COMMENT MÉMORISER LES LOIS DE PROBABILITÉ DISCRÈTES USUELLES SANS TROP SE FATIGUER

Edith KOSMANEK (*)

1.— LE TABLEAU DE CHOSSAT

Le tableau de CHOSSAT [1] présente 9 lois discrètes. Nous représentons sur le tableau ci-après les 7 qui sont unidimensionnelles. Pour ces lois, le modèle est celui du tirage dans une urne à 2 catégories de boules, les blanches en proportion p et les noires en proportion $q = 1 - p$.

En croisant 2 critères à 2 modalités chacun :
— mode de tirage : avec ou sans remise
— paramètre fixé : nombre n de tirages ou nombre r d'occurences de blanches
on récolte 4 lois principales : binômiale, hypergéométrique, Pascal avec ou sans remise, lesquelles admettent 3 autres lois comme cas particuliers : uniforme, Bernouilli, géométrique et la loi de Poisson comme loi limite. Des étudiants de 1er cycle, confrontés à ce tableau, émettent spontanément 2 remarques judicieuses :
— il manque la loi de Pascal sans remise (voir ci-dessous),
— pourquoi changer de modèle pour présenter la loi uniforme : boules numérotées au lieu de boules blanches et noires.

Loi	Valeurs possibles	Probabilités
Pascal **sans** remise $S[N, r, p]$	$[r, r+1, \ldots, k, Nq + r]$	$P[X = k] = \frac{C_{Np}^{r-1} C_{Nq}^{k-r}}{C_N^{k-1}} \frac{Np - r + 1}{N - k + 1}$

Espérance	Variance	Modèle
$r\frac{N+1}{Np+1}$	$rq\frac{N(N+1)(Np-r+1)}{(Np+1)^2(Np+2)}$	Dans une urne, deux catégories de boules : proportion p de blanches et q de noires. On tire, sans remise, jusqu'à obtenir r boules blanches. X est la variable aléatoire : nombre de tirages nécessaires.

Observations
On retrouve la loi de Pascal **avec** remise pour $N \to +\infty$ et la loi **uniforme** pour $p = \frac{1}{N}$ et $r = 1$.

(*) À Aimé FUCHS (Strasbourg), mon initiateur aux probabilités.

LOIS DE PROBABILITÉ :

LOI	VALEURS POSSIBLES	PROBABILITÉS DE CES VALEURS	ESPÉRANCE
BERNOULLI $\mathcal{B}(1,p)$	$\{0,1\}$	$P[X=0]=1-p=q$ $P[X=1]=p$ où $0 \leq p \leq 1$	p
BINOMIALE $\mathcal{B}(n,p)$	$\{0,1,\ldots n\}$	$P[X=k]=C_n^k p^k q^{n-k}$ où $0 \leq p \leq 1$ et $p+q=1$	np
GÉOMÉTRIQUE $R(1,p)$	$\{1,2,\ldots,k,\ldots,+\infty\}$	$P[X=k]=q^{k-1}p$ où $0 < p \leq 1$ et $p+q=1$	$\frac{1}{p}$
PASCAL (parfois appelée "*Binomiale négative*") $R(r,p)$	$\{r,r+1,\ldots+\infty\}$	$P[X=k]=C_{k-1}^{r-1}p^r q^{k-r}$ où $0 < p \leq 1$ et $p+q=1$	$\frac{r}{p}$
HYPER-GÉOMÉTRIQUE $H(N,n,p)$	$\sup(0,n-N_2) \leq K \leq \inf(n,N_1)$	$P[X=k]=C_{N_1}^k C_{N_2}^{n-k}/C_N^n$ $(N=N_1+N_2)$	np avec $p=\frac{N_1}{N}$
POISSON $\mathcal{P}(\lambda)$	$\{0,1,\ldots+\infty\}$	$P[X=k]=\frac{\lambda^k}{k!}e^{-\lambda}$	λ
UNIFORME SUR UN ENSEMBLE FINI $U(n)$	$\{1,2,\ldots,n\}$	$P[X=k]=\frac{1}{n}$	$\frac{n+1}{2}$

