252 : Loi binomiale - Loi de Poisson. Applications

Les variables aléatoire (v.a.) considérées dans la leçon sont implicitement définies sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On notera d'autre part i.i.d pour « indépendantes et identiquement distribuées ».

I - Loi binomiale

A) Définition et premières propriétés

[FF1] p68 **Définition** Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. La loi binomiale de paramètre (n,p) est la mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie par : $\mathcal{B}(n,p) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$. Pour $n=1$, on la notera $\mathcal{B}(p)$, loi de Bernoulli de paramètre p.

[OUV1] p70 **Proposition 2** $\mathcal{B}(n_1,p) * \mathcal{B}(n_2,p) = \mathcal{B}(n_1+n_2,p)$. Ainsi, si $X_1 \sim \mathcal{B}(n_1,p)$ et $X_2 \sim \mathcal{B}(n_2,p)$ sont indépendantes, alors $X_1 + X_2 \sim \mathcal{B}(n_1+n_2,p)$

$\longrightarrow$ La loi de Bernoulli modélise une expérience aléatoire à deux issues possibles, un succès et un échec. La loi binomiale est alors la loi du nombre de succès après n répétions indépendantes de cette expérience

Exemple 3 (tirage avec remise) Si la proportion de boules blanches dans une urne est p, la loi du nombre de boules blanches obtenues après n tirages avec remise est $\mathcal{B}(n,p)$.

Exemple 4 (tomber dans un borélien) Soient $(X_1, .., X_n)$ des v.a.r i.i.d et $A \in \mathcal{B}(\mathbb{R})$. Alors les v.a $\mathbb{1}_{\{X_1 \in A\}}, .., \mathbb{1}_{\{X_n \in A\}}$ sont i.i.d de loi $\mathcal{B}(\mathbb{P}(X_1 \in A))$ et $S_n = \sum_{i=1}^{n} \mathbb{1}_{\{X_i \in A\}} \sim \mathcal{B}(n, \mathbb{P}(X_1 \in A))$

Proposition 5 Si $X \sim \mathcal{B}(n,p)$ alors $\mathbb{E}[X] = np$ $\quad Var(X) = np(1-p)$ et $\varphi_X(t) := \mathbb{E}[e^{itX}] = (1-p+pe^{it})^n$

B) Théorèmes limites

Théorème 6 (loi faible des grands nombres) Si $X \sim \mathcal{B}(n,p)$ alors $\mathbb{P}(|\frac{X}{n} - p| > \varepsilon) \leq 1/4n\varepsilon^2$

DEV [Z-Q] p517 **Application 7** Approximation des fonctions continues sur $[0,1]$ par les polynômes de Bernstein

Proposition 8 Soit $(X_n)_{n \in \mathbb{N}}$ v.a iid de loi $\mathcal{B}(p)$ et $S_n = \sum_{i=1}^{n} X_i$. Alors

$$\frac{S_n}{n} \xrightarrow[n \to +\infty]{p.s} p \quad \text{et} \quad \sqrt{n}\left(\frac{S_n}{n} - p\right) \xrightarrow[n \to +\infty]{loi} \mathcal{N}(0, p(1-p))$$

$\longrightarrow$ $\frac{S_n}{n}$ est donc un estimateur fortement consistant et asymptotiquement normal (à vitesse $\sqrt{n}$) du paramètre p.

$\longrightarrow$ le caractère asymptotiquement normal est un cas particulier du théorème de la limite centrale, le théorème de Moivre - Laplace

C) Loi multinomiale

Soit $k \in \mathbb{N}^*$. Pour $n \in \mathbb{N}^*$, on considère une partition $(A_i^n)_{1 \leq i \leq k}$ de Ω. On suppose que les familles indexées sur n des éléments de ces partitions sont indépendantes. On suppose de plus que $\forall n \in \mathbb{N}^*$, $\forall i \in [\![1,k]\!]$, $\mathbb{P}(A_i^n) = p_i$ où $p_i > 0$ et $\sum_{i=1}^{k} p_i = 1$. On définit alors les v.a : $X^n = \begin{pmatrix} \mathbb{1}_{A_1^n} \\ \vdots \\ \mathbb{1}_{A_k^n} \end{pmatrix}$ et $Y^n = \sum_{j=1}^{n} X^j$ [OUV2] p79

Définition 9 La loi de Y^n est appelée loi multinomiale de paramètre $(n, p_1, ..., p_k)$.

