Probabilités Discrètes

Thème 1 : Lois Usuelles

1 La loi zêta

Pour tout entier $n \geq 1$, on désigne par A_n l'ensemble des multiples de n.
On note (p_n) la suite des nombres premiers.
On fixe un réel $s > 1$.
Une variable aléatoire X à valeurs dans $\mathbb{N}^*$ suit la loi zêta de paramètre s si $\forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = \dfrac{1}{\zeta(s).n^s}$

1- Justifier qu'il s'agit bien d'une probabilité.
2- Trouver espérance et variance de X.
3- Calculer $\mathbb{P}(A_n)$ pour tout $n \geq 1$.
4- Montrer que la famille (A_p), où p décrit l'ensemble des nombres premiers $\mathcal{P}$, est indépendante.
5- En déduire que $\mathbb{P}(1) = \lim\limits_n \prod\limits_{k=1}^{n} \left(1 - \dfrac{1}{p_k^s}\right)$ En résumé : $\mathbb{P}(1) = \prod\limits_{p \in \mathcal{P}} \left(1 - \dfrac{1}{p^s}\right)$
6- Montrer l'identité d'Euler : $\zeta(s) = \prod\limits_{p \in \mathcal{P}} \left(1 - \dfrac{1}{p^s}\right)^{-1}$
7- Montrer que $\sum \frac{1}{p_n}$ diverge.

2 Une démonstration probabiliste du théorème de Stone-Weierstrass

Soit $f : [0,1] \to \mathbb{R}$ une fonction continue et $n \in \mathbb{N}^*$.

1. Soit S_n une variable aléatoire suivant une loi binomiale de paramètres n et x, $x \in [0,1]$, on pose, pour tout $n \in \mathbb{N}^*$, $X_n = \dfrac{S_n}{n}$.
   (a) Déterminer $E(X_n)$ et $V(X_n)$ respectivement l'espérance et la variance de X_n.
   (b) Justifier que, pour tout $\delta > 0$, $P(|X_n - x| \geq \delta) \leq \dfrac{1}{4n\delta^2}$.
2. On introduit la variable aléatoire $Y_n = f(X_n)$ et on pose pour tout $x \in [0,1]$, $C_n(f)(x) = E(Y_n)$. Pour la suite de cette question, on se donne un réel $\varepsilon > 0$.
   (a) Vérifier que $x \mapsto C_n(f)(x)$ est une fonction polynomiale définie sur $[0,1]$.
   (b) D'après le théorème de Heine, comme f est continue sur $[0,1]$, alors il existe $\beta > 0$ tel que, pour tout $(x_1, x_2) \in [0,1] \times [0,1]$, $|x_1 - x_2| \leq \beta \Rightarrow |f(x_1) - f(x_2)| \leq \dfrac{\varepsilon}{2}$. (On ne vous demande pas de redémontrer ce résultat).
      i. Montrer que $\left| \sum\limits_{\left|\frac{k}{n} - x\right| \leq \beta} \left(f\left(\dfrac{k}{n}\right) - f(x)\right) P\left(X_n = \dfrac{k}{n}\right) \right| \leq \dfrac{\varepsilon}{2}$
      ii. Montrer que $\left| \sum\limits_{\left|\frac{k}{n} - x\right| > \beta} \left(f\left(\dfrac{k}{n}\right) - f(x)\right) P\left(X_n = \dfrac{k}{n}\right) \right| \leq \dfrac{M}{2n\beta^2}$, avec $M = \sup\limits_{t \in [0,1]} |f(t)|$.
   (c) En déduire que la suite $(C_n(f))_{n \geq 1}$ converge uniformément vers f sur $[0,1]$.

3 La loi binomiale négative

Soit $r \in \mathbb{N}^*$ et $p \in]0,1[$.

On répète une suite d'expériences de Bernoulli indépendantes de paramètre p. On définit une variable aléatoire X qui est le nombre d'échecs avant d'arriver à r succès.

1- Déterminer la loi de X.
2- Déterminer la fonction génératrice G_X. 3- Déterminer $E(X)$ et $V(X)$.