VARIABLES DISCRÈTES

VARIANCE	MODÈLE	OBSERVATIONS (*)
pq	Dans une urne, 2 catégories de boules : proportion p de boules blanches, q de noires. On effectue un tirage. X est la variable aléatoire "*nombre de boules blanches*".	
npq	Dans une urne, 2 catégories de boules : proportion p de boules blanches, q de noires. On effectue n tirages <u>avec</u> remise. X est la variable aléatoire "*nombre de boules blanches*".	Si $X_1 \rightarrow \mathcal{B}(n_1,p), X_2 \rightarrow \mathcal{B}(n_2,p)$ X_1 et X_2 indépendants alors $X_1 + X_2 \rightarrow \mathcal{B}(n_1+n_2,p)$.
$\frac{q}{p^2}$	Dans une urne, 2 catégories de boules : proportion p de boules blanches, q de noires. On tire <u>avec</u> remise jusqu'à obtenir une boule blanche. X est la variable aléatoire "*nombre de tirages nécessaires*".	
$\frac{rq}{p^2}$	Dans une urne, 2 catégories de boules : proportion p de boules blanches, q de noires. On tire <u>avec</u> remise jusqu'à obtenir r boules blanches. X est la variable aléatoire "*nombre de tirages nécessaires*".	On retrouve la loi géométrique avec $r = 1$. Si $X_1 \rightarrow$ Pascal de paramètre r_1, $X_2 \rightarrow$ Pascal de paramètre r_2. Si X_1 et X_2 sont indépendants, alors $X_1 + X_2$ suit une loi de Pascal de paramètre $r_1 + r_2$.
$npq\frac{N-n}{N-1}$ avec $q = \frac{N_2}{N}$	Dans une urne, 2 catégories de boules : nombre N_1 boules blanches, N_2 de noires. On tire n boules <u>sans</u> remise. X est la variable aléatoire "*nombre de boules blanches*".	On retrouve la loi binomiale à la limite où $N \rightarrow \infty, N_1 \rightarrow \infty, N_2 \rightarrow \infty$ avec $\frac{N_1}{N} \rightarrow p$, $\frac{N_2}{N} \rightarrow q = 1-p$.
λ	Loi limite d'une loi binomiale lorsque ($n > 50$) et $p < 0,10$. Loi des évènements "*rares*".	Si $X_1 \rightarrow \mathcal{P}(\lambda_1), X_2 \rightarrow \mathcal{P}(\lambda_2), X_1$ et X_2 sont indépendants alors : $X_1 + X_2 \rightarrow \mathcal{P}(\lambda_1+\lambda_2)$.
$\frac{n^2-1}{12}$	Dans une urne, n boules numérotées de 1 à n. On effectue un tirage, avec équiprobabilité de tirer chaque boule. X est la variable aléatoire "numéro de la boule tirée"	

(*) $X \rightarrow A$ signifie : X suit la loi de probabilité A.

Les résultats (page 1), obtenus sans trop de difficultés, répondent aux deux questions et manifestent le lien qui existe entre elles : la loi uniforme apparaît comme cas particulier de la loi Pascal sans remise et peut donc, elle aussi, se présenter à partir du modèle standard : on tire sans remise dans une urne contenant une seule blanche et $(N-1)$ noires jusqu'à l'occurence de cette unique blanche.

2.— L'ARBORESCENCE KOSMANIENNE

Le schéma unitaire ci-contre (p. 5) résume les considérations précédentes sous forme de graphe arborescent ; il met bien en évidence l'homogénéïté du mécanisme générateur des différentes lois.

3.— UN LOSANGE COMMUTATIF

En homogénéïsant les notations de façon évidente, on observe les particularités suivantes : on passe de la loi binômiale à la loi Pascal avec remise simplement en multipliant par le taux de "*succès*" $\frac{K}{n}$ (nombre d'occurences de blanches divisé par nombre total de tirages) :

$$C_n^K p^k q^{n-K}.\frac{K}{n} = C_{n-1}^{K-1} p^K q^{n-K}.$$

Idem pour le passage de la loi hypergéométrique à la loi Pascal sans remise.

$$\frac{C_{Np}^K C_{Nq}^{n-K}}{C_N^n}.\frac{K}{n} = \frac{C_{Np}^{K-1} C_{Nq}^{n-K}}{C_N^{n-1}}.\frac{Np-K+1}{N-n+1}$$

propriété généralisée au paragraphe 4.

L'arborescence précédente peut alors, compte tenu des observations du tableau de Chossat, se présenter sous forme de "*losange commutatif*" à rallonges (voir p. 6). Il suffit de **mémoriser une seule formule**, celle de la loi hypergéométrique ; toutes les autres s'en déduisent par trois opérations simples :
— multiplication par le taux de succès $\frac{K}{n}$;
— passage à la limite : $N \to +\infty, n \to +\infty, np \to \lambda$;
— particularisation : $n = 1, K = 1, p = \frac{1}{N}$.