Proposition 10 En notant $y = {}^t(y_1, ..., y_k) \in D_n = \{y \in [\![0,n]\!], \sum_{i=1}^{k} y_i = n\}$, on a $\mathbb{P}(Y^n = y) = \frac{n!}{y_1! \cdots y_k!} p_1^{y_1} \cdots p_k^{y_k}$. D'autre part, $\mathbb{E}[Y^n] = {}^t(np_1, ..., np_k)$ et $C_{Y^n} = (a_{ij})_{1 \leq i,j \leq n}$ avec $\begin{cases} a_{i,i} = np_i(1-p_i) \\ a_{i,j} = -np_ip_j \text{ si } i \neq j \end{cases}$

Exemple 11 Si une urne contient r_i boules de couleur c_i et si $p_i = r_i / (r_1 + \cdots + r_k)$ alors le vecteur aléatoire constitué du nombre de boules de chaque couleur obtenues après n tirages avec remise suit la loi multinomiale de paramètre $(n, p_1, ..., p_k)$ [FF1] p65

$\longrightarrow$ La loi multinomiale généralise à $\mathbb{N}^k$ la loi binomiale (chaque marginale d'une multinomiale suit une loi binomiale).

Théorème 12 (Karl-Pearson) Avec les notations de la définition 9, on pose, pour $j \in [\![1,k]\!]$, $N_j^n = \sum_{p=1}^{n} \mathbb{1}_{A_j^p}$ et $\chi^2_{k,n} = \sum_{j=1}^{k} \frac{(N_j^n - np_j)^2}{np_j}$. Alors $(\chi^2_{k,n})_{n \in \mathbb{N}}$ converge en loi vers la loi du chi-deux à $k-1$ degrés de liberté. [OUV2] p315

Application 13 tests du chi-deux

D) Loi binomiale négative (loi de Pascal)

Définition 14 Une v.a entière X suit une loi binomiale négative $\mathcal{B}^-(n,p)$ si $\mathbb{P}(X=k) = \begin{cases} 0 \text{ si } k \leq n-1 \\ \binom{k-1}{n-1} p^n (1-p)^{k-n} \text{ si } k \geq 1 \end{cases}$ [OUV1] p79

[OUV1] p79 **Proposition 15** Soient $(X_1,\dots,X_n)$ i.i.d de loi géométrie $G(p)$ (i.e $\mathbb{P}(X_1=k)=p(1-p)^{k-1}$ $k\in\mathbb{N}^*$) alors $X_1+X_2+\dots+X_n \sim \mathcal{B}^-(n,p)$

$\longrightarrow$ La loi binomiale négative est la loi du nombre de tirages effectués avant n succès (tirages avec remise)

Proposition 16 Si $X\sim\mathcal{B}^-(n,p)$ alors $\mathbb{E}[X]=n/p$ $Var(X)=n(1-p)/p^2$ et $\varphi_X(t)=\left(\frac{pe^{it}}{1-(1-p)e^{it}}\right)^n$

E) Loi hypergéométrique

[PP1] p68 **Définition 17** Soient n,M,N trois entiers positifs avec $n\leq N$ et $M\leq N$. La loi hypergéométrique de paramètres (n,M,N) est la mesure de probabilité $H(n,M,N)=\sum_k \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\delta_k$ où k est tel que $\max(0,n-(N-M))\leq k\leq \min(N,M)$

$\rightarrow$ Si on effectue n tirages sans remise dans une urne contenant N boules, dont M blanches, le nombre de boules blanches obtenues suit la loi $H(n,M,N)$

$\longrightarrow$ Intuitivement, si M et N sont grands devant n, il y a peu de différence entre un tirage sans remise et un tirage avec remise. Le théorème 19 va conforter cette intuition :

Proposition 18 Si $X\sim H(n,N,M)$ alors $\mathbb{E}[X]=n\frac{M}{N}$

$Var(X)=n\,\frac{N-M}{N-1}\cdot\frac{M}{N}\left(1-\frac{M}{N}\right)$

Théorème 19 Soit $n\in\mathbb{N}^*$ et deux suites $(r_j^1)_{j\in\mathbb{N}^*}$ et $(r_j)_{n\in\mathbb{N}^*}$ d'entiers croissantes tendant vers $+\infty$ et telles que $\lim\limits_{j\to+\infty}\frac{r_j^1}{r_j}=p\in]0,1[$.

Alors si $X_j\sim H(n,r_j,r_j^1)$, $(X_j)_j$ converge en loi vers la loi binomiale $\mathcal{B}(n,p)$.

