UNIVERSITE de ROUEN - S.C.U.R.I.F.F.

INSTITUT de RECHERCHE
sur
l' ENSEIGNEMENT des MATHEMATIQUES

IREM

tél : 35 14 61 41

LES PROBABILITES pour le Lycée 2

IREM de ROUEN

1, rue Thomas Becket - BP 153, 76135 Mont-Saint-Aignan Cédex

SOMMAIRE

III LA LOI DE POISSON

IV LOI DES ATTENTES

A ATTENTE DU PREMIER SUCCÈS

B ATTENTES PLUS LOINTAINES

INTRODUCTION

Ce deuxième fascicule, reprend les travaux du stage "Probabilités" qui s'est déroulé en Octobre 1993.

Il m'a semblé qu'il pouvait être profitable aux enseignants de Lycée, qui souvent, comme moi, n'avaient pas rencontré les probabilités pendant leur cursus universitaire avant la maîtrise, de dépasser les problèmes de Terminale, pour voir comment la théorie se développait ensuite.

Pourquoi?

D'abord, parce que les exercices de Terminale sont souvent des extraits ou des simplifications de modélisations discrètes générales, ensuite parce que considérer les mathématiques avec un certain recul, n'a jamais nui.

Les stagiaires, qui, je crois, ont suivi ce cours avec beaucoup d'attention et de bonne humeur, ont, pour certains découvert, pour d'autres retrouvé, des modélisations discrètes infinies qui ne sont jamais triviales.

Mais nous avons tous remarqué, qu' il vient très vite, dans ce domaine, même avec des problèmes simples, dans le cas fini (la collection d'un groupe d'objets par exemple), la nécessité de posséder des outils de l'analyse assez sophistiqués.
Les sommes de Riemann, les techniques de sommation, de majoration et de minoration, les suites géométriques arrivent ici très tôt.

Bien sûr, les probabilités discrètes nécessitent aussi l'introduction des séries, mais je n'ai pas, sauf exception (pour la formule de Stirling, en appendice) voulu utiliser l'analyse à un niveau dépassant celui de la classe de Terminale C.

Pour tout sujet existe un livre de référence. Pour les probabilités, c'est sans nul doute dans le livre de Feller[1] , qu'on trouve presque tout. Pour ceux qui n'aiment pas lire l'anglais, il existe chez Cedic, deux volumes qui reprennent un ouvrage d'Arthur Engel [2], hélas épuisés, particulièrement intéressants. J'ai tiré de nombreux exemples dans ces deux livres, mais il en reste beaucoup d'autres, qu'un lecteur sagace aimera chercher.

J'ai mis, pour finir, en appendice, un résumé d'un exposé qu'à fait un stagiaire, à propos des tableaux et des probabilités conditionnelles. Bien que ce sujet eût été mieux à sa place dans le premier fascicule, il m'a semblé profitable de l'insérer ici.
Jean Claude Jovet nous a en effet montré comment il utilisait pédagogiquement la symbolique des arbres, particulièrement dans des classes où les élèves sont souvent intéressés par les probabilités sans posséder ni la science, ni la technique mathématique suffisante.

Je remercie, enfin, une fois de plus, Thierry Hamel qui a bien voulu relire ce volume et lui apporter de nombreux commentaires.

---

[1] An Introduction to Probability Theory and its Applications. William Feller (John Wiley).

[2] L'enseignement des Probabilités et de la Statistique. Arthur Engel. (Cedic 1979).

I GENERALITES SUR LES VARIABLES ALEATOIRES DISCRETES.

1 INTRODUCTION

On suppose que (Ω,B,P) est un espace probabilisé.
Si Ω est fini on prendra la tribu $B=\mathcal{P}(\Omega)$.

Une variable aléatoire X est un caractère statistique "du futur", tel que les sous populations $\{\omega\in\Omega\ ,\ X(\omega)=x_i\}$ supportent la mesure par la probabilité P.

DEFINITION:

Soit X une application de Ω dans $\mathbb{R}$, telle que l'ensemble $X(\Omega) = \{\ x_i\ \}_{i\in\ I}$ soit fini ou dénombrable. On dit que X est une variable aléatoire discrète si et ssi pour tout $i\in I$ $\{\omega\in\Omega\ ,\ X(\omega)=x_i\} = [X=x_i]$ reste dans la tribu B.

REMARQUE: la condition est toujours vérifiée si Ω est fini.

On supposera désormais que $i<j$ entraîne $x_i<x_j$.

DEFINITION:

L'application qui à tout élément x de $X(\Omega)$ associe le réel $P[X=x]$ s'appelle la loi de probabilité de X. (ou distribution). [3]

Proposition:

La donnée d'une variable aléatoire sur Ω équivaut à la donnée d'une suite $(p_i)_{i\in I}$ de nombres <u>positifs</u> tels que $\sum_{i\in I} p_i= 1$.

Pour montrer le sens direct il suffit de remarquer que

$$\sum_{i\in I} p_i= \sum_{i\in I} P[X=x_i] = P(\bigcup_{i\in I} [X=x_i]) = P(\Omega) = 1.$$

Réciproquement, nous admettrons que pour toute suite $\{\ p_i\ \}_{i\in I}$, vérifiant les conditions précédentes peut être considérée comme la loi d'une variable aléatoire X.

---

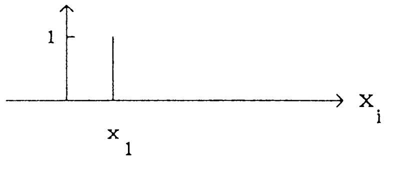
[3] La loi de X est donc la probabilité image définie sur $\{\ x_i\ \}_{i\in I}$ de la probabilité P par l'application $w \longrightarrow X(\omega)$.

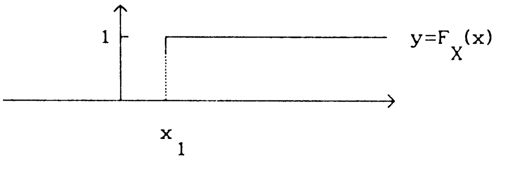
DEFINITION:

On appelle fonction de répartition de la variable aléatoire X, et l'on note F_X, l'application définie sur $\mathbb{R}$ par $F_X(x) = P[X\leq x] = \sum_{x_i \leq x} p_i$.

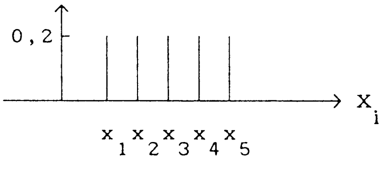
Exemples:

X est constante si et ssi elle ne prend qu'une seule valeur, ou quasi constante si elle prend une valeur avec la probabilité 1. Dans les deux cas on a la distribution et la fonction de répartition suivantes:





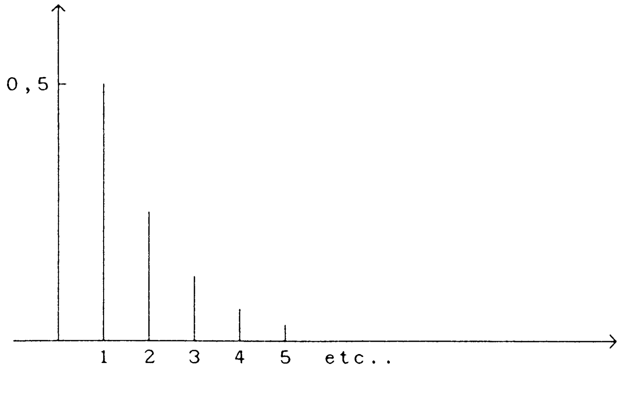
X suit une loi uniforme sur {1,..,n} si et ssi la suite des (p_i) est constante et donc égale à 1/n. D'où la distribution "en peigne":



Exercice: montrer que la fonction $f(i)= 1/2^i$ définie la loi de probabilité d'une variable aléatoire X. (sur $\mathbb{N}^*$).

Il suffit de calculer la somme $\sum_{i=1}^{+\infty} \frac{1}{2^i} = \frac{1}{2} \frac{1}{1-\frac{1}{2}} = 1.$

On a alors la distribution suivante:



REMARQUE: la fonction F est croissante et continue à droite en tout point de $\mathbb{R}$.

2 COMPARAISON AVEC LES STATISTIQUES.

Il est facile dans le cas où Ω est fini de comparer les notions introduites en probabilités avec, celles, beaucoup plus simples, qui relèvent des classes antérieures et des statistiques. On peut par exemple choisir la moyenne de maths des élèves de la classe de l'an passé, ce qui fait une belle série statistique et la mettre en relation avec celle que ces mêmes élèves obtiendront en Juin, ce qui définit une variable aléatoire.

STATISTIQUES	PROBABILITES
Passé	*Futur*
population Ω (finie). effectif n.	population Ω (finie).
caractère statistique.	X variable aléatoire.
$\{x_1,x_2,\dots,x_p\}$ valeurs du caractère.	$X(\Omega) = \{x_1,x_2,\dots,x_p\}$
la sous population attachée à la valeur x_i du caractère est $[\mathcal{X}=x_i]$	L'évènement $[X=x_i]$
son effectif est n_i	a pour probabilité
sa fréquence est $f_i = \dfrac{n_i}{n}$	$p_i=P[X=x_i]$
la fréquence cumulée croissante est $f_i^c = \sum\limits_{j\leq i} f_j$	la Fonction de répartition est $F_X(x_i)=P[X\leq x_i]$
la moyenne $\bar{x}$ est [4] $\bar{x} = \dfrac{1}{n}\sum\limits_{i\in I} x_i n_i = \sum\limits_{i\in I} x_i f_i$	l'espérance E(X) est $E(X) = \sum\limits_{i\in I} x_i p_i$
les moments d'ordre k ($k\in\mathbb{N}^*$). $m_k = \dfrac{1}{n}\sum\limits_{i\in I} x_i^k n_i = \sum\limits_{i\in I} x_i^k f_i$	les moments d'ordre k ($k\in\mathbb{N}^*$). $m_k = \sum\limits_{i\in I} x_i^k p_i$
les moments centrés d'ordre k. $\mu_k = \dfrac{1}{n}\sum\limits_{i\in I} (x_i-\bar{x})^k n_i = \sum\limits_{i\in I} (x_i-\bar{x})^k f_i$	les moments centrés d'ordre k. $\mu_k = \sum\limits_{i\in I} (x_i-\bar{x})^k p_i$

μ_2 s'appelle la variance notée Var.

---

4 La première formule est meilleure pour faire des calculs, car les erreurs d'arrondis sont minimisées (la division vient à la fin). Par contre, pour la théorie et le passage statistiques-probabilités, c'est la seconde qui est la plus intéressante.

REMARQUES:

1) pour définir la variance, en statistiques, on peut diriger les élèves vers les écarts. Après avoir constaté, généralement avec surprise, que la moyenne des écarts à la moyenne est toujours nulle, on peut tenter la moyenne des valeurs absolues des écarts à la moyenne que l'on appelle écart moyen. On préfère prendre la moyenne des carrés des écarts à la moyenne, car la fonction $x \longrightarrow x^2$ écrase les petits écarts et accroît les grands.

La quantité obtenue n'est toutefois pas homogène à la série de départ, d'où l'idée d'introduire l'écart type [5]:

$$\sigma = \sqrt{\mu_2}$$

2) Var(X) = 0 si et ssi la variable aléatoire est quasi constante.

En effet Var(X) est nul si et ssi tous les termes de la somme $\sum_{i \in I} (x_i - \bar{x})^2 p_i$ sont nuls (car ≥0). Comme un des p_i est forcément non nul, cela signifie que tous les p_i sont nuls sauf un qui vaut 1.

Dans ce cas on voit aussi facilement que si X est quasi constante égale à a,

$$E(X) = a.$$

3) Pour définir les moments pour une variable discrète X définie sur Ω quelconque, il suffit de garder les mêmes définitions mais de demander en plus la <u>convergence absolue</u> des séries concernées.

4) sous ces réserves

$$m_k = \sum_{i \in I} x_i^k p_i = E(X^k)$$

et

$$\mu_k = \sum_{i \in I} (x_i - \bar{x})^k p_i = E([X - E(X)]^k)$$

Exercice: calculons l'espérance et la variance de la variable uniforme sur {1,2,...,n}.

Pour tout i, $p_i = 1/n$

---

5 Comme l'a remarqué un stagiaire de Bernay l'écart absolu moyen est toujours inférieur à l'écart type, puisque la fonction $x \longrightarrow x^2$ est convexe, le domaine $\{M(x,y)/\ y > x^2\}$ est convexe et donc la moyenne des carrés est toujours supérieure au carré de la moyenne.

$$E(X) = \sum_{i=1}^{n} i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} i = \frac{n+1}{2}$$

$$\mu_2 = E([X-E(X)]^2) = \sum_{i \in I} (x_i - \bar{x})^2 p_i = \sum_{i=1}^{n} (i - \frac{n+1}{2})^2 \frac{1}{n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} (i^2 + \frac{(n+1)^2}{4} - i(n+1)) = \frac{1}{n} \sum_{i=1}^{n} i^2 + \frac{(n+1)^2}{4} - \frac{1}{n}(n+1) \sum_{i=1}^{n} i$$

$$= \frac{(n+1)(2n+1)}{6} + \frac{(n+1)^2}{4} - \frac{(n+1)^2}{2}$$

$$= \frac{n^2-1}{12}$$

3 PARENTHÈSE SUR LES MOYENNES.

Nous venons déjà (en statistiques comme en probabilités) de voir plusieurs moyennes. Par exemple, l'écart type σ est, en fait, la moyenne quadratique, rendue à nouveau homogène par la fonction réciproque de $x \longrightarrow x^2$, la racine.

On peut ainsi généraliser la notion de moyenne au travers d'un "filtre" qui est introduit en fonction de l'utilité ou du sens statistique, ici nous prendrons une fonction arbitraire φ.

> Soit φ une bijection d'un intervalle I de $\mathbb{R}$ sur un intervalle J ($I \subseteq J$).
> Soit X une variable aléatoire (ou statistique....) prenant ses valeurs (finies[6]) dans I. On appelle moyenne, modulo φ le réel
>
> $$m_\varphi = \varphi^{-1}(E(\varphi(X))).$$

Remarquons d'abord qu'avec nos hypothèses l'espérance de $\varphi(X)$ est définie puisque X prend un nombre fini de valeurs. De plus une moyenne faite avec des éléments de J reste dans J puisque cet intervalle est convexe et donc on peut lui appliquer φ^{-1}.

L'idée est de faire la moyenne ordinaire des éléments vus au travers du filtre φ puis ensuite de revenir en arrière avec φ^{-1}.

Si φ est l'identité on retrouve $m_{Id} = E(X)$

Si φ est $\begin{array}{ccc} [0,+\infty[& & [0,+\infty[\\ x & \longrightarrow & x^2 \end{array}$ on retrouve la moyenne quadratique $m_2^{1/2}$

---

6 on pourrait facilement généraliser au cas discret en envisageant la convergence absolue des séries, mais ce serait compliquer inutilement.

Si φ est $x \xrightarrow{\quad} x^k$ ($[0,+\infty[\to [0,+\infty[$) on retrouve $m_k^{1/k}$.

Si φ est $x \xrightarrow{\quad} \ln x$ ($]0,+\infty[\to]-\infty,+\infty[$) on trouve la moyenne géométrique.

$$m_g = \exp\Big(\sum_{i\in I} \ln(x_i) p_i\Big) = \exp\Big(\sum_{i\in I} \ln(x_i^{p_i})\Big)$$

$$m_g = \prod_{i\in I} x_i^{p_i} = \Big(\prod_{i\in I} x_i^{n_i}\Big)^{1/n}$$

Si φ est $x \xrightarrow{\quad} 1/x$ ($]0,+\infty[\to]0,+\infty[$) on trouve la moyenne harmonique m_h .

$$m_h = \varphi^{-1}(E(\varphi(X)).$$

donc

$$\varphi(m_\varphi) = (E(\varphi(X)).$$

$$\frac{1}{m_h} = \sum_{i\in I} \frac{1}{x_i} p_i$$

ou

$$\frac{n}{m_h} = \sum_{i\in I} \frac{n_i}{x_i}$$

Cas particulier des fonctions φ concaves (ou convexes):.