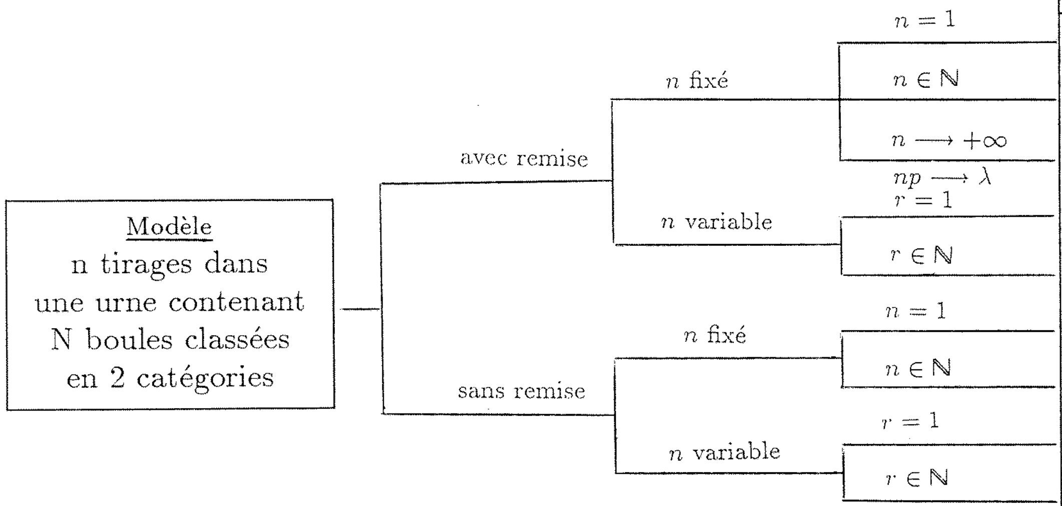
4.— UNE GÉNÉRALISATION (*)

Notations :

- (A_i) : suite d'événements associée à une suite d'épreuves,
- 1_{A_i} : variable aléatoire indicatrice de l'événement A_i,
- $S_n = \sum_{i=1}^n 1_{A_i}$: nombre d'événements réalisés en n épreuves,
- $T_K = \text{Min } \{n | S_n = K\}$: nombre d'épreuves nécessaires à l'occurence de K événements.

(*) Avec l'aide de P.-L. HENNEQUIN.

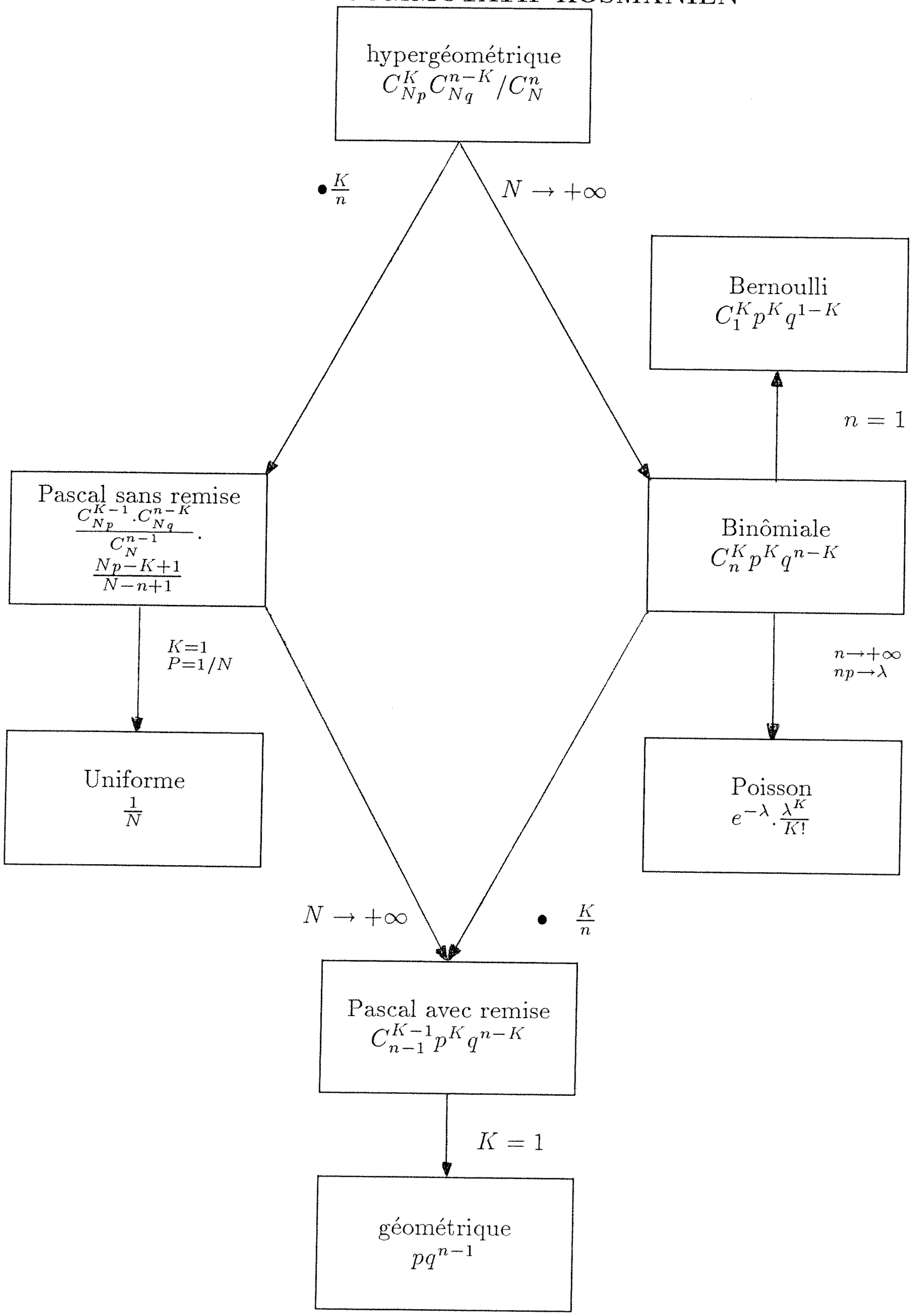
SCHÉMA ARBORESCENT RELATIF AUX VARIABLES ALÉATOIRES DISCRÈTES USUELLES À UNE DIMENSION.