Thème 1 : Théorèmes de convergence et d'approximation

I Inégalité de Bienaymé-Tchebychev

Théorème 1
Soit X une VAR admettant un moment d'ordre 2. Alors :

$$\forall \varepsilon > 0, \quad P\left(|X - E(X)| \geqslant \varepsilon\right) \leqslant \frac{V(X)}{\varepsilon^2}$$

Démontrer le théorème cité ci dessus

II Convergence d'une suite de VAR

1 *Loi faible des grands nombres*

Théorème 2
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de VAR indépendantes, admettant une même espérance m et une même variance σ^2. On pose $\overline{X}_n = \dfrac{X_1 + \cdots + X_n}{n}$. Alors on a :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P\left(|\overline{X}_n - m| \geqslant \varepsilon\right) = 0$$

On dit que $\overline{X}_n$ converge en probabilité vers la variable constante égale à m.

Démontrer le théorème cité ci dessus

2 *Convergence en loi*

Définition 1
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de VAR. On note F_{X_n} la fonction de répartition de X_n.
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ **converge en loi** vers la VAR X, de fonction de répartition F_X, si pour tout $x \in \mathbb{R}$ **où F_X est continue**, la suite $(F_{X_n}(x))_{n\in\mathbb{N}}$ converge vers $F_X(x)$.

Propriété 1
Soit (X_n) une suit de VAR convergeant en loi vers la VAR X. Pour tous points a et b de continuité de F_X tels que $a < b$, on a :

$$\lim_{n\to+\infty} P(a < X_n \leqslant b) = P(a < X \leqslant b)$$

Démontrer le théorème cité ci dessus

Théorème 3
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de VAR **discrètes** et X une VAR **discrète** telles que pour tout entier n, $X_n(\Omega) \subset \mathbb{N}$ et $X(\Omega) \subset \mathbb{N}$.
Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X ssi :

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(X_n = k) = P(X = k)$$

Démontrer le théorème cité ci dessus

3 Théorème de la limite centrée

Théorème 4
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires réelles définies sur une même espace probabilisé et **indépendantes**. On suppose que les X_n ont toutes la même loi et qu'elles admettent une variance non nulle.
On note :
$$E(X_1) = m \quad V(X_1) = \sigma^2$$
$$\forall n \in \mathbb{N}^*, \quad S_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad Y_n = \frac{S_n}{n}$$
On a donc $S_n^* = \dfrac{S_n - nm}{\sigma\sqrt{n}} = \dfrac{\sqrt{n}(Y_n - m)}{\sigma} = Y_n^*$
Alors la suite (S_n^*) (et donc la suite (Y_n^*)) converge en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$.

Démontrer le théorème cité ci dessus

III Approximation de variables aléatoires

1 Approximation d'une loi hypergéométrique par une loi binomiale

Théorème 5
Soient n un entier fixé et $p \in]0;1[$. On pose $I = \{N \in \mathbb{N} / Np \in \mathbb{N}\}$. Soit $(X_N)_{N\in I}$ une suite de VAR telle que pour tout N, X_N suit la loi hypergéométrique $\mathcal{H}(N,n,p)$. Alors (X_N) converge en loi vers une VAR X qui suit la loi binomiale $\mathcal{B}(n,p)$.

2 Approximation d'une loi binomiale par une loi de Poisson

Théorème 6
Soit λ un réel strictement positif et $(X_n)_{n\in\mathbb{N}}$ une suite de VAR discrètes telles que X_n suit la loi binomiale de paramètre $\left(n, \dfrac{\lambda}{n}\right)$. Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une VAR X qui suit une loi de Poisson de paramètre λ.

3 Approximation d'une loi binomiale par une loi normale

Théorème 7
Soit $p \in]0;1[$, $q = 1-p$ et $(S_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires telle que S_n suit la loi binomiale $\mathcal{B}(n,p)$. Alors la suite de variable aléatoire $S_n^* = \dfrac{S_n - np}{\sqrt{npq}}$ converge en loi vers la loi normale centrée réduite.