Prop:("petite" inégalité de Jensen)

Si φ est convexe sur I, on a $\varphi(E(X)) \le E(\varphi(X))$

En effet si φ est convexe le domaine $\{M(x,y)/\ y>\varphi(x)\}$ est convexe et donc la moyenne des images par φ est toujours supérieure à l'image de la moyenne. En particulier si φ est strictement croissante on déduit, puisque la réciproque φ^{-1} l'est aussi, que

$$E(X) \le \varphi^{-1}(E(\varphi(X))) \text{ dans le cas convexe etc...}$$

On retrouve ainsi que pour tout entier k:

$$m_1 \le m_k^{1/k}$$
$$m_g \le m_1$$
$$m_h \le m_1$$

Mais on peut aussi arriver à comparer des moyennes filtrées ensemble. par exemple m_g et m_h.

$$m_g = \varphi^{-1}(E(\varphi(X))) \qquad \text{avec } \varphi \text{ le logarithme.}$$

$$m_h = \psi^{-1}(E(\psi(X))) \qquad \text{avec } \psi \text{ le passage à l'inverse.}$$

Prenons comme nouvelle variable aléatoire $Y = \varphi(X)$

$$m_g = \varphi^{-1}(E(Y)) \qquad m_h = \psi^{-1}(E(\psi(\varphi^{-1}(Y))))$$

Or la fonction $\psi o \varphi^{-1} : x \longrightarrow e^{-x}$ est décroissante et convexe sur $\mathbb{R}$.

On peut ainsi écrire que $E(Y)) \geq (\psi o \varphi^{-1})^{-1}(E(\psi(\varphi^{-1}(Y)))$

$$E(Y)) \geq (\varphi o \psi^{-1})(E((\psi o \varphi^{-1})(Y)))$$

et en composant par φ^{-1} de chaque côté

$$m_g = \varphi^{-1}(E(Y)) \geq \psi^{-1}(E(\psi(\varphi^{-1}(Y))) = m_h$$

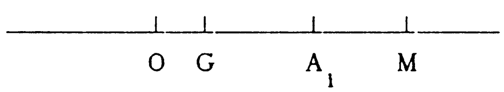
On trouve ainsi que

$$m_h \leq m_g \leq m_1 \leq m_k^{1/k}$$

4 THEOREME DE KOENIG-HUYGHENS.

Il s'agit, bien sûr, d'un théorème de mécanique, que l'on applique aux probabilités. On peut aussi le relier aux lignes de niveau et donc à la géométrie.

$$Var(X) = \sum_{i \in I} (x_i - E(X))^2 p_i = \sum_{i \in I} GA_i^2 \, p_i \quad \text{donc une fonction de "Leibniz"}$$



$$\overline{OA_i} = x_i \qquad \overline{OG} = E(X)$$

Nous appelons G le barycentre du système (A_i, p_i)

Pour tout point M de la droite

$$\sum_{i \in I} MA_i^2 \, p_i = \sum_{i \in I} (\overrightarrow{MG} + \overrightarrow{GA_i})^2 p_i = \sum_{i \in I} p_i \, \overrightarrow{MG}^2 + 2 \sum_{i \in I} \overrightarrow{MG}.\overrightarrow{GA_i} p_i + \sum_{i \in I} \overrightarrow{GA_i}^2 p_i$$

Comme le second terme est nul

$$\sum_{i \in I} MA_i^2 \, p_i = MG^2 + \sum_{i \in I} \overrightarrow{GA_i}^2 p_i = MG^2 + Var(X).$$ [7]

On en déduit que la dispersion est minimale quand on la calcule par rapport au point moyen.

De plus quand on prend M à l'origine du repère $\sum_{i \in I} OA_i^2 \, p_i$ représente m_2 et OG la

---

[7] Ce qui signifie que le moment d'inertie autour de tout point M est la somme du moment d'inertie autour du centre de gravité G, augmenté de la distance MG^2 (affectée de la masse totale qui fait 1).

moyenne, donc:

$$\mathrm{Var}(X) = m_2 - m_1^2 = E(X^2) - E^2(X).$$

On pourra d'ailleurs redémontrer cette formule, une fois le formalisme des probabilités, un peu avancé.

Proposition: Si elle existe

$$E(X) = \sum_{\omega\in\Omega} X(\omega)\ P(\omega).$$

La démonstration, dans le cas général, va faire appel à la notion de sommabilité qui dépasse le cadre de ce stage. On peut toutefois la rédiger dans le cas où Ω est un ensemble fini.

$$\text{On a } p_i = P[\ \omega\in\Omega\ /\ X(\omega) = x_i] = \sum_{\omega\in X^{-1}(x_i)} P(\omega).$$

$$\text{et donc} \qquad p_i x_i = \sum_{\omega\in X^{-1}(x_i)} P(\omega)X(\omega).$$

$$\sum_{i\in I} x_i p_i = \sum_{i\in I}\ \sum_{\omega\in X^{-1}(x_i)} P(\omega)X(\omega) = \sum_{\omega\in\Omega} X(\omega)\ P(\omega).$$

Car les $X^{-1}(x_i)$ forment une partition de Ω.

Corollaire: Si X et Y possède des espérances finies alors X+Y a une espérance finie et $E(X+Y) = E(X)+E(Y)$.

On se contente, ici encore, de la démonstration dans le cas fini.

$$E(X)+E(Y) = \sum_{\omega\in\Omega} X(\omega)\ P(\omega) + \sum_{\omega\in\Omega} Y(\omega)\ P(\omega) = \sum_{\omega\in\Omega} (X+Y)(\omega)\ P(\omega) = E(X+Y)$$

REMARQUE: on a facilement $E(\lambda X) = \lambda E(X)$.

En effet si la série $\sum_{i\in I} x_i p_i$ est absolument convergente, pour tout réel λ, il en va de même pour la série $\sum_{i\in I} \lambda x_i p_i$.
Comme la variable λX prend les valeurs λx_i avec les probabilités p_i.
On a $E(\lambda X) = \sum_{i\in I} \lambda x_i p_i = \lambda \sum_{i\in I} x_i p_i = \lambda E(X)$.

Exercice: On peut retrouver la formule de Huyghens, en utilisant le nouveau formalisme.

$$E(\ (X - E(X))^2\ ) = E(\ X^2 - 2XE(X) + E^2(X)\ ) = E(X^2) - 2E(XE(X)) + E(E^2(X))$$

Comme E(X) est un réel constant on retrouve

$$Var(X) = E(X^2) - 2E(X)E(X) + E^2(X) = E(X^2) - E^2(X).$$

5 INDÉPENDANCE

DEFINITION: Deux variables X et Y sont dites "indépendantes" si et ssi quelles que soient les valeurs x_i et y_j de $X(\Omega)$ et $Y(\Omega)$, les événements $[X=x_i]$ et $[Y=y_j]$ sont indépendants.

DEFINITION: Les variables aléatoires $(X_i)_{i \in I}$ sont dites "mutuellement indépendantes", si et ssi pour toute sous famille J de I et toute suite x_{i_j} de $X_i(\Omega)$, les événements ($[X_i = x_{i_j}]$, $j \in J$) sont mutuellement indépendants.

Dans le second stage nous n'avons pas étudié en détails ces notions, et nous nous sommes contentés d'admettre le théorème suivant:

THEOREME: Soit $(X_i)_{i \in I}$ une suite de variables aléatoires deux à deux indépendantes

Alors $$Var(\sum_{i \in I} X_i) = \sum_{i \in I} Var(X_i)$$

II LA LOI BINOMIALE.

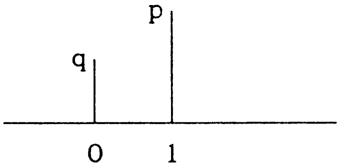
1 LOI DE BERNOULLI

On appelle loi de Bernoulli une variable aléatoire prenant deux valeurs.
La plus simple est donc X telle que $X(\Omega) = \{0, 1\}$.

On écrit $X \simeq \mathcal{B}(p)$

p étant le paramètre défini par les égalités:

$P[X=0]=q \quad P[X=1]=p$ avec p, q positifs et $p+q=1$.



REMARQUE1: quel que soit le réel α $\quad X^{\alpha} = X$.

On a facilement $E(X) = p$.

Compte tenu de la remarque précédente, c'est aussi le moment d'ordre 2.

donc $\quad Var(X)=E(X^2) - E^2(X) = p - p^2 = pq$.

REMARQUE2: Pour une variable de Bernoulli générale on a $Y(\Omega) = \{a, b\}$[8].

et, bien sûr, l'espérance est directement $E(Y) = qa + pb$, résultat que l'on retrouve en écrivant, que $\dfrac{Y-a}{b-a} \simeq \mathcal{B}(p).(b \neq a)$

$$p = E(X) = E\left(\frac{Y-a}{b-a}\right) = \frac{1}{b-a}\big(E(Y)-a\big)$$

On a aussi $pq = Var(X) = Var\left(\dfrac{Y-a}{b-a}\right) = \left(\dfrac{1}{b-a}\right)^2 Var(Y-a) = \left(\dfrac{1}{b-a}\right)^2 Var(Y)$

Donc $Var(Y) = (b-a)^2 pq$.

2 LOI BINÔMIALE.

On compte le nombre de réussites, dans ce qu'on appelle, en Terminale, le schéma de Bernoulli, suite de n épreuves indépendantes à deux issues, 0 ou 1.

On peut appeler X_i le résultat de la ième épreuve.

Le nombre total de réussites est $X = \sum_{i=1}^{n} X_i$.

---

[8] Il est fréquent de prendre a=-1 et b=1.

DEFINITION: On appelle loi binômiale de paramètres n et p et l'on note $\mathcal{B}(n,p)$ la loi d'une variable aléatoire qui est la somme de n variables de Bernoulli <u>indépendantes</u> de loi $\mathcal{B}(p)$.

On a donc alors $X(\Omega) = \{0,1,\dots,n\}$.

Proposition: la loi de $X \simeq \mathcal{B}(n,p)$ sur $X(\Omega) = \{0,1,\dots,n\}$ est donnée par

$$P[X=k] = C_n^k \; p^k q^{n-k} \qquad \text{si } k \in X(\Omega)$$

Démonstration:

Soit S_k l'événement "obtenir k succès".

s_k l'événement "obtenir le succès au kième tirage".

σ_k l'événement "obtenir k succès dans un ordre donné".

e_k l'événement "obtenir l'échec au kième tirage".

Quelque soit l'ordre que l'on s'est fixé [9]

$$P(\sigma_k) = P[s_1 \cap s_2 \cap \dots \cap s_k \cap e_{k+1} \cap \dots \cap e_n] = P(s_1)P(s_2)\dots P(s_k)P(e_{k+1})\dots P(e_n)$$

Puisque les événements sont mutuellement indépendants.

$$P(\sigma_k) = p^k q^{n-k}$$

Comme S_k est l'union disjointe des événements σ_k quand tous les ordres varient, il nous suffit de compter les anagrammes du mot SSSSSEEEEE pour conclure. Il y en a $\dfrac{n!}{p!(n-p)!} = C_n^p$.

$$\text{Ainsi } P(S_k) = C_n^k \; p^k q^{n-k}.$$

REMARQUE: on a en corollaire, puisque la somme de n variables aléatoires est une variable aléatoire que, $\sum_{k=0}^{n} C_n^k \; p^k q^{n-k} = \sum_{k=0}^{n} P[X=k] = 1$, ce qui est un cas particulier de la formule du binôme de Newton.

COMMENTAIRES:

Dans les anciens programmes de Terminale, la loi binômiale était la seule loi au programme et donc la seule présente dans les problèmes de Bac. La reconnaître revenait, dans ces conditions, à deviner la présence du schéma de Bernoulli, ce qui, me semble-t-il, est très peu formateur. En effet, si l'on élève un enfant en ne lui montrant que des exemples de poissons plats, on aura bien du mal à lui faire deviner qu'il peut tenir au bout de sa canne à pêche, un saumon ou un cabillaud.

Le schéma de Bernoulli contient des hypothèses précises qui sont toutes capitales pour la loi.

---

[9] Par exemple on peut fixer les k succès au début, puis faire les anagrammes du mot $\underbrace{\text{SSSSSSS}}_{k}\underbrace{\text{EEEEEEEEEEE}}_{n-k}$.

Exercice 1: On opère cinq tirages, <u>sans remise</u>, d'une boule dans une urne qui contient 10 boules numérotées de 1 à 10.

On appelle succès l'obtention d'une boule supérieure strictement à 5.
Soit X_i la variable de Bernoulli qui vaut 1 en cas de succès au ième tirage.

Calculons son paramètre.

$$P[X_i=1] = \frac{\overset{\downarrow}{5}\, A_9^4}{A_{10}^5}$$

les cinq boules favorables de la ième place

← les quatre autres boules.

$$P[X_i=1] = \frac{1}{2}$$

Ainsi X le nombre de succès est la somme de 5 variables de Bernoulli de paramètre 1/2 et de plus $X(\Omega) = \{0,1,\ldots,5\}$, mais X ne suit pas une loi binômiale car il n'y a pas d'indépendance.

$$\text{On a } P[X=k] = \frac{C_5^k\ C_5^{10-k}}{C_{10}^5}$$ [10]

Exercice 2: Maintenant, dans la même suite de tirages sans remise, dans la même urne, on appelle succès le fait qu'au ième tirage le plus grand des numéros tirés soit à la ième place. Par exemple 427 donne un succès au premier et au troisième tirage. On appelle Y_i la variable de Bernoulli qui compte le succès au ième tirage.

Prenons un modèle de i tirages sans remise. Une fois les i numéros connus il y a (i-1)! possibilités de les répartir, le plus grand restant à la dernière place. Donc $(i-1)!C_{10}^i$

$$P[Y_i=1] = \frac{(i-1)!C_{10}^i}{A_{10}^i}$$

← choix des i premières boules.

$$P[Y_i=1] = \frac{1}{i}$$

Y le nombre total de succès (ie le nombre de fois ou le plus grand des numéros tirés se trouve à la dernière place) est la somme des variables X_i. De plus Y prend des valeurs entières allant de 1 pour le tirage 10 9 8 7 6.. 2 1, à 10 pour le tirage 1 2 ....8 9 10.

Y est la somme de variables de Bernoulli de paramètre différents 1/i, et non indépendantes.

---

[10] cf plus loin loi hypergéométrique.

Il faut donc bien faire remarquer l'importance des hypothèses, tout comme celle de la nature de la variable recherchée [10] . Si cette loi a été introduite dans le second cycle, c'est sans doute d'abord à cause du rôle primordial qu'elle joue dans les sciences de la nature.... comme le montrent les deux exercices suivants:

Exercice 3: Une épidémie atteint 25% d'un élevage. On choisit 3 échantillons comportant, pour le premiers 10 éléments, le second 17 et le troisième 23 sur lesquels on teste trois nouveaux sérums a, b, c.

Les résultats sont

	a	b	c
X	0	1	2

X nbre de malades dans l'échantillon.

Les sérums sont ils efficaces (au seuil des 5%)? [11]

Pour répondre à la question, cherchons la probabilité d'obtenir ces mêmes résultats sans intervention des sérums.

X_a, X_b, X_c le nombre de malades dans chaque échantillon suivent des lois binômiales $X_a \simeq \mathcal{B}(10,\frac{1}{4})$, $X_b \simeq \mathcal{B}(17,\frac{1}{4})$ et $X_c \simeq \mathcal{B}(23,\frac{1}{4})$.

Donc $P[X_a=0] \cong 0{,}05631$.

$P[X_b \leq 1] \cong 0{,}05011$.

$P[X_c \leq 2] \cong 0{,}04920$.

Conclusion le seul sérum efficace au seuil des 5% est le c.

Exercice 4: Dans un jury de 25 membres, cinq sont de parti pris.
Quelle est la probabilité que le jugement soit dans le sens des cinq qui sont de mauvaise foi, si les autres membres se déterminent au hasard.
Soit X le nombre de personnes sans parti pris allant dans le même sens que les cinq.

Soit J l'événement: "le jugement va dans le sens des cinq".

$P(J) = P[X \geq 8] \cong 0{,}868$ puisque $X \simeq \mathcal{B}(18,\frac{1}{2})$.

Exercice 5: Est il plus facile d'obtenir au moins un as avec 6 dés ou deux as avec 12?