Variable X	Notation	Appellation
Nombre de blanches	$B(1,p)$	Bernoulli
Nombre de blanches	$B(n,p)$	Binômiale
Nombre de blanches	$P(\lambda)$	Poisson
Nombre de tirages	$R(1,p)$	Géométrique
Nombre de tirages	$R(r,p)$	Pascal
Nombre de blanches	$H(N,1,p) = B(1,p)$	Bernoulli
Nombre de blanches	$H(N,n,p)$	Hypergéométrique
Nombre de tirages	$U(N) = S(N,1,\frac{1}{N})$	Uniforme
Nombre de tirages	$S(N,r,p)$	Pascal sans remise

ARBORESCENCE KOSMANIENNE

LOSANGE COMMUTATIF KOSMANIEN



Hypothèse

$$P(A_i|S_n = K) = C \qquad \forall i = 1, 2, \ldots, n.$$

Proposition

$$\boxed{P(T_K = n) = P(S_n = K) \bullet \frac{K}{n}, 1 \leq K \leq n.}$$

Démonstration

On montre d'abord que la constante C est égale au taux de succès : $\frac{K}{n}$; en effet :

$$\sum_{i=1}^{n} P(A_i|S_n = K) = nC = \sum_{i=1}^{n} E(1_{A_i}|S_n = K) =$$

$$= E(\sum_{i=1}^{n} 1_{A_i}|S_n = K) = E(S_n|S_n = K) = K \Longrightarrow C = \frac{K}{n}.$$

Il est clair, par ailleurs, que l'événement $(T_K = n)$ est équivalent à l'événement : $(S_{n-1} = K - 1$ et $S_n = K)$ donc $P(T_K = n) = P(S_n = K)P(S_{n-1} = K - 1|S_n = K) = P(S_n = K)P(A_n|S_n = K) = P(S_n = K) \bullet C = P(S_n = K) \bullet \frac{K}{n}$.

Application

Considérons comme suite d'épreuves celle du modèle des urnes. L'événement A_i est l'occurence d'une blanche $\forall i$. Que le tirage soit avec ou sans remise, l'hypothèse $P(A_i|S_n = K) = C$ est valable.

Pour le tirage avec remise :
— S_n suit la loi binômiale,
— T_K suit la loi Pascal avec remise.

Pour le tirage sans remise :
— S_n suit la loi hypergéométrique,
— T_K suit la loi Pascal sans remise.

On retrouve donc le losange.

Bibliographie

[1] Aide-mémoire de l'étudiant en économie et gestion (Tableau de Chossat), éd. Bordas-Dunod
[2] KOSMANEK E. : Simple remarque à propos d'un tableau d'aide-mémoire (Bull. A.P.M.E.P., sept. 1989)
[3] KOSMANEK E. : Y a-t-il un probabiliste dans l'avion? (Revue Quadrature n° 4, 1990)