) = 0{,}8867$. Ce calcul exige que la variable X suive une loi de Poisson, ce qui suppose que les variables $X_1, \ldots, X_6$ sont indépendantes. Le sont-elles ? Si les six voyageurs sont en fait trois couples, il y aura des dépendances 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, $(x_1 ; x_2 ; \ldots ; x_6)$, 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, $\{R_1 ; R_2 ; \ldots ; R_k\}$, le résultat R_i avec probabilité p_i, $i = 1, \ldots, k$.

Considérons les k variables aléatoires $X_1, \ldots, X_k$, où X_i est le nombre de fois (parmi n) où le résultats R_i se produit ($i = 1, \ldots, k$). Ensemble, ces k variables constituent ce qu'on appelle un *vecteur aléatoire* $\boldsymbol{X}$: $\boldsymbol{X} = [X_1 ; \ldots ; X_k]$.

Remarque *Les probabilités $p_1, p_2, \ldots, p_k$ et les variables $X_1, X_2, \ldots, X_k$ 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

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

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 *Loi multinomiale*

Considérons une expérience constituée de n épreuves indépendantes, chacune donnant lieu à l'un de k résultats $R_1; \ldots; R_k$ mutuellement exclusifs, avec probabilités $p_1, \ldots, p_k$, respectivement. Le vecteur aléatoire $\boldsymbol{X} = [X_1; \ldots; X_k]$, où X_i ($i = 1, \ldots, k$) 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 $p_1, p_2, \ldots, p_k$.

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

$$\mathcal{M}\left(m; \frac{p_1}{p_1+\ldots+p_j};\ldots;\frac{p_j}{p_1+\ldots+p_j}\right) \text{ et } \mathcal{M}\left(n-m; \frac{p_{j+1}}{p_{j+1}+\ldots+p_k};\ldots;\frac{p_k}{p_{j+1}+\ldots+p_k}\right), \text{ respectivement.}$$

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,

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

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) $E[X] = np$; ii) $\mathrm{Var}[X] = np(1-p) = npq$; iii) $M_X(t) = (q + pe^t)^n$.

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

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

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

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

ii) Pour déterminer la variance de X, notons que

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

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

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

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

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

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) $E[X] = np = n\dfrac{N_1}{N}$; 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 $\binom{N}{n} = \frac{N}{n}\binom{N-1}{n-1} = \frac{N(N-1)}{n(n-1)}\binom{N-2}{n-2}$.

i) L'espérance de X est

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

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

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 $\mathrm{E}[X] = \frac{nN_1}{N} = np$.

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

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

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

, $M_1 = N_1$-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 $\mathrm{E}[X(X\text{-}1)] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)}$.

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

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

Théorème 2.3.1

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

i) $\mathrm{E}[X] = \frac{1}{p}$; 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 $\mathrm{E}(X) = \sum_{x=1}^{\infty} xq^{x-1}p = p\left(\sum_{x=1}^{\infty}\frac{d}{dq}q^x\right) = p\frac{d}{dq}\left(\sum_{x=0}^{\infty}q^x - 1\right) =$

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

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

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

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

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

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

Théorème 2.4.1

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

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 $M_X(t) = \frac{pe^t}{1-qe^t}$.

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

Théorème 2.5.1

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

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

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

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

$e^{-\lambda}\lambda^2\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} = e^{-\lambda}\lambda^2e^{\lambda} = \lambda^2$. 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) $M_X(t) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{e^t\lambda} = e^{\lambda(e^t-1)}$

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

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

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

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

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

Théorème 2.5.3

À montrer que si $X_i \sim \mathcal{P}(\lambda_i)$, $i = 1, 2, \dots, n$, 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 allons montrer que $Z \sim \mathcal{P}(\lambda)$, où $\lambda = \lambda_1 + \lambda_2$, c'est-à-dire, que pour $z = 0, 1, \dots,$ on a

$$P(Z = z) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}.$$

$P(Z = z) = P(X + Y = z) = P(X=0;Y=z)+P(X=1;Y=z-1)+...+P(X=z;Y=0) = \sum_{j=0}^{z}\frac{e^{-\lambda_1}\lambda_1^j}{j!}\frac{e^{-\lambda_2}\lambda_2^{z-j}}{(z-j)!}$

$= \frac{1}{z!}e^{-\lambda_1}e^{-\lambda_2}\sum_{j=0}^{z}\frac{z!}{j!(z-j)!}\lambda_1^j\lambda_2^{z-j} = \frac{1}{z!}e^{-(\lambda_1+\lambda_2)}\sum_{j=0}^{z}\binom{z}{j}\lambda_1^j\lambda_2^{z-j} = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^z}{z!}$, 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 $\mathcal{B}(n;\, p)$	$x \in \{0, 1, \ldots, n\}$	$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$	np	$np(1-p)$	$(q + pe^t)^n$
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 $\mathcal{B}^{-}(p)$	$x \in \{n, n+1, \ldots\}$	$p(x) =$ $\binom{x-1}{n-1} p^n (1-p)^{x-n}$	$\dfrac{n}{p}$	$\dfrac{n(1-p)}{p^2}$	$\left(\dfrac{pe^t}{1-(1-p)e^t}\right)^n$