"Obtenir un as avec 6 dés" (resp 12) est noté A (resp B).

---

[11] Ainsi, même avec schéma de Bernoulli (exemple succession de pile-face), je peux m'intéresser à tout autre chose que le nombre de pile. Le rang du premier triple PPP par exemple ou tout autre chose.

[12] Une hypothèse H est dite réalisée au seuil des 5% si et ssi la probabilité de $\bar{H}$ est inférieure à 0,05.

$$P(\bar{A}) = \frac{5^6}{6^6}$$

$$P(A) = 1 - \frac{5^6}{6^6} \cong 0,665$$

$P(\bar{B}) = P[Z \le 1]$ $\qquad$ $Z \simeq \mathcal{B}(12, 1/6)$ $\qquad$ (Z nbre d'as)

$$P(B) = 1 - \frac{5^{12}}{6^{12}} - \frac{12 x 5^{11}}{6^{12}} = 0,617$$

$0 < P(A)-P(B) < 0,047$

C'est donc A qui est le plus probable, mais la différence est piètre.

C'est comprendre le coefficient C_n^k, dans l'étude de la loi Binômiale, qui est important et pas savoir reconnaître le schéma de Bernoulli, contrairement à ce qui va "être payant" pour le Bac.

Pour vérifier cette compréhension, rien n'est meilleur que d'observer l'attitude des élèves devant une généralisation: la loi trinômiale.

3 LOI TRINÔMIALE

On fait une succession de n épreuves, à trois issues, indépendantes de probabilités p, q, r pour les issues A, B, C. (p+q+r=1). [13]

Calculons la probabilité d'obtenir i résultats A et j B. $(i+j \le n)$
Le raisonnement est <u>le même</u> que pour la loi binômiale.
La probabilité d'obtenir dans un ordre donné les i A et les j B est de

$p^i q^j r^{n-i-j}$ (et donc ne dépend pas de cet ordre).

Pour trouver la probabilité finale il suffit de trouver le nombre d'ordres possibles, c'est à dire le nombre d'anagrammes du mot formé avec i fois la lettre A et j fois B et n-i-j fois C. On en a $\frac{n!}{i!j!(n-i-j)!}$

Donc si on appelle X, Y, Z les nombres de A, B, C sortis on a:

$$P[X=i \cap Y=j] = \frac{n!}{i!j!(n-i-j)!} p^i q^j r^{n-i-j}$$

[13] Il est facile de présenter ceci par des tirages de boules de trois couleurs avec remise, ou des choix indépendants à opérer sur 3 produits etc...

On peut aussi ne s'intéresser qu'au nombre de A. Ces trois issues nous donneront donc à nouveau une loi binômiale en "confondant" les B et les C.

$X \simeq \mathcal{B}(n,p)$

$$P[X=k] = C_n^k \; p^k(q+r)^{n-k}$$

Exercice: On peut retrouver cette loi trinômiale en utilisant les probabilités conditionnelles:

Supposons qu'à la sortie d'une station de métro, se trouvent trois correspondances. On appelle X_1, X_2, X_3 les nombres de personnes allant choisir chaque issue parmi un flot de n personnes descendant de la rame. (les choix étant considérés comme indépendants).

1°) Quelles sont les lois de X_1, et X_1+X_2?

2°) Trouver trois méthodes pour calculer $P[X_1=i \cap X_2=j]$.

1°) X_1 suit donc $\mathcal{B}(n,1/3)$ $\quad X_1+X_2 \simeq \mathcal{B}(n,2/3)$

2°) ✱ méthode "conditionnelle":

$$P[X_1=i \cap X_2=j] = P[X_2=j \;/X_1=i]P[X_1=i]$$

Or si l' connaît le flot des personnes ayant choisi la première correspondance les n-i autres se répartissent équitablement donc

$$X_2 \;/[X_1=i] \simeq B(n-i,1/2)$$

donc $P[X_1=i] = C_n^i \; (\frac{1}{3})^i(\frac{2}{3})^{n-i}$ et

$$P[X_2=j \;/X_1=i] = C_{n-i}^j \; (\frac{1}{2})^j(\frac{1}{2})^{n-i-j} = C_{n-i}^j \; (\frac{1}{2})^{n-i}$$

$$P[X_1=i \cap X_2=j] = C_n^i(\frac{1}{3})^i(\frac{2}{3})^{n-i}C_{n-i}^j \; (\frac{1}{2})^{n-i} = C_n^iC_{n-i}^j \; (\frac{1}{3})^n$$

✱✱ méthode multinômiale (cf ci dessus). On compte les mots avec i A et j B:

$$P[X_1=i \cap X_2=j] = \frac{n!}{i!j!(n-i-j)!}(\frac{1}{3})^n$$

✱✱✱ méthode multinômiale modifiée:

Pour un ordre donné du flot des sortants, avec i choix en A et j en B on a une probabilité de $(\frac{1}{3})^n$.

Pour compter les configurations on calcule les places possibles des A: C_n^i puis ensuite les places des B: C_{n-i}^j [14] . Donc en tout $C_n^iC_{n-i}^j \; (\frac{1}{3})^n$.

---

[14] C'est ici que l'on peut constater si les élèves ont bien compris. Dans le ca

Un problème de Générations:

On admet que la loi du nombre d'enfants d'un individu quelconque d'une population est constant à travers les âges et est une variable aléatoire de loi binômiale $\mathcal{B}(a,\frac{1}{2})$, a désignant un paramètre entier fixé. On suppose que les descendances de chaque personne sont indépendantes.

Pour $n \in \mathbb{N}^*$, p_n désigne la probabilité pour un individu de ne pas avoir de descendants à la $n^{\text{ième}}$ génération. (on pose $p_0=0$).

1°) Calculer p_1.

2°) Trouver une relation de récurrence entre p_n et p_{n+1}.

3°) En déduire l'étude de $(p_n)_{n\geq 1}$.

1°) Appelons I notre premier individu et A_1 l'événement "I est sans enfant".

$$p_1 = P[A_1] = \left(\frac{1}{2}\right)^a$$

2°) Soit A_n l'événement "I est sans postérité à la nième génération".

Supposons que X le nombre d'enfants de I soit k.

A_{n+1} est donc l'intersection des événements suivants,"les k enfants n'ont eux mêmes pas de descendants à la nième génération". Puisque ces k descendances sont indépendantes et suivent la même loi on peut écrire

$$P[A_{n+1}/X=k] = P[A_n]^k = p_n^k$$

En appliquant la formule des probabilités totales au système complet [X=0], ..[X=n]:

$$p_{n+1} = P[A_{n+1}] = \sum_{k=0}^{n} P[A_{n+1}/X=k]\ P[X=k] = \sum_{k=0}^{n} p_n^k \left(\frac{1}{2}\right)^a C_n^k$$

Ainsi $p_{n+1} = \left(\dfrac{p_n + 1}{2}\right)^a$ avec $p_0 = 0$ [15].

Pour étudier cette suite récurrente, on trace la représentation graphique de la fonction f_a définie sur [0,1] par $f_a(x) = \left(\dfrac{x + 1}{2}\right)^a$

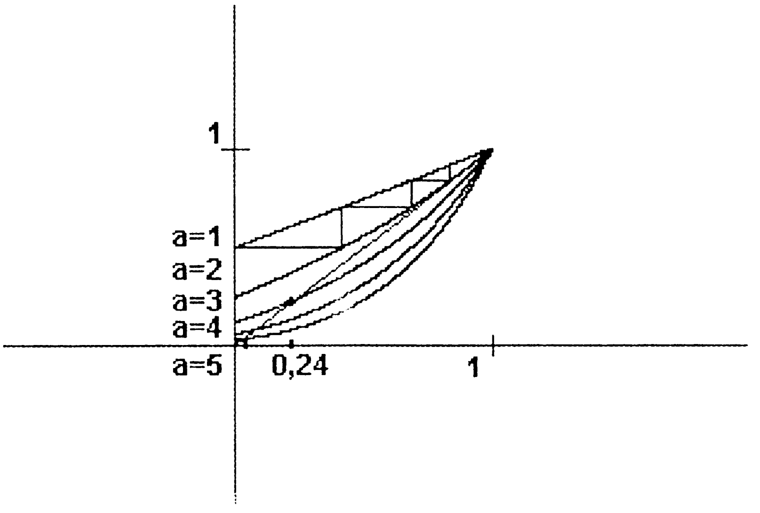
---

contraire ils reprendraient C_n^j!

[15] Puisqu'on vérifie qu'avec cette hypothèse gratuite on retrouve bien p_1.

elle dépend du paramètre entier a.

Traçons la représentation graphique de cette fonction.



On s'aperçoit graphiquement que si le nombre d'enfants paramètre de la loi binômiale est 1 ou 2, la suite converge vers 1. Asymptotiquement, l'extinction est donc presque sûre. Par contre si a=3, ou si a=5 cette suite converge vers deux valeurs dont les approximations sont $2{,}4.10^{-1}$ et 4.10^{-2}.

Un exercice classique.

On admet qu'un constructeur de moteurs fournit une compagnie d'aviation qui possède des biréacteurs et des quadriréacteurs. La probabilité d'incident sur un moteur est estimée à p, et l'on supposera que ces incidents sont indépendants les uns des autres.

On demande de comparer, en fonction de p, la fiabilité d'un biréacteur et d'un quadriréacteur, si l'on sait qu'un avion s'écrase dès qu'il a plus de la moitié de ses moteurs en panne.

Appelons X et Y le nombre de moteurs en panne pour le bi (resp quadri) réacteur, et A (resp A') l'événement l'avion s'écrase.

$$P(A) = P[X=2] = p^2$$

On remarque que X et Y suivent les lois binômiales $\mathcal{B}(2,p)$ et $\mathcal{B}(4,p)$.

$$P(A') = P[Y\geq 3] = p^4 + 4p^3q$$

$$\Delta = P(A')-P(A) = p^4 + 4p^3q - p^2 = p^4+4p^3q-p^2 = p^2(1-p)(3p-1)$$

Δ est donc du signe de 3p-1.

Si p< 1/3 il est donc moins risqué de voyager en quadriréacteur.

4 ESPÉRANCE ET VARIANCE.

Puisque $X \simeq \mathcal{B}(n,p)$ est la somme de n variables indépendantes de Bernoulli de paramètre p notées X_i on a

$$E(X) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = np$$

$$Var(X) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = npq \qquad \text{par indépendance.}$$

Exercice [16] :

On possède 100 pièces, 50 sont bien équilibrées et 50 sont truquées, donnant pile à trois chances contre quatre.
Pour éliminer les mauvaises pièces on décide d'effectuer l'opération suivante. On lancera chacune des pièces quatre fois et si le nombre de pile excède trois on jettera la pièce suspecte.

1°) Combien éliminerait-on de pièces en moyenne si toutes étaient saines?
2°) Quelle est la probabilité d'éliminer une pièce saine?
3°) Quelle est la probabilité de garder une mauvaise pièce?
4°) Trouver la loi du nombre de pièces éliminées, et du nombre de pièces saines éliminées. Donner les moyennes.

1°) Appelons N le nombre de pile qui sort pendant un jeu de quatre lancers. Si toutes les pièces étaient saines $N \simeq \mathcal{B}(4,\frac{1}{2})$.
On éliminerait une pièce donnée avec la probabilité

$$p' = P[N \geq 3] = \left(\frac{1}{2}\right)^4 + 4\left(\frac{1}{2}\right)^4 = \frac{5}{2^4} = \frac{5}{16}$$

Comme on recommence le jeu sur chacune des 100 pièces, le nombre des pièces éliminées $X' \simeq \mathcal{B}(100,p')$.

En moyenne on élimine ainsi $E(X') = 100p' = 31{,}25$ pièces.

2°) Appelons E l'événement "éliminer la pièce" et T l'événement "la pièce est truquée". N le nombre de pile parmi les quatre lancers.

On a $E = [N \geq 3]$

On ne peut pas connaître a priori la loi de N, par contre il est facile de noter que la loi de $N/\bar{T}$ reste $\mathcal{B}(4,\frac{1}{2})$ alors que N/T suit $\mathcal{B}(4,\frac{3}{4})$.

Donc

$$P(\bar{T} \cap E) = P([N \geq 3] \cap \bar{T}) = P([N \geq 3] / \bar{T}).P(\bar{T}) = \left[\left(\frac{1}{2}\right)^4 + 4\left(\frac{1}{2}\right)^4 \right]\left(\frac{1}{2}\right) = \frac{5}{32}$$

3°) Avec les mêmes notations

$$P(T \cap \bar{E}) = P([N \leq 2] \cap T) = P([N \leq 2] / T)\ P(T) = \left(\left(\frac{1}{4}\right)^4 + 4\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right) + 6\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2 \right)$$

---

[16] Cet exercice est fabriqué à partir d'un sujet de l'ISG 1984.

$$P(T\cap \bar{E}) = \frac{67}{512}$$

4°) Pour calculer le nombre de pièces éliminées il faut d'abord calculer p la probabilité de P(E).

On voit que $p=P(E) = P((E\cap T)\cup(E\cap \bar{T})) = P(E\cap T)+P(E\cap \bar{T})$

(événements disjoints)

$$p = P(T)-P(\bar{E}\cap T) + P(E\cap \bar{T}) = \frac{1}{2} - \frac{67}{512} + \frac{5}{32} = \frac{269}{512}$$

Appelons X le nombre de pièces éliminées. $X \simeq \mathcal{B}(100,p)$.
En moyenne on éliminera donc 100p soit 52,34 pièces.

Si nous appelons X_1 et X_2 le nombre de pièces saines et truquées éliminées pendant le jeu il est facile de voir que:

$$X_1 \simeq \mathcal{B}(100,p_1) \quad \text{et} \quad X_2 \simeq \mathcal{B}(100,p_2).$$

avec $p_1 = P(E\cap \bar{T}) = \frac{5}{32}$

donc en moyenne $100p_1 \cong 15,62$ pièces saines

et $p_2 = P(E\cap T) = \frac{189}{512}$

donc en moyenne $100p_2 \cong 36,91$ pièces truquées.

Avec une opération on a donc éliminé, en moyenne 73,82% des mauvaises pièces mais cela a occasionné une perte de 30,72% des bonnes ce qui n'est pas très efficace.

5 ESTIMATION DU MODE

Appelons b(n,p,k) le kième coefficient $C_n^k\, p^k q^{n-k}$

$$\frac{b(n,p,k)}{b(n,p,k-1)} = \frac{n!}{(n-k)!k!}\ \frac{(n-k-1)!(k-1)!}{n!}\ \frac{p}{q} = \frac{n-k+1}{k}\ \frac{p}{q}$$

$$\frac{b(n,p,k)}{b(n,p,k-1)} = \frac{n-k+1}{k}\ \frac{p}{q} = \frac{(n+1)p}{kq} - \frac{p}{q} = \frac{(n+1)p}{kq} - \frac{1}{q} + 1 = \frac{(n+1)p-k}{kq} + 1$$

Le coefficient binômial est donc croissant tant que ce rapport est supérieur à 1, c'est à dire tant que $k \leq (n+1)p$

Si $(n+1)p$ est un entier alors il y a deux modes obtenus pour $(n+1)p$ et son successeur.

Sinon appelons m la partie entière de $(n+1)p$, la suite est croissante jusqu'à m qui représente donc le mode.

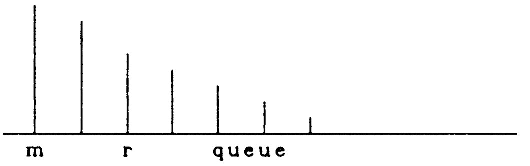
REMARQUE: si on joue 100 fois à pile ou face, le mode vaut E(101/2) = 50[17] . Mais la probabilité correspondante est faible, puisque $P[X=50] \cong 0,08$.

---

[17] Attention, c'est la partie entière.

6 ESTIMATION DE LA QUEUE.

On peut d'abord comparer la "queue" à partir du rang r au coefficient b(r,n,p) (on fixe r strictement supérieur à np pour être à droite de m et strictement inférieur à n).