Application 20 (compter les poissons) Un étang contient un nombre inconnu $N\geq1$ de poissons. On effectue une première pêche de $r\geq1$ poissons que l'on marque avant de les relacher. On procède à une seconde pêche de $n\geq1$ poissons qui compte un nombre aléatoire X de poissons marqués. Alors $X\sim H(n,N,r)$ et on peut estimer N par la méthode du maximum de vraisemblance.

II - Loi de Poisson

A) Définition et premières propriétés

[PP1] p72 **Définition 21** Soit $\lambda\in\mathbb{R}_+^*$. La loi de Poisson de paramètre λ est la mesure définie par $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ par $\mathcal{P}(\lambda)=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}\delta_k$

Proposition 22 Si $X\sim\mathcal{P}(\lambda)$ alors $\mathbb{E}[X]=Var(X)=\lambda$ et $\varphi_X(t)=e^{\lambda(e^{it}-1)}$

Exemple 23 La loi de Poisson modélise des phénomènes de comptage : nombre de désintégration radioactives se produisant dans un intervalle de temps, nombre de gouttes de pluie touchant une surface donnée, ... [OUV1] p243

Proposition 24 $\mathcal{P}(\lambda_1)*\mathcal{P}(\lambda_2)=\mathcal{P}(\lambda_1+\lambda_2)$ [OUV1] p65

Application 25 Calcul de $\lim\limits_{n\to+\infty}e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}$ à l'aide du théorème de la limite centrale. [OUV1] p236

B) Lien avec la loi binomiale : évènements rares

Théorème 26 Soit $(p_n)_{n\in\mathbb{N}}\in[0,1]^{\mathbb{N}^*}$ telle que $\lim\limits_{n\to+\infty}np_n=\lambda>0$. Si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de v.a telle que $X_n\sim\mathcal{B}(n,p_n)$ alors $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$. [OUV2] p310 [OUV1] p226

$\longrightarrow$ le théorème fournit donc une approximation de la loi binomiale pour p petit et n grand (en pratique, utilisé pour $n\geq30$ et $p<0,1$). Ainsi, si $X\sim\mathcal{B}(n,\lambda)$ avec λ petit, X suit approximativement une loi $\mathcal{P}(\lambda)$. Pour cette raison, la loi de Poisson est parfois appelée « loi des évènement rares ». Le théorème 26 se généralise comme suit :

Théorème 27 (des évènements rares) Soit pour tout $n\in\mathbb{N}^*$ une famille finie $\{A_{n,j},\ 1\leq j\leq M_n\}$ d'évènements indépendants. On pose $\mathbb{P}(A_{n,j})=p_{n,j}$ et $S_n=\sum_{j=1}^{M_n}A_{n,j}$. On suppose que $(M_n)_n$ tend en croissant vers $+\infty$, que $\max\limits_j p_{n,j}\xrightarrow{n\to+\infty}0$ et que $\sum_{j=1}^{M_n}p_{n,j}\to\lambda>0$

Alors la suite $(S_n)_{n\in\mathbb{N}}$ converge en loi vers $\mathcal{P}(\lambda)$.

C) Lois de Poisson composées

Définition 28 Si $(X_n)_{n\in\mathbb{N}}$ est une suite de v.a i.i.d et si $N\sim\mathcal{P}(\lambda)$ avec $N\perp\!\!\!\perp(X_n)_{n\in\mathbb{N}}$, alors $Y=\sum_{j=1}^{N}X_j$ suit une loi de Poisson composée [OUV1] p84

Proposition 29 Avec ces notations si $X_i\sim\mathcal{B}(p)$ alors $Y\sim\mathcal{P}(\lambda p)$

Exemple 30 On considère une route à sens unique se séparant entre deux voies, A et B. On suppose que le nombre de véhicules arrivant à l'embranchement durant un temps T donné suit une loi $\mathcal{P}(\lambda)$ et que chaque véhicule peut emprunter la voie A avec une probabilité p (et la voie B avec probabilité $1-p$), indépendamment des autres véhicules.

Alors le nombre de véhicules à partir sur la voie A durant T suit la loi $\mathcal{P}(\lambda p)$.

③

⊕ ou remplacement : • développer plus les I et II
• ⊕ sur les lois de Poisson composées
• Markov, Ehrenfest

• files d'attente (...)
• Marches aléatoires simples (ballot theorem / th. des scrutins)

III - Processus de Poisson [si on fait des prob... – c'est hors-programme –]

A/ Définition et premières propriétés

[FF2] p. 27 On cherche à étudier la répartition dans le temps d'instants auxquels se produisent des événements spécifiques, comme l'émission d'une particule radioactive, l'arrivée d'un client à un guichet, ... On appellera ces événements des tops.