La queue est donc $P[X\geq r] = \sum_{i=0}^{n-r} b(n,p,r+i)$

Or nous avons vu au § précédent sur le mode que

$$\frac{b(n,p,k)}{b(n,p,k-1)} = \frac{(n+1)p-k}{kq} +1$$

donc si i est un entier supérieur ou égal à 1: (p-i<0)

$$b(n,p,r+i) = b(n,p,r+i-1)\left(1 +\frac{(n+1)p-r-i}{(r+i)q}\right) \leq b(n,p,r+i-1)\left(1 +\frac{np-r}{rq}\right)$$

On démontre facilement par récurrence que quelque soit i positif:

$$b(n,p,r+i) \leq \left(1 +\frac{np-r}{rq}\right)^i b(n,p,r)$$

et donc

$$P[X\geq r] = \sum_{i=0}^{n-r} b(n,p,r+i) \leq \sum_{i=0}^{n-r} \left(1 +\frac{np-r}{rq}\right)^i b(n,p,r) \leq \sum_{i=0}^{+\infty} \left(1 +\frac{np-r}{rq}\right)^i b(n,p,r)$$ [18]

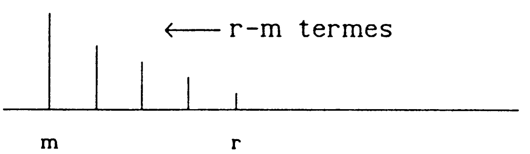
$$P[X\geq r] \leq b(n,p,r) - \frac{1}{1 -\left(1 +\frac{np-r}{rq}\right)} = b(n,p,r) \frac{rq}{r-np}$$

Exercice: Une estimation "globale de la queue".

Montrer que $P[X\geq r] \leq \frac{rq}{(r-np)^2}$

Pour ce faire il faut chercher une majoration de b(n,p,r).

Soit m le rang du terme central. Entre le rang m et le rang r il y a r-m termes.



Tous ces termes sont inférieurs à b(n,p,r) donc

$$(r-m)b(n,p,r) \leq 1 \qquad \text{et } b(n,p,r) \leq \frac{1}{r-m} \leq \frac{1}{r-np}$$ [19]

On trouve bien le résultat $P[X\geq r] \leq \frac{rq}{(r-np)^2}$.

---

[18] Puisque cette série géométrique est convergente car

$r(1-q) < np$ donc $r-np < rq$ et $0< 1 +\frac{np-r}{rq} < 1$

[19] Rappelons nous que m est la partie entière de (n+1)p et donc m≥np.

7 SOMME DE DEUX VARIABLES BINÔMIALES.

Proposition: Soient X_1 et X_2 deux variables indépendantes de loi $\mathcal{B}(n_1,p)$ et $\mathcal{B}(n_2,p)$. Alors la variable X_1+X_2 suit la loi $\mathcal{B}(n_1+n_2,p)$.

Démonstration: Il est clair que $(X_1+X_2)(\Omega) = \{0,1,\ldots,n_1+n_2\}$

On a $P[X_1+X_2=k] = P[\bigcup_{i=0}^{k}(X_1=i)\cap(X_2=k-i)]$ les événements sont disjoints

$$= \sum_{i=0}^{k} P[(X_1=i)\cap(X_2=k-i)]$$

$$= \sum_{i=0}^{k} P[X_1=i]\ P[X_2=k-i] \quad \text{(par indépendance)}$$

$$= \sum_{i=0}^{k} C_{n_1}^{i}\ C_{n_2}^{k-i} p^k q^{n_1+n_2-k} = p^k q^{n_1+n_2-k} \sum_{i=0}^{k} C_{n_1}^{i}\ C_{n_2}^{k-i}$$

Ceci étant fait on peut adopter deux attitudes.

On peut procéder à une série de n_1+n_2 épreuves de Bernoulli indépendantes de même paramètre p.
Si l'on s'arrête pour faire une pause au bout de n_1 tentatives, on peut appeler X_1 le nombre de succès qu'on aura obtenus alors.
De même X_2 représentera le nombre de succès obtenus ensuite. X_1+X_2 suit donc une loi $\mathcal{B}(n_1+n_2,p)$ et $P[X_1+X_2=k]= C_{n_1+n_2}^{k}\ p^k q^{n_1+n_2-k}$
On retrouve donc une preuve probabiliste de la formule de Vandermonde:

$$\sum_{i=0}^{k} C_{n_1}^{i}\ C_{n_2}^{k-i} = C_{n_1+n_2}^{k}$$

A l'envers on peut aussi supposée la formule de Vandermonde connue, pour retrouver analytiquement la loi de X_1+X_2.

8 LOI BINÔMIALE A PLUSIEURS TENTATIVES.

On cherche à joindre n correspondants (les appels étant indépendants). On estime à p la probabilité de pouvoir joindre un correspondant donné. On appelle aussi X le nombre de correspondants appelés, et l'on <u>recommence</u> ensuite n-X essais vers les n-X absents. Soit Y le nombre de succès au second essai, et Z le nombre total de succès. Quelle est la loi de Z?

Ecrivons d'abord que $Z(\Omega) = \{0,1,..,n\}$
Soit k un entier fixé de $Z(\Omega)$:

$P[Z=k] = \sum_{i=0}^{n} P[Z=k\ /X=i]P[X=i]$ En appliquant la formule des probabilités totales au système complet d'événements $[X=i]$, $é\in\{0,1..,é\}$.

$$P[Z=k] = \sum_{i=0}^{k} P[Y=k-i/X=i]P[X=i] = \sum_{i=0}^{k} C_{n-i}^{k-i}\, p^{k-i}q^{n-k}\, C_n^i p^i q^{n-i}$$

Or nous avons déjà vu que $C_n^i\, C_{n-i}^{k-i} = C_n^k\, C_k^i$ [20].

Donc

$$P[Z=k] = C_n^k\, p^k q^{2(n-k)} \sum_{i=0}^{k} C_k^i\, q^{k-i} = C_n^k\, p^k q^{2(n-k)} \left(1+q\right)^k$$

$$P[Z=k] = C_n^k\, \left(p(1+q)\right)^k \left(q^2\right)^{(n-k)}$$

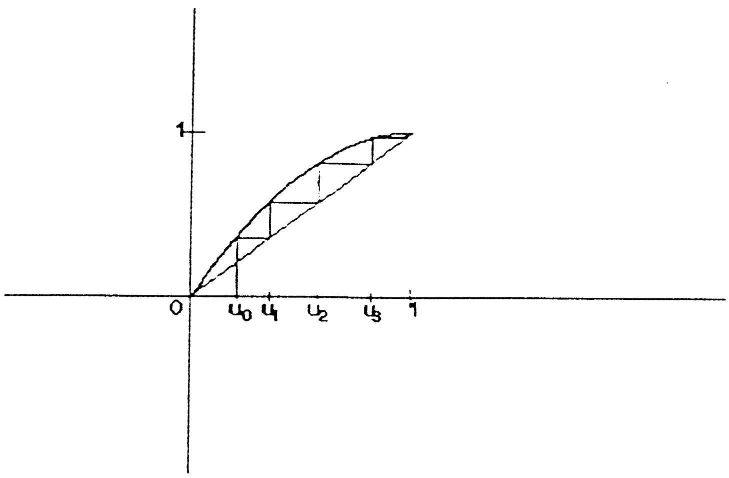
On reconnaît une loi binômiale de paramètre $p(1+q)$ sous réserve de vérifier que la somme $p(1+q) + q^2 = q(p+q)+p = q+p = 1$.

En itérant nos tentatives on obtient ainsi un processus qui reste binômial, mais on améliore le paramètre.
Ainsi si $p=0{,}2$ on trouve comme nouveau paramètre $p'=0{,}36$.

D'une façon plus générale on peut, surtout si n est grand, songer à itérer le processus. Le rang r+1 est défini à partir du rang r par la relation suivante:
Si $X_r \simeq \mathcal{B}(n,u_r)$ $\quad X_{r+1} \cong \mathcal{B}(n,u_{r+1})$

$$\text{avec } u_{r+1} = u_r(1+(1-u_r)) = u_r(2-u_r) \qquad \text{et } u_0 = p.$$

Pour $p=0{,}2$ on a tracé sur le graphique qui suit les premiers termes de la suite des paramètres:



[20] Le second membre s'interprète comme le nombre de comités de k personnes élisant, en leur sein, un directoire de i présidents. Alors que celui de droite s'interprète comme la désignation des i présidents parmi les n personnes, nommant directement leur k-i subordonnés.

9 LE PROBLÈME DE L'ÉQUIRÉPARTITION

On lance une pièce équilibrée 2n fois. Quelle est la probabilité de l'équirépartition parfaite, a_n? Que devient cette probabilité quand n tend vers $+\infty$?

Appelons X_n le nombre de pile obtenus lors des 2n lancers.

X_n suit la loi $\mathcal{B}(n,\frac{1}{2})$. [21]

$$\text{et donc } a_n = P[X=n] = C_{2n}^{n} \left(\frac{1}{2}\right)^{2n}$$

Mais si on ne prend pas garde, si n est grand le numérateur ou le dénominateur dépasse la capacité de la calculatrice, nous empêchant de deviner la réponse à la deuxième question.

Pour pouvoir faire un calcul récurrent on remarque que:

$$\frac{a_n}{a_{n-1}} = \frac{2n-1}{2n} \qquad \text{et } a_0=1 \quad a_1=\frac{1}{2}$$

On peut ainsi dans tout langage programmer l'algorithme suivant:

Saisir n.
$P \longleftarrow 1.$
Pour i allant de 1 à n $\quad P \longleftarrow \frac{2i-1}{2i} * P$
Afficher P.

Avec une classe qui connaît le calcul intégral, on peut aussi utiliser les intégrales de "Wallis".

On pose $\quad I_n = \int_0^{\pi/2} \sin^n x \, dx$

On a $\quad I_0 = \int_0^{\pi/2} dx = \frac{\pi}{2}$

$$I_1 = \int_0^{\pi/2} \sin x \, dx = \Big[-\cos x\Big]_0^{\pi/2} = 1$$

Soit n un entier supérieur ou égal à 2.

On pose $\quad u = \sin^{n-1}x \qquad v' = \sin x$

$u' = (n-1)\cos x \, \sin^{n-2}x \qquad v = -\cos x \qquad$ u et v sont de classe C_1 sur $\mathbb{R}$.

---

[21] On peut aussi, avec des élèves qui ignorent la loi binômiale, facilement compter les cas possibles 2^{2n}, les cas favorables étant le nombre de mots que l'on peut écrire avec n P et n F.

$$I_n = \int_0^{\pi/2} \sin^n x \, dx = \Big[-\cos x \sin^{n-1} x\Big]_0^{\pi/2} + (n-1) \int_0^{\pi/2} \cos^2 x \sin^{n-2} x \, dx$$

$$I_n = (n-1) \int_0^{\pi/2} \cos^2 x \sin^{n-2} x \, dx = (n-1) \int_0^{\pi/2} (1-\sin^2 x) \sin^{n-2} x \, dx$$

$$I_n = (n-1)(I_{n-2} - I_n)$$
$$nI_n = (n-1) I_{n-2}$$

Ainsi $\dfrac{I_{2k}}{I_{2k-2}} = \dfrac{2k-1}{2k}$

On constate donc que la suite des rangs pairs extraite de (I_n) obéit à la même relation de récurrence que (a_k).

On a donc facilement en comparant les premiers termes que

$$I_{2k} = a_k \frac{\pi}{2}$$

$\dfrac{I_{2k+1}}{I_{2k-1}} = \dfrac{2k}{2k+1}$ donc

$$\frac{I_{2k-1}}{I_{2k+1}} = \frac{2k-1}{2k} \cdot \frac{2k+1}{2k-1}$$

$$\frac{I_{2k-1}}{I_{2k+1}} \cdot \frac{2k-1}{2k+1} = \frac{2k-1}{2k}$$

Et donc la suite $\left(\dfrac{1}{(2k+1)\, I_{2k+1}}\right)$ obéit à la même loi de formation que la suite (a_k). En comparant les premiers termes on trouve:

$$a_k = \frac{1}{(2k+1)\, I_{2k+1}}$$

Comme $I_{n+1} - I_n = \displaystyle\int_0^{\pi/2} \sin^n x \, (\sin x - 1) \, dx$, intégrale d'une fonction négative sur l'intervalle $[0, \frac{\pi}{2}]$, est négatif, la suite (I_n) est décroissante.

Pour tout entier n on a:

$$I_{2n+2} \le I_{2n+1} \le I_{2n}$$

et donc

$$a_{n+1}\frac{\pi}{2} \le \frac{1}{(2n+1)a_n} \le a_n\frac{\pi}{2}$$

$$a_n\frac{2n+1}{2n+2}\frac{\pi}{2} \le \frac{1}{(2n+1)a_n} \le a_n\frac{\pi}{2}$$

$$\frac{2n+1}{2n+2}\frac{\pi}{2} \le \frac{1}{(2n+1)a_n^2} \le \frac{\pi}{2}$$

$\frac{1}{(n+1)a_n^2}$ est encadré par deux suites qui convergent vers la même limite $\frac{\pi}{2}$

et converge donc vers cette limite.

$$\frac{1}{(n+1)a_n^2} \cong \frac{\pi}{2}$$

$$a_n \cong \frac{1}{\sqrt{\pi n}}$$

Quand n tend vers $+\infty$, la probabilité de l'équirépartition tend donc vers 0.
Quand on pose la question aux élèves, ils s'attendent généralement à trouver 1, car ils confondent la loi des grands nombres (intuitive) avec l'équirépartition: si on lance un grand nombre de fois la pièce, seulement la <u>fréquence</u> des pile sera voisine de celle des face ce qui ne nous dit rien sur le nombre exact de pile.

Pour généraliser ce que nous venons de faire, cherchons un équivalent du terme central de $\mathcal{B}(n,p)$.

10 UN ÉQUIVALENT DU TERME CENTRAL.

Nous avons vu que dans le cas général le rang du terme central était m, avec m la partie entière de (n+1)p.

$$m = np + \delta \qquad\qquad -q \le \delta \le p$$

Le terme central $b(n,p,m) = C_n^m\, p^m q^{n-m}$

or $p-1 \le \delta \le p$ donc $p-1 \le m-np \le p$

$$-p \le np-m \le 1-p$$

et enfin

$$\frac{m}{n+1} \leq p \leq \frac{m+1}{n+1}$$

p est donc équivalent quand n et m tendent vers l'infini à $\frac{m}{n}$ et q à $1-\frac{m}{n}$.

Pour regarder quand $p=\frac{m}{n}$, on utilise la formule de Stirling (cf Appendice 2) pour trouver d'abord un équivalent du coefficient binômial.

$$C_n^m = \frac{n!}{m!(n-m)!} \cong \frac{1}{\sqrt{2\pi}} \frac{n^{n+\frac{1}{2}} e^{-n+m-(n-m)}}{m^{m+\frac{1}{2}} (n-m)^{(n-m)+\frac{1}{2}}}$$

$$\cong \frac{1}{\sqrt{2\pi n}} \left(\frac{n}{m}\right)^{m+\frac{1}{2}} \left(\frac{n}{n-m}\right)^{n-m+\frac{1}{2}}$$

$$\cong \frac{1}{\sqrt{2\pi n}} \frac{1}{p^{m+\frac{1}{2}} q^{n-m+\frac{1}{2}}}$$

Et donc $b(n,p,m) = C_n^m\, p^m q^{n-m} \cong \frac{1}{\sqrt{2\pi pqn}}$

11 LA LOI BINÔMIALE COMME LOI D'APPROXIMATION

Lorsqu'un pêcheur chanceux prélève dans un lac, quelques specimen de l'art halieutique, il ne procède pas à une suite d'épreuves de Bernoulli de même paramètre, sauf à remettre chaque prise à l'eau, ce qui paraît-il se fait maintenant. Mais il est clair que si le nombre de poissons est très grand par rapport au nombre de prélèvements, la loi du nombre de poissons pêchés et comestibles sera très proche d'un "schéma de Bernoulli".
Nous allons essayer de préciser ces notions intuitives.

Supposons qu'il y ait en tout N boules. p la proportion des blanches, q celle des noires.
Nous avons donc Np boules blanches et Nq noires.
Appelons X le nombre de boules blanches obtenues, lors de n tirages.

$X(\Omega)$ est inclus dans $\{0,1,\ldots,n\}$. [22]

On a $P[X=k] = \frac{C_{Np}^k\, C_{Nq}^{n-k}}{C_N^n}$

[22] En fait il est exactement égal à $\{Max(0,n-Nq)\ldots,Min(n,Np)\}$.

Nous vérifions que nous avons bien une loi de probabilité puisque d'après la formule de Vandermonde $\sum_{k=0}^{n} C_{Np}^{k} C_{Nq}^{n-k} = C_{N}^{n}$.

Pour obtenir l'espérance sans trop de calculs on peut faire l'exercice suivant:
Exercice: On appelle X_i la variable aléatoire (de Bernoulli) qui vaut 1 si le ième tirage a été un succès, c'est à dire blanc, et 0 sinon. Quelle est la loi de X_i. En déduire la loi de X.

Pour calculer le paramètre de cette variable, il faut prendre un modèle de n listes sans répétition c'est à dire d'arrangements.
Les cas possibles sont alors A_N^n.
Les cas favorables sont de Np choix pour la ième case, les N-1 autres boules devant ensuite s'arranger sur les n-1 autres cases. [23]

$$P[X_i=k] = \frac{Np \quad A_{N-1}^{n-1}}{A_N^n} = p$$

X compte le nombre total de succès et vaut donc $\sum_{i \in I} X_i$.
Ainsi $E(X) = E(\sum_{i \in I} X_i) = \sum_{i \in I} E(X_i) = np$.

REMARQUE: La loi de X s'appelle la loi hypergéométrique de paramètre N, n, p.
$X \cong \mathcal{H}(N,n,p)$.

Regardons maintenant ce qui se passe lorsque n est très petit devant N, c'est à dire, mathématiquement, lorsque n est fixé et N tend vers l'infini.

Soit k un entier inférieur à n et donc fini.
Pour simplifier notons M=Np et donc Nq=N-M

$$\frac{C_M^k \, C_{N-M}^{n-k}}{C_N^n \, C_n^k} = \frac{M!}{(M-k)!k!} \, \frac{(N-M)!}{(N-M-(n-k))!(n-k)!} \, \frac{(N-n)!n!}{N!} \, \frac{k!(n-k)!}{n!}$$

$$= \frac{M!}{(M-k)!} \, \frac{(N-k)!}{N!} \, \frac{(N-n)!}{(N-k)!} \, \frac{(N-M)!}{(N-M-(n-k))!}$$

---

23 C'est ici que l'on comprend pourquoi on a introduit ce modèle d'arrangements. S nous avions pris un modèle de combinaison, avec le même raisonnement, on aurait trouvé Np choix pour le ième tirage, mais en reprenant les N-1 autres boules, on aurait fait l'erreur dite "du roi de pique" dans le premier volume.

$$\text{Or} \quad \frac{M!}{(M-k)!}\,\frac{(N-k)!}{N!} = \frac{M(M-1)\ldots\ldots(M-k+1)}{N(N-1)\ldots\ldots(N-k+1)} \cong \left(\frac{M}{N}\right)^k = p^k$$

Car nous avons exactement k facteurs en haut et en bas.

$$\text{et} \quad \frac{(N-n)!}{(N-k)!}\,\frac{(N-M)!}{(N-M-(n-k))!} = \frac{(N-n)(N-n-1)\ldots\ldots(N-n-k+1)}{(N-M-(n-k))\ldots\ldots\ldots(N-M+1)} \cong \left(\frac{N-M}{N}\right)^k = q^k$$

$$\text{Enfin} \quad \frac{C_M^k\, C_{N-M}^{n-k}}{C_N^n} \underset{n\to+\infty}{\cong} C_n^k\, p^k q^{n-k}$$

En pratique on recommande, quand $N \geq 10n$, de confondre les lois du tirage avec ou sans remise, et donc de remplacer $\mathcal{H}(N,n,p)$ par $\mathcal{B}(n,p)$.

III LA LOI DE POISSON

Nous venons de voir que la loi binômiale sert pour approcher la loi du tirage sans remise, appelée loi hypergéométrique, regardons maintenant ce qui se passe pour $\mathcal{B}(n,p)$ lorsque n est grand, p petit et le produit np raisonnable.

1 APPROXIMATION PAR LA LOI DE POISSON

THEOREME et DEFINITION: Soit (X_n) une suite de variables aléatoires suivant toutes une loi $\mathcal{B}(n,p_n)$ telles que, quand n tend vers $+\infty$ np_n converge vers un réel λ non nul.
Alors pour tout entier k fixé

$$P[X_n=k] = C_n^k\, p_n^k\, (1-p_n)^{n-k} \qquad \text{converge vers le réel} \qquad e^{-\lambda}\, \frac{\lambda^k}{k!}.$$

La fonction définie sur $\mathbb{N}$ par $f(k)= \lim_{n\to+\infty} P[X_n=k]$ est la loi de probabilité d'une variable aléatoire X [24] appelée distribution de Poisson.

Démonstration:
Remarquons d'abord que λ est nécessairement strictement positif (limite de réels positifs) et que p_n converge vers 0. [25]

Soit k un entier naturel fixé.

$$P[X_n=k] = C_n^k\, p_n^k\, (1-p_n)^{n-k} = \frac{1}{k!}\; \frac{n!}{(n-k)!\,n^k}\, (np_n)^k (1-p_n)^{n-k}$$

(dans toute la suite ε_n désigne une suite convergente vers 0.)

$$(1-p_n)^{n-k} = \exp(\ (n-k)\ \ln(1-p_n)\ ) = \exp\ (\ (n-k)\ (-p_n + p_n\varepsilon_n)\ )$$
$$= \exp\ (-np_n + \varepsilon'_n)$$

et donc la suite $(1-p_n)^{n-k}$ converge vers $e^{-\lambda}$.

Comme la fraction rationnelle $\frac{n!}{(n-k)!\,n^k}$ converge vers 1 et la suite $(np_n)^k$ converge vers λ^k par continuité de la fonction puissance, finalement $P[X_n=k]$ converge vers $e^{-\lambda}\, \frac{\lambda^k}{k!}$.

Pour vérifier que l'on obtient bien une loi de probabilité il suffit de remarquer que

---

[24] On dit aussi que la suite de variables aléatoires (X_n) converge en loi vers la variable X suivant la loi de Poisson.

[25] $np_n = \lambda + \varepsilon_n$ avec $\varepsilon_n \longrightarrow 0$ donc $p_n = \lambda/n + \varepsilon_n/n$.

$$\sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1$$

$$\text{puisque } \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{\lambda}.$$

En pratique, on demande aux candidats des écoles de commerce d'appliquer le théorème avec les hypothèses suivantes:

$$\begin{cases} n \geq 30 \\ p \leq 0,1 \\ np \leq 10 \end{cases}$$

Il est pourtant facile de voir qu'on obtient d'excellentes approximations même au delà de ce champ d'application, qui a donc simplement un rôle normalisateur.

Exercice: Quelle est la probabilité pour que k personnes d'une assemblée de 5000 fêtent leur anniversaire le jour de Noël?

Soit X la variable qui compte le nombre de personnes nées le jour de Noël. X est une somme de variables de Bernoulli indépendantes de même paramètre 1/365. Donc $X \cong \mathcal{B}(5000, 1/365)$.

Si l'on cherche à approximer cette loi, on constate que n et p sont bien dans le champ d'application mais pas np qui vaut environ 14.

k	2	3	4	5	6	7
$\mathcal{P}(14)$	0,0001	0,0004	0,0013	0,0037	0,0087	0,0174
$\mathcal{B}$in	0,00010	0,00048	0,00163	0,00449	0,01026	0,02011
$\mathcal{P}(\frac{5000}{365})$	0,00011	0,00048	0,00165	0,00452	0,01032	0,02019

L'approximation, surtout quand on choisit le paramètre exact pour la loi de Poisson, est malgré tout, excellente.[26]

2 LA LOI DE POISSON POUR ELLE MÊME

Nous venons de trouver que si X suit une loi de Poisson de paramètre λ, réel strictement positif, on écrit $X \cong \mathcal{P}(\lambda)$:

$$X(\Omega) = \mathbb{N} \qquad P[X=k] = e^{-\lambda} \frac{\lambda^k}{k!} = p(k,\lambda)$$

Il est facile de voir que

[26] La première ligne a été lue dans une table, les deux autres calculées.

$$E(X)=\sum_{k=0}^{+\infty} kP[X=k] = \sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!} = \lambda$$

Pour calculer la variance, le plus simple est de calculer E(X(X-1))

$$E(X(X-1)) = \sum_{k=0}^{+\infty} k(k-1)P[X=k] = \sum_{k=2}^{+\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

$$Var(X) = E(X(X-1)) + E(X) - E^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Si nous cherchons le mode de cette distribution, il faut écrire:

$$\frac{p(k,\lambda)}{p(k-1,\lambda)} = \frac{\lambda}{k}$$

Si $\lambda<1$ λ/k est toujours inférieur à 1, et donc la loi est strictement décroissante, le mode est 0.

Si $\lambda>1$ la distribution est croissante jusqu'à la partie entière de λ, et décroissante ensuite, le mode est la partie entière de λ. [27]

Exercice: Dans un grand magasin , le nombre de magnétoscopes vendus au cours d'une journée , X, suit une loi de Poisson de paramètre 4.
Les ventes sont supposées 2 à 2 indépendantes.

a) Calculer la probabilité de chacun des événements suivants:
- La vente journalière est au plus égale à 5.
- La vente journalière est au plus égale à 2 ou au moins égale à 6
- La vente journalière est au plus égale à 6 sachant qu'elle est au moins égale à 2.

b) Calculer la probabilité que 2 jours de suite la vente soit au moins égale à 5.

c) Calculer la probabilité que la somme des ventes de 2 jours consécutifs soit inférieure ou égale à 2.

d) Le directeur de ce magasin décide d'effectuer pendant une semaine une campagne publicitaire sur les magnétoscopes.
Il estime que pendant cette semaine la vente suivra toujours une loi de Poisson et que son paramètre μ sera égal à 6 avec la probabilité 2/3 ou bien sera égal à 8 avec la probabilité 1/3.
Quelle devrait alors être la probabilité que pendant cette campagne la vente journalière de magnétoscopes soit au moins 5 ?

Pour traiter cet exercice de calcul numérique on utilise les tables de la loi de Poisson, fournissant la fonction de répartition $F(k)=P[X\leq k]$ [28] .

---

27 Si λ est entier, le quotient peut faire 1 et donc la distribution possède alors deux valeurs modales.

28 Il est encore plus commode d'utiliser une calculatrice mais attention, il vous faudr sans doute alors passer par la fonction décroissante $G(k)=P[X\geq k]$.

a) Nous donnons les approximations à 0,001 près.

$P[X\leq 5] = F(5) \simeq 0,785$

$P[(X\leq 2)\cup(X\geq 6)] = P[X\leq 2] + P[X\geq 6] = F(2) + 1 - F(5)$ [29] $\simeq 0,453$

$$P[(X\leq 6)/(X\geq 2)] = \frac{P[(X\leq 2)\cap(X\geq 6)]}{P[X\leq 2]} = \frac{P[X\leq 6]-P[X\leq 2]}{F(2)} = \frac{F(6)-F(2)}{F(2)} \simeq 0,878$$

b) Appelons X_1 et X_2 les variables aléatoires (indépendantes) des ventes du premier et du second jour. Les deux variables ont la même loi.

$$P[(X_1\geq 5)\cap((X_2\geq 5)] = P[X_1\geq 5]^2 = (1-F(5))^2 \simeq 0,138$$

c) $P[X_1+X_2\leq 2] = P[X_1+X_2=0] + P[X_1+X_2=1] + P[X_1+X_2=2]$

$= P[(X_1=0)\cap(X_2=0)] + P[(X_1=0)\cap(X_2=1)]+P[(X_1=1)\cap(X_2=0)] +$

$P[(X_1=0)\cap(X_2=2)] + P[(X_1=1)\cap(X_2=1)\cup(X_1=0)\cap(X_2=2)]$

On utilise alors l'indépendance des variables aléatoires pour trouver f désignant la distribution de $\mathcal{P}(4)$

$$f(0)^2+2f(0)f(1)+ 2f(0)f(2) +f(1)^2=$$
$$2f(0)[F(2)] + F(1)[f(1)-f(0)] \simeq 0,014.$$

c) En appliquant la formule des probabilités totales au système complet $([\mu=6];[\mu=8])$ et en remarquant que désormais ce sont les variables conditionnées $X/[\mu=6]$ et $X/[\mu=8]$ qui suivent les lois $\mathcal{P}(6)$ et $\mathcal{P}(8)$

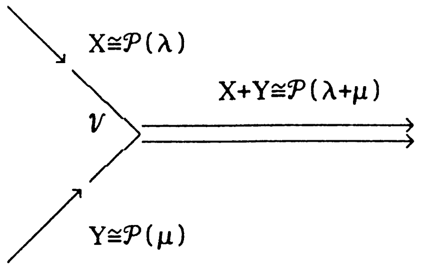
$$P[X\geq 5] = P[X\geq 5/\mu=6]P[\mu=6]+P[X\geq 5/\mu=8]P[\mu=8]$$
$$= (1 - F_6(4))2/3 + (1 -F_8(4))1/3 \simeq 0,777$$

La loi de Poisson donne de bonnes modélisations pour les problèmes de flux, trafics, et autres feux rouges. En tout cas le théorème suivant, s'interprète aisément avec cette image de flux.

THEOREME: Si X et Y sont deux variables indépendantes suivant des lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, alors X+Y suit encore une loi de Poisson de paramètre $\lambda+\mu$.

REMARQUE: Si X et Y sont les flux de voitures circulant sur des routes arrivant à la ville V et si ces flux sont indépendants, alors le flux sortant sera encore de Poisson, et les paramètres préservent les moyennes.

[29] ou encore 1-G(3)+G(6).



Démonstration:

$(X+Y)(\Omega) = \mathbb{N}$

Pour tout entier k naturel

$$P[X+Y=k] = P[\bigcup_{i=0}^{k} ((X=i)\cap(Y=k-i))] = \sum_{i=0}^{k} P[(X=i)\cap(Y=k-i)]$$

Cette dernière somme vaut puisque les variables sont indépendantes:

$$\sum_{i=0}^{k} P[X=i]P[Y=k-i] \text{ donc } \sum_{i=0}^{k} e^{-\lambda}\frac{\lambda^{i}}{i!}e^{-\mu}\frac{\mu^{k-i}}{(k-i)!} = e^{-\lambda}e^{-\mu}\frac{1}{k!}\sum_{i=0}^{k} C_k^i\,\lambda^i\mu^{k-i}$$

et en appliquant la formule du Binôme on trouve $e^{-(\lambda+\mu)}\dfrac{(\lambda+\mu)^k}{k!}$.

3 BINÔMIALE SACHANT POISSON

THEOREME: Soit X une variable aléatoire, suivant une loi de Poisson de paramètre λ et Y une variable aléatoire définie sur $\mathbb{N}$. p est un réel fixé de $]0,1[$.
Si quelque soit l'entier naturel n, $Y/[X=n]$ suit $\mathcal{B}(n,p)$ alors Y suit $\mathcal{P}(\lambda p)$.

Démonstration:

En appliquant la formule des probabilités totales au système complet ($[X=n]$ pour $n\in\mathbb{N}$) on a pour tout entier k

$$P[Y=k] = \sum_{n=0}^{+\infty} P[Y=k/X=n]P[X=n] = \sum_{n=k}^{+\infty} C_n^k\, p^k q^{n-k}\, e^{-\lambda}\frac{\lambda^n}{n!}$$

$$P[Y=k] = e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{n=k}^{+\infty}\frac{(\lambda q)^{n-k}}{(n-k)!} = e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{-\lambda q} = e^{-\lambda p}\frac{(\lambda p)^k}{k!}$$

ce qui prouve que Y suit bien une loi de Poisson de paramètre λp.

Application:
On admet que le nombre de personnes fréquentant chaque heure un grand magasin suit une loi de Poisson .La fréquentation moyenne par jour (10h) est de 2200. Le gérant décide d'organiser un jeu promotionnel. Les gagnants seront les tickets tirés dont l'addition se termine par 77. Le nombre de gagnants est calculé à chaque heure. Quelle sera la loi du nombre de gagnants. Combien y aura-t-il en moyenne de gagnants par heure?
(Il admet que le montant de toutes les factures est supérieur à 100F).