Définition 30 On appelle processus de comptage un processus $\{N(t), t\geq 0\}$ tel que (i) $\forall t\geq 0$, $N(t) \in \mathbb{N}$
(ii) $t \mapsto N(t)$ est croissante

→ $N(t)$ représente alors le nombre de tops se produisant dans l'intervalle $[0,t]$, et si $b>a\geq 0$, $N(b)-N(a)$ est le nombre de tops dans $]a,b]$.

Soient $I_1,\dots,I_k$ des intervalles et $N_1,\dots,N_k$ le nombre de tops dans chacun d'eux. On effectue sur les $I_1,\dots,I_k$ une même translation pour obtenir des intervalles $I'_1,\dots,I'_k$ dont les nombres de tops sont $N'_1,\dots,N'_k$. $(N_1,\dots,N_k)$ et $(N'_1,\dots,N'_k)$ sont des v.a dans $\mathbb{N}^k$ de lois respectives $\mathcal{L}$ et $\mathcal{L}'$.

Définition 31 $\{N(t), t\geq 0\}$ est dit à accroissements stationnaires si pour tout k, quelle que soit la suite $(I_1,\dots,I_k)$ et quelle que soit la translation, $\mathcal{L}=\mathcal{L}'$.

Définition 32 Un processus de comptage est dit à accroissements indépendants si le nombre de tops se produisant dans des intervalles disjoints sont indépendants. Il est dit localement continu en probabilité si $\forall t\geq 0$, $\lim_{h\to 0} \mathbb{P}(N(t+h)-N(t)\geq 1)=0$

Définition 33 (processus de Poisson)
Un processus de comptage $\{N(t), t\geq 0\}$ est dit de Poisson si
(i) $N(0)=0$ (ii) il est à accroissements indépendants
(iii) $\forall s\geq 0, \forall t\geq 0$, $N(s+t)-N(s) \sim \mathcal{P}(\lambda t)$

Exemple 34 Soit $(X_n)_{n\in\mathbb{N}}$ v.a.r i.i.d de loi $\mathcal{E}(\lambda)$, $\lambda>0$
[COT] p. 97 On pose $S_0=0$ et $S_n = X_1+\dots+X_n$ pour $n\geq 1$. Pour $t>0$, on pose $N(t)=\sum_{n\geq 1} \mathbb{1}_{\{S_n\leq t\}}$. Alors $\{N(t), t\geq 0\}$ est un processus de Poisson

Proposition 35 Un processus de Poisson est [FF2] p. 28
- à accroissements stationnaires
- localement continu en probabilité

Si $\{N(t), t\geq 0\}$ alors
$\forall t>0$, $\mathbb{P}(N(t)=\infty)=0$ $\qquad \dfrac{\sigma(N(t))}{\mathbb{E}[N(t)]} = \dfrac{1}{\sqrt{\lambda t}}$

$\dfrac{N(t)}{t} \xrightarrow[t\to+\infty]{p.s} \lambda$

→ Le paramètre $\lambda = \dfrac{\mathbb{E}[N(t)]}{t}$ est la fréquence moyenne des tops.

B/ Caractérisation d'un processus de Poisson

Théorème 36 Soit $\{N(t), t\geq 0\}$ un processus de comptage vérifiant (i) $N(0)=0$ [FF2] p. 41
(ii) le processus est à accroissements stationnaires
(iii) le processus est à accroissements indépendants
(iv) le processus est localement continu en probabilité.
Alors c'est un processus de Poisson

Dev.

C/ Temps d'attente

Soit $\{N(t), t\geq 0\}$ un processus de Poisson. On pose
T_1 = temps d'attente du 1er top
$n\geq 1$: T_n = délai entre le $(n-1)$ième et le n-ième top
$S_0=0$ $\quad n\geq 1$: $S_n = T_1+\dots+T_n$: temps du n-ième top

Théorème 37 [FF2] p. 31
(i) $(T_n)_{n\geq 1}$ est une suite de va i.i.d de loi $\mathcal{E}(\lambda)$, $\lambda>0$
(ii) $S_n \sim \Gamma(n,\lambda)$

→ on peut ainsi proposer une définition alternative et constructive des processus de Poisson

[FF1] : Foata-Fuchs « Calcul des probabilités »
[FF2] : " « Processus stochastiques »
[OUV1] : Ouvrard « Probabilités 1 »
[OUV2] : " « Probabilités 2 »
[COT] : Cottrell, « Exercices de probabilités »