Soit X le nombre journalier de client. On sait que X suit une loi de Poisson. Comme on nous a gentiment indiqué que la moyenne journalière est 2200, on nous a fourni par la même occasion le paramètre λ, 220.

$X \cong \mathcal{P}(220)$.

Si n personnes sont présentes dans le magasin, chaque facture a une probabilité de 1/100 d'être gagnante, et donc le nombre des factures gagnantes suit une loi binômiale $\mathcal{B}(n;0,01)$.

En appliquant le théorème précédent, on trouve que la loi du nombre de gagnants est $\mathcal{P}(2,2)$. Il y a donc aussi 2,2 gagnants par heure.

4 OÙ L'ON RETROUVE LES BIJECTIONS SANS POINT FIXE.

Nous avons démontré dans le premier fascicule (page 27) que la probabilité pour qu'une bijection d'un ensemble de n éléments soit sans point fixe est $p_n = \sum_{i=0}^{n} \frac{(-1)^i}{i!}$.

On peut maintenant se donner n et considérer la variable aléatoire X qui compte le nombre de points fixes d'une permutation donnée de $\mathfrak{S}_n$.

$X(\Omega) = \{1,2,\dots,n\}$

Introduisons les événements:

A_k= " les k éléments donnés $\{x_1,x_2\dots,x_k\}$ sont fixes."

B_k=" les k éléments donnés $\{x_1,x_2\dots,x_k\}$ sont fixes et ce sont les seuls."

En effet P[X=k] sera égal à $C_n^k\ P(B_k)$, le coefficient binômial correspondant au choix des k éléments qui seront fixes.

$P(B_k) = P(B_k/A_k)P(A_k)$

L'ensemble des cas possibles est $\mathfrak{S}_n$, de cardinal n!

Le cardinal de A_k est (n-k)! car on fait tourner les n-k éléments non désignés.

Ainsi $\qquad P(A_k) = \dfrac{(n-k)!}{n!}$

$P(B_k/A_k)$ est la probabilité de faire tourner <u>sans point fixe</u> les n-k éléments non désignés et donc $P(B_k/A_k) = p_{n-k}$

$$P[X=k] = C_n^k P(A_k)p_{n-k} = C_n^k \frac{(n-k)!}{n!}\ p_{n-k} = \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!}.$$

REMARQUE: quand n tend vers $+\infty$, P[X=k] converge vers $\frac{1}{k!}\sum_{i=0}^{+\infty}\frac{(-1)^i}{i!} = \frac{e^{-1}}{k!}$.

On peut donc approcher la loi de X quand n est suffisamment grand par la loi de Poisson $\mathcal{P}(1)$.

Or l'erreur que l'on commet alors est $\left|\frac{1}{k!}\sum_{i=n-k+1}^{+\infty}\frac{(-1)^i}{i!}\right|$ inférieure à

$\frac{1}{k!}\frac{1}{(n-k+1)!}$, qui converge très rapidement vers 0.

Application: On présente à un soi-disant voyant une liste de 6 écrivains dont le nom se trouve dans une enveloppe numérotée 1,2,..,6.
L'homme est chargé, par "double vue" d'associer la bonne enveloppe avec le nom qu'elle contient. Il réussit deux rencontres...
La probabilité que la réussite soit due au hasard est

$$P[X\geq 2] \simeq e^{-1}\left(\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}\right)\simeq 0,275$$

ce qui est trop important pour qu'on puisse croire sans acte de foi dans les capacités du médium. [30]
Il faut enfin remarquer que l'approximation que l'on utilise revient à considérer que la probabilité du nombre de rencontres ne dépend pas de n dès que n est assez grand.

---

[30] Cet exercice est repris à partir du livre d'Arthur Engel (Cedic).

IV LOI DES ATTENTES

Comme nous l'avons déjà remarqué, pendant l'étude de la loi binômiale, quand on fait une succession d'épreuves de Bernoulli (ou autres) on ne cherche pas toujours à compter le nombres de succès. On cherche souvent le premier succès, ou encore si les échecs sont très coûteux, on comptera le nombre d'échecs avant le premier succès, etc...

A ATTENTE DU PREMIER SUCCÈS

1 CAS DU SCHÉMA DE BERNOULLI

On procède à une série d'épreuves de Bernoulli, indépendantes, de même paramètre $p \in]0,1[$. On appelle X le rang du premier succès.

THEOREME et DEFINITION: Avec les hypothèses précédentes, on appelle loi géométrique de paramètre p et l'on note $\mathcal{G}(p)$, la loi du rang du premier succès. $X(\Omega) = \mathbb{N}^*$

et pour tout entier k non nul $P[X=k] = q^{k-1}p$

En effet appelons (X_n) la suite des variables indépendantes de Bernoulli correspondant à notre schéma. Puisque les événements $[X_1=0]$ etc..sont mutuellement indépendants:

$$P[X=k]=P[(X_1=0)\cap(X_2=0)\cap\ldots.(X_{k-1}=0)\cap(X_k=1)] =P[(X_k=1) \prod_{i=1}^{k-1} P[(X_i=0)= p\ q^{k-1}$$

Vérification: $\sum_{k=1}^{+\infty} q^{k-1}p = p\frac{1}{1-q} = 1.$(la série étant absolument convergente puisque $q \in]0,1[$).

On va montrer par le calcul que

$$E(X) = \frac{1}{p}$$

et admettre que $Var(X) = \dfrac{q}{p^2}$.

$$E(X) = \sum_{k=1}^{+\infty} kq^{k-1}p$$ [31]

On obtient tout de suite le résultat si on sait que la série entière est dérivable dans le disque de convergence et que donc $\sum_{k=1}^{+\infty} kq^{k-1}p = p\ \frac{1}{(1-q)^2}$.

$$E_n(X) = \sum_{k=1}^{n} kq^{k-1}(1-q) = \sum_{k=1}^{n} kq^{k-1}(1-q) = \sum_{k=1}^{n} kq^{k-1} - \sum_{k=1}^{n} kq^{k}$$

$$= \sum_{k=1}^{n} kq^{k-1} - \sum_{k=2}^{n+1} (k-1)q^{k-1}$$

$$= 1 - nq^n + \sum_{k=2}^{n+1} q^{k-1}$$

$$= 1 - nq^n + q\ \frac{1 - q^{n-1}}{p}$$

Comme les suites géométriques q^{n-1} et nq^n convergent vers 0 ($q\in]0,1[$).
La suite $E_n(X)$ converge vers $1 + q/p = 1/p$.

REMARQUE: Loi géométrique avec "abandon".

On a supposé dans le modèle précédent que le temps n'était pas compté pour l'expérimentateur et qu'il avait donc l'éternité devant lui pour attendre la réussite. Dans la pratique, on peut se fixer un rang n à partir duquel on décidera d'arrêter l'expérience, quelle qu'en soit l'issue.

On obtient alors une nouvelle variable aléatoire X' avec $X'(\Omega) = \{1,2,\dots,n\}$.

Si k est inférieur à n-1 $P[X'=k] = P[X=k] = q^{k-1}p$.
Par contre on a $X'=n$ si et ssi les n-1 tentatives sont des échecs quelque soit le résultat de la nième

$$P[X'=n]=P[(X_1=0)\cap(X_2=0)\cap\dots(X_{n-1}=0)] = \prod_{i=1}^{n-1} P[(X_i=0)= q^{n-1}$$

Calculons maintenant la moyenne de X' pour voir ce que nous coûte, en moyenne, de continuer au delà du rang n fixé.

$$E(X') = \sum_{k=1}^{n-1} kq^{k-1}p = E_{n-1}(X) + nq^{n-1} = 1 - (n-1)\ q^{n-1} + q\ \frac{1 - q^{n-2}}{p} + nq^{n-1}$$

$$= \frac{1}{p} - \frac{q^{n-1}(1-p)}{p} = \frac{1}{p} - \frac{q^n}{p}$$

La différence est donc de $\frac{q^n}{p}$.
Exemple: si p=0,1 le rang moyen d'apparition du premier succès est 10.
Voilà les nouvelles moyennes si l'on abandonne à partir de 5, 10, 15, 20.

n	5	10	15	20
E(X')	4,1	6,5	7,9	8,8

2 LOIS GÉOMÉTRIQUES

Pour bien comprendre comment fonctionne la loi géométrique, on va la "décaler" d'un cran.

DEFINITION: X suit une loi de type géométrique de paramètre p, que nous noterons $\mathcal{T}(p)$, si et ssi X+1 suit la loi géométrique $\mathcal{G}(p)$.

REMARQUE: Ainsi, au lieu de compter le rang du premier succès, on compte le nombre <u>d'échecs</u> avant le premier succès.

Proposition: X suit $\mathcal{T}(p)$ si et ssi

$$X(\Omega) = \mathbb{N}$$

et pour tout entier k non nul $P[X=k] = q^k p$

REMARQUE: on trouve alors facilement la moyenne et la variance:

$$E(X) = E(Y-1) = \frac{1}{p} - 1 = \frac{q}{p} \quad (\text{avec } Y \cong \mathcal{G}(p)\ )$$

Et puisque la variance ne change pas par translation $Var(X) = \frac{q}{p^2}$

THEOREME: Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$

$X \cong \mathcal{T}(p)$ si et ssi pour tout entier k $P[X \geq k] = q^k$

$\Rightarrow$) dans le sens direct: Si X est géométrique, comme l'événement $[X \geq k]$ correspond à dire que les k premiers essais sont des échecs:

$$P[X \geq k] = q^k$$

$\Leftarrow$) Réciproquement pour tout entier k

$$P[X=k] = P[X \geq k] - P[X \geq k+1] = q^k - q^{k+1} = q^k p.$$

THEOREME: Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$

$X \cong \mathcal{T}(p)$ si et ssi pour tout couple d'entiers r et s $P[X \geq r+s / X \geq r] = P[X \geq s]$.

$\Rightarrow$) si $X \cong \mathcal{T}(p)$, d'après le théorème précédent

$$P[X \geq r+s / X \geq r] = \frac{P[(X \geq r+s) \cap (X \geq r)]}{P[X \geq r]} = \frac{P[X \geq r+s]}{P[X \geq r]} = \frac{q^{r+s}}{q^r} = q^s = P[X \geq s]$$

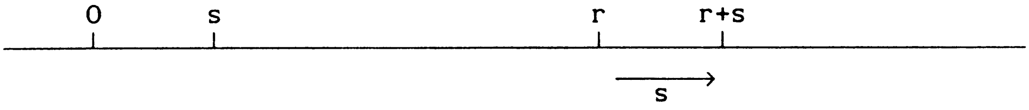
$\Leftarrow$) Réciproquement si pour tout couple d'entiers $P[X \geq r+s / X \geq r] = P[X \geq s]$

on pose $u_r = P[X \geq r]$ $\quad u_0 = P[X \geq 0] = 1 \quad u_1 = P[X \geq 1] = q$ réel de $]0,1[$

$$P[X \geq r+1 / X \geq r] = P[X \geq 1] = \frac{P[X \geq r+1]}{P[X \geq r]}$$

et donc $u_{r+1} = u_r q$ $\quad$ la suite (u_n) est donc géométrique de raison q et de premier terme 1 $\quad u_r = P[X \geq r] = q^r$ $\quad$ et donc $\quad X \cong \mathcal{T}(p) \quad p = 1-q$.

REMARQUE: ce théorème traduit ce qui se passe pendant le schéma:



$P[X \geq r+s / X \geq r] = P[X \geq s]$ signifie que, une fois les r premiers essais ratés, on a l'événement "les s suivants sont des échecs", avec la même probabilité que l'événement "les s premiers sont des échecs".
Le phénomène est donc "sans mémoire".

3 LA COLLECTION

Nous avons tous des enfants, qui collectionnent les figurines, drapeaux ou autre porte-clés que les professionnels de la consommation associent à leurs produits. Nous nous posons alors, la question suivante:" combien faudra-t-il acheter de paquets pour réaliser la collection de r figurines". Appelons N la collection complète.
Appelons S_r le nombre de tirages qu'il sera donc nécessaire d'opérer. [32]
Appelons X_k le temps d'attente entre le kième et le (k+1)ième succès.

$X_k(\Omega) = \mathbb{N}^*$. X_k est le temps d'attente du premier succès dans le schéma de Bernoulli, correspondant à la somme des variables de Bernoulli indépendantes réalisant le succès si une nouvelle figurine est acquise et donc de paramètre $\frac{N-k}{N}$.

$X_k \cong \mathcal{G}(\frac{N-k}{N}) \qquad E(X_k) = \frac{N}{N-k}$

Pour tout r compris entre 1 et N:

$$\sum_{i=1}^{r-1} X_i = \sum_{i=1}^{r-1} S_{i+1} - S_i = S_r - S_1 = S_r - 1$$

puisque la variable S_1 est sûrement égale à 1.

En posant $X_0 = 1$ on obtient $S_r = \sum_{i=0}^{r-1} X_i$

$$E(S_r) = E(\sum_{i=0}^{r-1} X_i) = \sum_{i=0}^{r-1} E(X_i) = \sum_{i=0}^{r-1} \frac{N}{N-i} = N \sum_{i=0}^{r-1} \frac{1}{N-i}$$

$$E(S_r) = N\ (\frac{1}{N} + \frac{1}{N-1} + \ldots\ldots + \frac{1}{N-r+1})$$

Ainsi si le publicitaire a eu l'idée de faire collectionner les dix commandements [33] (sur bouts de cartons multicolores ou pin's) il vous faudra attendre, en moyenne, $E(S_5) \simeq 6{,}5$ paquets, pour voir apparaître cinq commandements différents, et il faudra

---

Nous associons à l'achat d'un paquet de produit, un tirage avec remise. considérant que la taille de l'échantillon prélevé est négligeable devant la masse des produits offerts. Nous considérerons également les tirages comme indépendants.

Ce qu'il faut bien reconnaître est assez peu probable!

manger, en moyenne, $E(S_{10}) \cong 29{,}3$ paquets pour obtenir la collection complète.

Si l'on veut un critère plus systématique on peut utiliser le développement asymptotique suivant:

$$\text{on pose } s_n = \sum_{i=1}^{n} \frac{1}{i}$$

Il est bien connu que s_n est équivalent au voisinage de $+\infty$ à $\ln(n)$. [33]

$$s_n = \gamma + \ln(n) + \varepsilon(n)$$

Mais il nous faudra, si l'on veut être précis une estimation du ε.
Avec les outils des séries $a_n = s_n - \ln(n)$

$$a_n - a_{n-1} = \frac{1}{n} + \ln(1 - \frac{1}{n}) \cong \frac{-1}{2n^2}$$

et donc la série de terme général $(a_n - a_{n-1})$ est convergente, soit $\gamma - 1$ sa limite:

comme $a_n - a_1 = \sum_{i=1}^{n} a_i - a_{i-1}$, la suite (a_n) converge vers γ et

$$a_n = \gamma + \varepsilon(n)$$

$$\text{Or} \quad \varepsilon(n) = \sum_{i=n+1}^{+\infty} a_i - a_{i-1}$$

Si n est assez grand $\forall\ i > n$ $\quad (a_i - a_{i-1}) = \frac{-1}{2i^2}(1 + \varepsilon(i))$

avec $|\varepsilon(i)| \leq 1$ $\quad$ donc $|a_i - a_{i-1}| \leq \frac{1}{i^2}$

$$|\varepsilon(n)| = \left|\sum_{i=n+1}^{+\infty} a_i - a_{i-1}\right| \leq \sum_{i=n+1}^{+\infty} |a_i - a_{i-1}| \leq \sum_{i=n+1}^{+\infty} \frac{1}{i^2} \leq \sum_{i=n+1}^{+\infty} \frac{1}{i(i-1)}$$

$$|\varepsilon(n)| \leq \sum_{i=n+1}^{+\infty} \frac{1}{(i-1)} - \frac{1}{i} = \frac{1}{n+1} \leq \frac{1}{n}$$

Comme $E(S_r) = N\,(-s_N - s_{N-r})$

Avec des outils de second cycle:

pour tout i compris entre 2 et n

$$\frac{1}{i} \leq \int_{i-1}^{i} \frac{1}{t}\,dt \leq \frac{1}{i-1} \qquad \text{en sommant}$$

$$\sum_{i=2}^{n} \frac{1}{i} \leq \int_{1}^{n} \frac{1}{t}\,dt \leq \sum_{i=2}^{n} \frac{1}{i-1}$$

$$s_n - 1 \leq \ln(n) \leq s_{n-1}$$

$$\ln(n+1) \leq s_n \leq \ln(n) + 1$$

$$= N\left(\ln \frac{N}{N-r} + \varepsilon(N)\right) = N\ln \frac{N}{N-r}\left(1+\varepsilon(N)\left(\ln \frac{N}{N-r}\right)^{-1}\right)$$

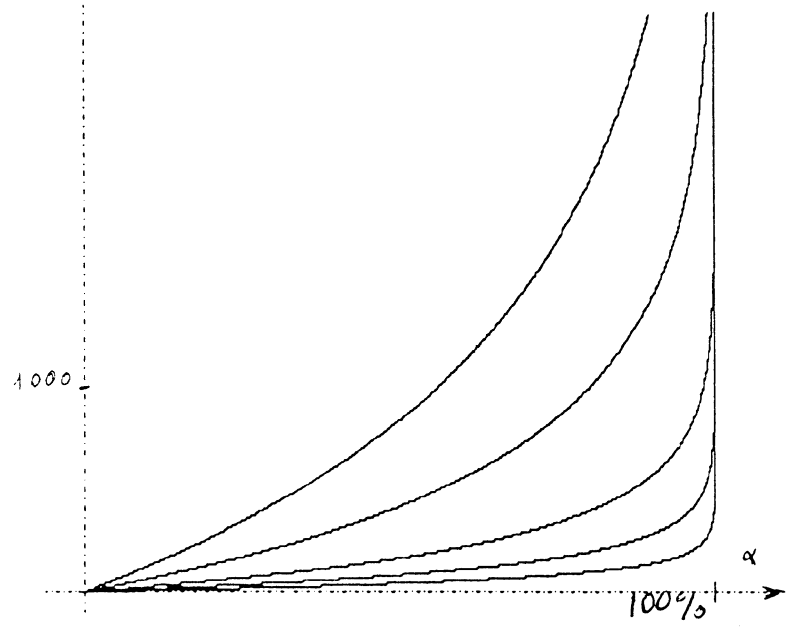
$$\left|\varepsilon(N)\left(\ln \frac{N}{N-r}\right)^{-1}\right| \leq \left(N\left(\ln \frac{N}{N-r}\right)\right)^{-1} \xrightarrow[N \to +\infty]{} 0$$

Et donc $E(S_r) \simeq N\ln \frac{N}{N-r}$. [33]

Ainsi donc pour obtenir une certaine proportion α de la collection il faut que $r \geq \alpha N$ donc $\frac{1}{N-r} \geq \frac{1}{N-\alpha N}$

et ainsi $E(S_r) \simeq N\ln \frac{N}{N-r} \geq N\ln \frac{1}{1-\alpha} = -N\ln(1-\alpha)$.

Sur la courbe qui suit, on a tracé en fonction de α, l'évolution de ce minorant de l'espérance pour les valeurs de N dans {50,100,200,500,1000}.



2 CAS DU TIRAGE SANS REMISE.

On considère une urne qui contient r boules blanches $(1 \leq r \leq n)$ et n-r boules noires. On commence une suite de tirages, <u>sans remise</u>, jusqu'à l'obtention de la première boule blanche dont le rang sera noté X.

$$X(\Omega) = \{1,2,\ldots,n-r+1\}$$

---

[33] cf Feller.

$$P[X=k] = \frac{\overset{\text{les } k-1 \text{ échecs noirs.}}{A_{n-r}^{k-1}} \; A_r^1 \longleftarrow \text{le succès: la blanche.}}{A_n^k}$$

UN CAS PARTICULIER:

Si l'urne ne possède qu'une boule blanche, r=1
$X(\Omega) = \{1,2,\ldots,n\}$

$$P[X=k] = \frac{A_{n-1}^{k-1}}{A_n^k} = \frac{1}{n}$$

On retrouve, ainsi que X suit une loi uniforme $\mathcal{U}(n)$.
Dans ce cas nous avons vu que $E(X) = \frac{n+1}{2}$ [35]

Exercice: A propos de tirages, mais ici avec remise, on peut chercher l'exercice suivant: Une urne contient N boules numérotées de 1 à N, et l'on procède à une suite de n tirages avec remise. X représente alors le plus grand numéro sorti. Trouver la limite de l'espérance mathématique de cette variable quand N tend vers l'infini.

$X(\Omega) = \{1,2,\ldots,N\}$.

Ici il est plus simple de commencer par calculer la fonction de répartition:

$P[X\le k]$ est la probabilité que les n tirages soient pris parmi les k premiers numéros: $P[X\le k]=\left(\frac{k}{N}\right)^n$

Et donc pour tout entier k de $X(\Omega)$

$$P[X=k] = P[X\le k] - P[X\le k-1] = \frac{k^n - (k-1)^n}{N^n}.$$

$$E(X) = \sum_{k=1}^{N} k \frac{k^n - (k-1)^n}{N^n} = \sum_{k=1}^{N} \frac{k^{n+1} - k(k-1)^n}{N^n} = \frac{1}{N^n} \sum_{k=1}^{N} k^{n+1} - (k-1)^{n+1} - (k-1)^n$$

$$E(X) = \frac{1}{N^n}\left(N^{n+1} - \sum_{k=1}^{N}(k-1)^n\right) = \left(N - \sum_{k=1}^{N}(1-\frac{k}{N})^n\right) = N\left(1 - \frac{1}{N}\sum_{k=1}^{N}(1-\frac{k}{N})^n\right)$$

Ceci vaut d'être comparer avec l'attente quand il y a remise. La probabilité du succès étant 1/n, nous avons vu au paragraphe précédent que l'espérance est 1/p donc ici n. L'écart "en moyenne" de l'attente avec ou sans remise vaut donc $n - \frac{n+1}{2} = \frac{n-1}{2}$.

On reconnaît quand N tend vers $+\infty$ la limite d'une somme de Riemann: [36]

$$\lim_{N \to +\infty} \frac{1}{N} \sum_{k=1}^{N} \left(1-\frac{k}{N}\right)^n = \int_0^1 (1-x)^n dx = \frac{1}{n+1}$$

$$\lim_{N \to +\infty} E(X) = N\left(1 - \frac{1}{n+1}\right) = N\frac{n}{n+1}.$$ [37]

B ATTENTES PLUS LOINTAINES

1 LOI BINÔMIALE NÉGATIVE

Soit r un entier fixé ($r \geq 1$) et p un réel de $]0,1[$.
On procède à une succession d'épreuves de Bernoulli indépendantes de même paramètre p.
On appelle X le nombre d'échecs qui précèdent le rième succès.
X suit une loi binômiale négative de paramètre r et p: $X \cong \mathcal{J}(r,p)$.

$X(\Omega) = \mathbb{N}$.

Appelons A_k l'événement obtenir le rième succès après k échecs, les r-1 premiers succès possédant des places bien déterminées.

$$P(A_k) = p^r q^k$$

Pour trouver $P[X=k]$ il suffit de multiplier A_k par le nombre de places possibles pour les r-1 succès sur les k+r-1 premières places, la dernière place étant réservée pour un succès.

$$P[X=k] = C_{k+r-1}^{r-1} \; p^r q^k$$

On peut maintenant, bien sûr, vérifier par le calcul que $\sum_{k \in \mathbb{N}} C_{k+r-1}^{r-1} \; p^r q^k = 1$.

Mais pour être certains d'obtenir une variable aléatoire, il nous suffit de regarder si nous n'avons pas oublié d'événements.

A priori nous avons omis l' événement "il y a moins de r succès".
Appelons F_n:"il y a moins de r succès pour n coups joués". ($n \geq r$)

Soit f une fonction continue sur [0,1], alors $\frac{1}{N}\sum_{k=1}^{N} f(1-\frac{k}{N})$ converge, quand N tend vers $+\infty$, vers $\int_0^1 f(t)dt$.

Feller indique, dans son livre, que cette formule était utilisée pendant la seconde guerre mondiale, par les anglais pour estimer l'effectif global de la flotte ennemie, connaissant les numéros d'identification de quelques appareils abattus. Il était facile de trouver E(X) et à rebours d'estimer N.

Soit Y le nombre de réussites sur les n coups joués, $Y \cong \mathcal{B}(n,p)$

$$P(F_n) = P[Y<r] = \sum_{i=0}^{r-1} C_n^i p^i q^{n-i}.$$

Or quand n tend vers $+\infty$ cette somme garde un nombre fini de termes et comme **de plus** pour tout i $\lim_{n \to +\infty} C_n^i p^i q^{n-i} = 0$. $P(F_n)$ converge vers 0.

Comme la suite des F_n est décroissante $P(F) = P(\bigcap_{n\in\mathbb{N}} F_n) = \lim_{n\to+\infty} P(F_n) = 0$.

X est donc bien une variable aléatoire et en prime on obtient le résultat suivant

$$\sum_{k\in\mathbb{N}} C_{k+r-1}^{r-1} p^r q^k = 1. \qquad (*)$$

REMARQUE: pour r=1 on remarque que $\mathcal{J}(1,p) = \mathcal{T}(p)$.

Calcul de l'Espérance:

1°) On peut ramener le calcul de l'espérance à l'égalité (*) si l'on se souvient que $(k+r)C_{k+r-1}^{r-1} = rC_{k+r}^{r}$ [38]

$$E(X) = \sum_{k=0}^{+\infty} kC_{k+r-1}^{r-1} p^r q^k = \sum_{k=0}^{+\infty} (k+r)\, C_{k+r-1}^{r-1} p^r q^k - \sum_{k=0}^{+\infty} rC_{k+r-1}^{r-1} p^r q^k$$

$$= \sum_{k=0}^{+\infty} rC_{k+r}^{r} \quad p^r q^k - r$$

$$= \frac{r}{p} \sum_{k=0}^{+\infty} C_{k+r+1-1}^{r+1-1} p^{r+1} q^k \quad - r$$

On retrouve encore la formule (*) avec cette fois la valeur r+1

$$E(X) = \frac{r}{p} - r = \frac{rq}{p}$$

2°) Appelons X_r le nombre d'échecs avant le r ième succès. et (Y_n) la suite de variables aléatoires ainsi définies

$$Y_1 = X_1 \qquad\qquad Y_i = X_i - X_{i-1} \quad \text{pour } i\geq 1$$

Ainsi Y_1 est le nombre d'échecs précédant le premier succès et Y_i est la nombre d'échecs séparant le (i-1)ème et le ième succès.

Comme le schéma de Bernoulli est sans mémoire, toutes les variables Y_i ont la même loi que celle de Y_1 c'est à dire une loi de type géométrique $\mathcal{T}(p)$.

Pour tout i $E(Y_i) = \frac{q}{p}$.

En posant $X_0 = 0$ sûrement, on peut écrire que

$$X_r = \sum_{i=1}^{r} X_i - X_{i-1} = \sum_{i=1}^{r} Y_i$$

Probabilités pour le lycée 1 page 10 (les chefs!).

$$E(X_r) = E(\sum_{i=1}^{r} Y_i) = \sum_{i=1}^{r} E(Y_i) = \frac{rq}{p}$$

L'intérêt de cette méthode est qu'elle s'applique pour la variance puisque les variables (Y_i) sont indépendantes:

$$Var(X_r) = Var(\sum_{i=1}^{r} Y_i) = \sum_{i=1}^{r} Var(Y_i) = \frac{rq}{p^2}$$

REMARQUE: si l'on s'interroge sur le nom de cette nouvelle loi, il faut se rappeler la définition du coefficient binômial généralisé:

Pour x réel et k entier

$$\binom{x}{k} = \frac{x(x-1).........(x-k+1)}{k!}$$ [38]

et l'on vérifie que

$$C_{k+r-1}^{r-1}\, p^r q^k = \binom{-r}{k}\, p^r(-q)^k$$

2 LOI DE PASCAL

Dans les mêmes conditions que pour le paragraphe 2 on appelle loi de Pascal de paramètre r et p et l'on note $\mathcal{P}(r,p)$ le rang d'apparition du rième succès lors du schéma de Bernoulli.

REMARQUE: si r vaut 1, on s'aperçoit que $\mathcal{P}(1,p) = \mathcal{G}(p)$.

Proposition: $X \cong \mathcal{P}(r,p)$ si et ssi $X-r \cong \mathcal{J}(r,p)$

ce qui nous permet facilement de retrouver à la fois la loi et les moments de cette nouvelle variable.

$X(\Omega) = \mathbb{N} \cap [r,+\infty[$.

Si $k \geq r$ $\quad P[X=k] = P[X-r=k-r] = C_{k-1}^{r-1}\, p^r q^{k-r}$.

Comme $E(X-r) = \frac{rq}{p}$ $\quad E(X) = \frac{rq}{p} + r = \frac{r}{p}$

et comme la variance est invariante par translation:

$$Var(X) = \frac{rq}{p^2}$$

---

[38] On retrouve si x est naturel que $\binom{x}{k} = C_x^k$.

3 OU L'ON RETROUVE UNE BOÎTE D'ALLUMETTES CONNUE

Monsieur B. possède dans chacune de ses poches n allumettes. Pour allumer ses cigarettes il pioche indifféremment dans chaque poche. Quelle est la loi de X le nombre d'allumettes qu'il lui reste quand, pour la première fois il s'aperçoit qu'une des poches est vide [40] .

Enlevons, provisoirement la symétrie du problème. On ne s'intéresse, pour l'instant, qu'à l'épuisement de la poche droite.
Le nombre d'échecs Z avant le n+1 ème succès dans cette poche suit donc une loi $\mathcal{J}(n+1,\ 1/2)$ et l'on a $X = n - Z$.

$X(\Omega) = \{0,1,\dots,n\}$

à cause de la symétrie la symétrie

$$P[X=r] = 2\ P[Z=n-r] = 2\ C_{2n-r}^{n-r} \left(\frac{1}{2}\right)^{n+1} \left(\frac{1}{2}\right)^{n-r} = C_{2n-r}^{n} \left(\frac{1}{2}\right)^{2n-r}$$

Pour calculer l'espérance de X il faut démontrer le résultat intermédiaire:

$$2(n-r)P[X=r] = (2n+1)P[X=r+1] - (r+1)P[X=r+1] \qquad (*)$$

En effet, on remarque que $\dfrac{P[X=r+1]}{P[X=r]} = 2\left(\dfrac{n-r}{2n-r}\right)$

et donc $2(n-r)P[X=r] = (2n-r)P[X=r+1]$, ce qui démontre (*).

En sommant (*)

$$\sum_{r=0}^{n} 2(n-r)P[X=r] = \sum_{r=0}^{n} (2n+1)P[X=r+1] - (r+1)P[X=r+1]$$

$$2n \sum_{r=0}^{n} P[X=r] - 2 \sum_{r=0}^{n} rP[X=r] = (2n+1) \sum_{r=0}^{n} P[X=r+1] - \sum_{r=0}^{n} (r+1)P[X=r+1]$$

$$2n - 2E(X) = (2n+1)\ (1 - P[X=0]) - E(X)$$

Ainsi $E(X) = (2n+1)\ P[X=0] - 1$

$$E(X) = (2n+1)\ C_{2n}^{n} \left(\frac{1}{2}\right)^{2n} - 1$$

On peut ainsi chercher un équivalent de cette espérance, lorsque n tend vers $+\infty$. Nous avons vu au § 2 que

---

Il s'aperçoit qu'une des poches est vide quand il essaie de prendre la n+1 ième allumette, cf Probabilités pour le Lycée 1 page 22.

$$C_{2n}^{n} \left(\frac{1}{2}\right)^{2n} \cong \frac{1}{\sqrt{\pi n}}$$

$$E(X) \cong \sqrt{\frac{2n}{\pi}}$$

APPENDICE 1

Jean Claude Jovet a montré aux stagiaires comment il utilisait la symbolique des arbres pour aider les élèves. On trouve bien sûr déjà cette idée dans quelques manuels, mais sa démarche a l'avantage d'être <u>systématique</u> et de s'appuyer sur une <u>clarification</u> des notations qui, je suis convaincu, peut intéresser de nombreux professeurs de Lycée et il m'a semblé profitable de reproduire une partie de son document en appendice.

(**********

Pré-requis: Notions élémentaires sur les ensembles.
Notions élémentaires sur les probabilités.

Schématisation d'une probabilité $\left\langle \begin{array}{l} A_1 \\ \text{— } A_i \\ A_n \end{array} \right.$

$\bigcup_{i\in I} A_i=\Omega$

$0\le p_i\le 1$

$\sum_{i\in I} p_i = 1$

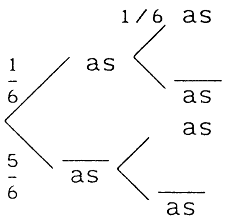
Préambule: Mon expérience pédagogique dans l'initiation des probabilités, m'a convaincu qu'il semblait très bien perçu par les élèves de traduire la conjonction de deux événements par un schéma du type:

$\xrightarrow[\text{relatif à A}]{p_A}$ $\xrightarrow[\text{relatif à B}]{p_B}$ produisant $A\cap B$

$P(A\cap B)= p_A p_B$

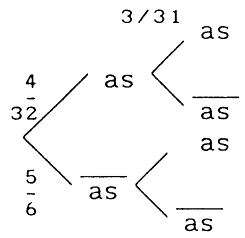
Exemples:

1 On lance deux dés, probabilité d'obtenir deux as?



$P(\text{"2as"}) = \frac{1}{6}\times\frac{1}{6} = \frac{1}{36}$

2 On tire <u>successivement</u> deux cartes d'un jeu de 32, probabilité d'obtenir deux as?



$$P(\text{"2as"}) = \frac{4}{32} \times \frac{3}{31} = \frac{3}{248}$$

On s'aperçoit alors qu'on a utilisé le même graphique pour deux situations différentes, dans l'exemple 1 les événements sont indépendants, alors que, dans le second, les flèches traduisent les probabilités conditionnelles.
Il est donc intéressant de proposer une notation différente pour les deux situations:

Evènements indépendants:

$$\xrightarrow{\quad p_A \quad} A \xrightarrow{\quad p_B \quad} A \cap B$$

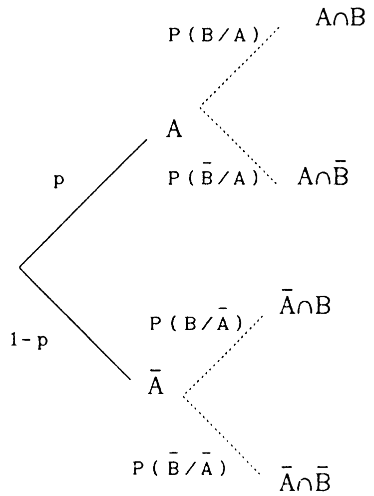
Evènements liés:

$$\xrightarrow{\quad p_A \quad} A \overset{P(B/A)}{\dashrightarrow} A \cap B$$

Remarque: cette notation s'adapte bien également au cas de produits d'espaces de probabilités.

A PROBABILITÉS CONDITIONNELLES.

On a P(A).P(B/A) = P(A∩B)



le schéma établit que P(./A) est une probabilité.

(mise en évidence de Ω et $\sum p_i = 1$)

Propriété: $\qquad P(A)P(B/A) = P(B)P(A/B)$

$$\begin{tikzcd} & & A\cap B \quad = \quad B\cap A & & \\ \cdot \arrow[r, "p_A"] & A \arrow[ur, dotted, "P(B/A)"'] & & B \arrow[ul, dotted, "P(A/B)"] & \cdot \arrow[l, "p_B"'] \end{tikzcd}$$

La généralisation au cas d'un nombre quelconques d'événements:

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1)P(A_2/A_1)\ldots\ldots\ldots\ldots\ldots\ldots P(A_n/A_1\cap A_2\cap\ldots\cap A_{n-1})$$

est facilement perceptible sur le dessin (que je ne fais pas pour des raisons typographiques).

Avec ce formalisme, il est possible de présenter, ou d'interpréter toutes les formules fondamentales des probabilités conditionnelles:

B PROBABILITÉS TOTALES

Pour tout système complet d'événements, $A_i)_{i\in I}$:

$$p(B) = \sum_{i\in I} p(B\cap A_i) = \sum_{i\in I} p(A_i)p(B/A_i)$$

Correspond au schéma suivant:

$$\begin{array}{l} \diagup\ A_1 \\ \text{—}\ A_i \\ \diagdown\ A_n \end{array} \qquad \text{pour indiquer que le système est complet} \qquad \begin{array}{l} \bigcup_{i\in I} A_i=\Omega \\ 0\leq p_i\leq 1 \\ \sum_{i\in I} p_i =1 \end{array}$$

Puis

$$\begin{array}{l} \diagup\ A_1 \xrightarrow{P(B/A_1)} A_1\cap B \\ \text{—}\ A_i \xrightarrow{P(B/A_i)} A_i\cap B \\ \diagdown\ A_n \xrightarrow{P(B/A_n)} A_n\cap B \end{array}$$

"Pour retrouver B il faut mettre en évidence $B\cap(\bigcup_{i\in I} A_i)$, c'est-à-dire $B\cap\Omega$".

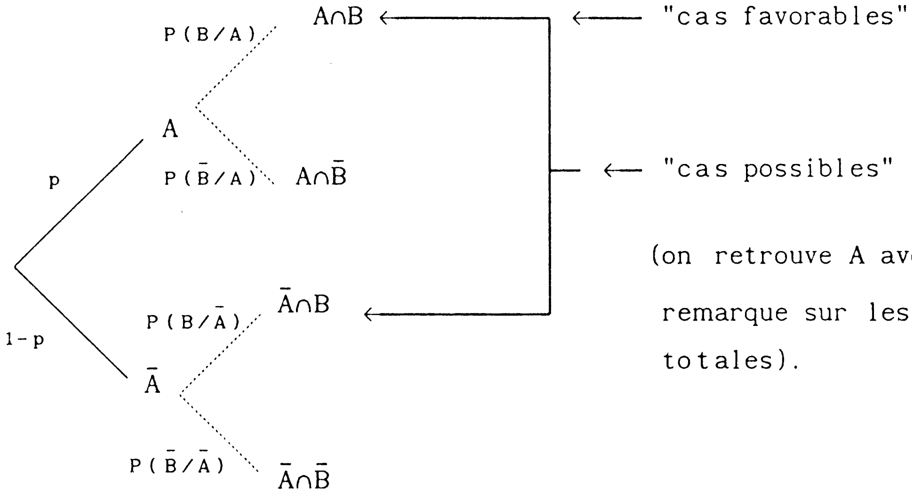
Sur le dessin on "retrouve Ω sur toutes les flèches pointillées conditionnées par A_i".

C FORMULE DE BAYES

Dans les mêmes conditions que dans le paragraphe précédent:

$$p(A_{i_0}/B) = \frac{P(A_i)P(B/A_i)}{\sum_{i\in I} P(A_i)P(B/A_i)}$$

Toujours pour des raisons typographiques, on n'a traité ici que le cas élémentaire.



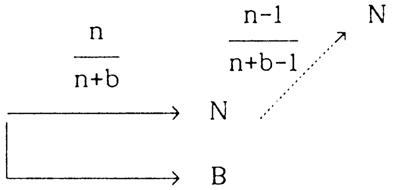
(on retrouve A avec la remarque sur les probabilités totales).

D EXEMPLES

1 J'ai n boules noires et b boules blanches. J'en tire successivement deux sans remise, quelle est la probabilité que les deux soient noires?

Référence: probabilités pour le Lycée 1, page 24.

On élague l'arbre général, pour extraire:

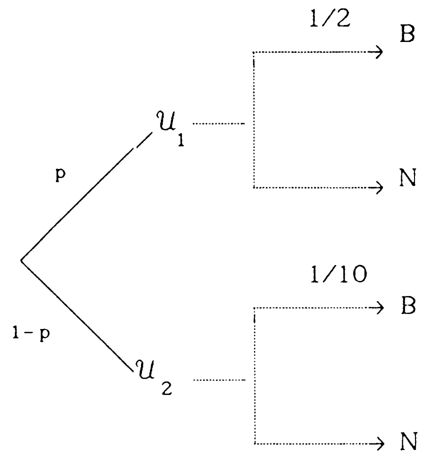


réponse: $\frac{n}{n+b}\,\frac{n-1}{n+b-1}$

2 J'ai deux urnes: $\mathcal{U}_1$ contient 5 boules noires et 5 boules blanches, $\mathcal{U}_2$ contient 1 blanche et 9 noires. La probabilité de tirer dans $\mathcal{U}_1$ est p (resp q pour $\mathcal{U}_2$). Je choisis

d'abord une urne, puis une boule. Quelle est la probabilité que la boule tirée soit blanche?

Référence: probabilités pour le Lycée 1, page 25.



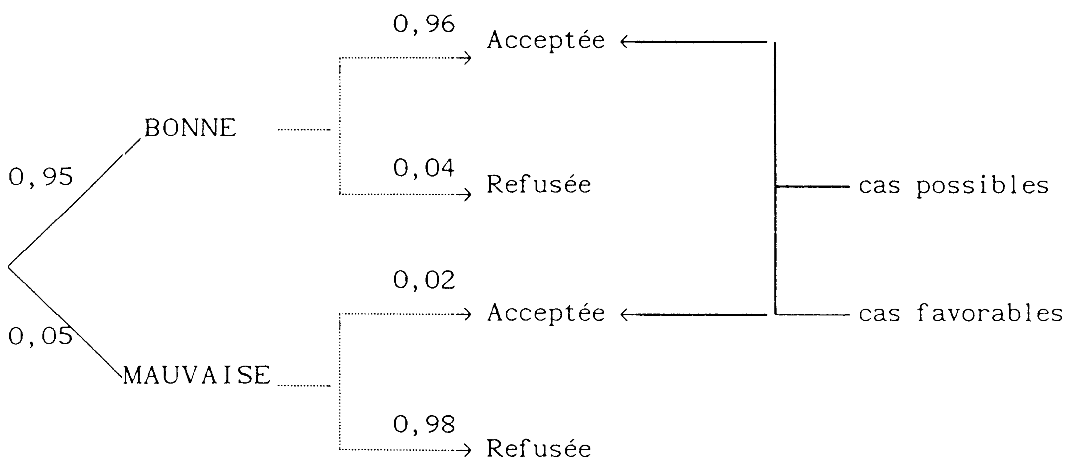
Réponse: $\dfrac{p}{2} + \dfrac{1-p}{10} = \dfrac{1+4p}{10}$

3 Dans un lot de pièces, il y a 5% de pièces défectueuses. Un procédé de contrôle accepte une bonne pièce dans 96% des cas, et refuse une mauvaise pièce dans 98% des cas. On choisit une pièce au hasard, quelle est la probabilité des événements suivants:

1°) il y a une erreur de contrôle

2°) Une pièce acceptée soit mauvaise.
est

Référence: HEC probabilités Breal page 62.



L'erreur de contrôle correspond à la probabilités des événements disjoints Bonne et refusée, et Mauvaise et Acceptée, c'est-à-dire, par lecture directe de ces intersections:

$$0,95 \times 0,04 + 0,05 \times 0,02 = 0,039.$$

La seconde question, correspond à la probabilité conditionnelle, Mauvaise sachant Acceptée.

Il suffit donc de se mettre dans l'univers des pièces Acceptées, et de prendre les cas favorables correspondants pour trouver:

$$\frac{0,05 \times 0,02}{0,05 \times 0,02 + 0,95 \times 0,96} = \frac{1}{913}$$

***********)

Jean Claude Jovet, a poursuivi par d'autres exemples "d'arbres", pour les "remontées dans le temps" et les chaînes de Markov.

APPENDICE 2

Une démonstration, classique, de la formule de Stirling.

L'idée consiste à essayer d'approcher la somme $S_n=\ln(n!)=\sum_{i=1}^{n}\ln(i)$.

Si on utilise, la méthode des trapèzes pour approcher $\int_1^n \ln x\,dx$, on construira la nouvelle somme:

$$\sum_{i=1}^{n-1}\frac{\ln(i)+\ln(i+1)}{2} = S_n - \frac{1}{2}\ln(n).$$

Un calcul direct de l'intégrale donne $n(\ln(n)-1)+1$.

D'où l'on tire la différence des termes prépondérants de ces deux quantités:

$$d_n = S_n - \frac{1}{2}\ln(n)-n(\ln(n)-1)$$

Suite que l'on va exprimer comme une série:

$$\text{si } a_k = d_{k+1}-d_k$$

$$\sum_{k=1}^{n-1} a_k = d_n - d_1$$

Or

$$\begin{aligned}
a_k &= S_{k+1}-S_k-(k+1)(\ln(k+1)-1)+k(\ln(k)-1)-\frac{1}{2}\ln(k+1)+\frac{1}{2}\ln(k)\\
&= \ln(k+1)-k(\ln(k+1)-\ln(k))-\ln(k+1)+1-\frac{1}{2}(\ln(k+1)-\ln(k))\\
&= 1-(\ln(k+1)-\ln(k))(k+\frac{1}{2})\\
&= 1-(\ln(1+\frac{1}{k})(k+\frac{1}{2})\\
&= 1-(k+\frac{1}{2})\left(\frac{1}{k}-\frac{1}{2k^2}+\frac{1}{3k^3}+o\left(\frac{1}{k^3}\right)\right)\\
&= -\frac{1}{12k^2}+o\left(\frac{1}{k^2}\right)
\end{aligned}$$

La série de terme général (a_k) est donc convergente, et la suite (d_n), qui en est une somme partielle, aussi.

Donc $S_n - \frac{1}{2}\ln(n) - n(\ln(n)-1) = \ln\left(\frac{n!}{n^{n+\frac{1}{2}}\, e^{-n}}\right)$ converge vers une limite finie, et par continuité de la fonction exponentielle, la suite $\left(\frac{n!}{n^{n+\frac{1}{2}}\, e^{-n}}\right)$ converge vers d.

Nous avons vu, au chapitre II 9, que l'expression $a_n = C_{2n}^{n} \left(\frac{1}{2}\right)^{2n} \cong \frac{1}{\sqrt{\pi n}}$.

En appliquant la formule de Stirling (partielle), on calcule un autre équivalent de a_n.

$$C_{2n}^{n} \cong \frac{(2n!)}{n!n!} \cong \frac{(2n)^{2n+\frac{1}{2}}}{d(n^{n+\frac{1}{2}})^2}$$

$$a_n = \frac{(2n)^{2n+\frac{1}{2}}}{d(n^{n+\frac{1}{2}})^2} \frac{1}{2^{2n}} \cong \frac{1}{d}\sqrt{\frac{2}{n}}$$

Et donc la valeur de d est $\sqrt{2\pi}$

Finalement:

$$n! \cong \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n}$$

Titre : LES PROBABILITES POUR LE LYCEE – 2 –

Auteur : Luc SINEGRE

Public concerné : Professeurs de Lycée, Bts, prépa.

Niveau : Lycée, BTS, Prépa.

Résumé : On présente les principales lois des variables aléatoires discrètes, en particulier la loi binômiale, on finit par l'étude des temps d'attente et par les applications de la formule de Stirling.

Mots clefs :
- Variable aléatoire.
- Loi binômiale.
- Loi géométrique.
- Loi de Poisson.
- Formule de Stirling.

Date : Juin 1993

Nombre de pages : 56
Format : A 4
Prix : 37.00 F
Publication : IREM DE ROUEN

Dépôt légal : ISBN 2-86239-050 -x

****

BON DE COMMANDE

M. , Mme, Mlle : ____

Adresse : ____

Libellé	Prix	Quantité	Total
LES PROBABILITES POUR LE LYCEE – 2	37.00 F	....	....
Frais d'envoi : 10 F pour le 1er livre et 5 F par livre supplémentaire (France)			....
Frais réels pour l'étranger			....
SOMME DUE :			....

Les chèques de règlement seront libellés à l'ordre de :

L'AGENT COMPTABLE DE L'UNIVERSITE DE ROUEN

Et adressés directement à l'I.R.E.M. – B.P. 153 – 76135 MONT SAINT AIGNAN

Pour tout renseignement complémentaire Tél. : 35.14.61.41.

RIB : TP ROUEN TG 10071 76000 00044004056 81

****

DATE : SIGNATURE :