Démontrer le théorème cité ci dessus

4 Approximation d'une loi de Poisson par une loi normale

Théorème 8
Soit α un réel strictement positif et $(S_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires telle que S_n suit la loi de Poisson $\mathcal{P}(n\alpha)$. Alors la suite de variable aléatoire $S_n^* = \dfrac{S_n - n\alpha}{\sqrt{n\alpha}}$ converge en loi vers la loi normale centrée réduite.

Corrigé

Thème 1 : Lois Usuelles

1 La loi zêta

1- Il est clair que

$$\sum_{n=1}^{\infty} \frac{1}{\zeta(s).n^s} = 1$$

2- Si $s > 2$:

$$E(X) = \frac{\zeta(s-1)}{\zeta(s)}$$

Si $s > 3$:

$$V(X) = \frac{\zeta(s-2)}{\zeta(s)} - (E(X))^2$$

3-

$$\mathbb{P}(A_n) = \frac{1}{n^s}$$

4- Soit $r \geq 1$, $q_1, ..., q_r$ des nombres premiers distincts ; notons

$$B_j = A_{q_j}$$

L'intersection

$$B = \bigcap_{j=1}^{r} B_j$$

est l'ensemble des multiples de $n = q_1...q_r$; avec 2 :

$$\mathbb{P}(B) = \frac{1}{n^s} = \prod_{j=1}^{r} \frac{1}{q_j^s} = \prod_{j=1}^{r} \mathbb{P}(B_j)$$

5- Soit $p_1, ..., p_r$ les r premiers nombres premiers ; soit C_r l'ensemble des entiers $n \geq 1$ qui ne sont divisibles par aucun des nombres $p_1, ..., p_r$.

$$\mathbb{P}(C_r) = \prod_{j=1}^{r} \left(1 - \frac{1}{p_j^s}\right)$$

car c'est une intersection d'évènements mutuellement indépendants.

$$\lim_{r} \mathbb{P}(C_r) = \mathbb{P}(1)$$

d'après le théorème de continuité décroissante.

6- Découle de

$$\mathbb{P}(1) = \frac{1}{\zeta(s)}$$

7- Supposons que $\sum \frac{1}{p_n}$ converge. Alors $\sum \ln\left(1 - \frac{1}{p_n}\right)$ converge. Donc :

$$\forall s > 1,\ \ln \zeta(s) = -\sum_{k=1}^{\infty} \ln\left(1 - \frac{1}{p_k^s}\right) \leq -\sum_{k=1}^{\infty} \ln\left(1 - \frac{1}{p_k}\right)$$

On obtient une contradiction car

$$\lim_{1^+} \zeta = +\infty$$

3 La loi binomiale négative

1-

$$\mathbb{P}(X = k) = p_k = \binom{k+r-1}{r-1} p^r (1-p)^k$$

2- Utiliser le développement en série entière de

$$u \to (1-u)^{-r}$$

3- Notons $q = 1 - p$.
On trouve

$$G(t) = \left(\frac{p}{1-qt}\right)^r,\ E(X) = r.\frac{q}{p},\ V(X) = r.\frac{q^2}{p}$$

2 Une démonstration probabiliste du théorème de Stone-Weier

1. (a) *Puisque S_n suit la loi binomiale de paramètres n et x, alors $S_n(\Omega) = [\![0\ ,\ n]\!]$ et $\forall k \in [\![0\ ,\ n]\!]$, $P(S_n = k) = B_{n,k}(x)$.*
Donc

$$E(S_n) = \sum_{k=0}^{n} kB_{n,k}(x) = nx \text{ et } V(S_n) = E(S_n^2) - (E(S_n))^2 = \sum_{k=0}^{n} \frac{k^2}{n^2} B_{n,k}(x) - n^2x^2 = nx(1-x).$$

Donc $E(X_n) = x$ et $V(X_n) = \dfrac{1}{n^2}V(S_n) = \dfrac{x(1-x)}{n}$.

(b) *Soit $\delta > 0$. On a d'apèes l'inégalité de Bienaymé-Tchebychev,*

$$P(|X_n - x| \geq \delta) \leq \frac{V(X_n)}{\delta^2} = \frac{x(1-x)}{n\delta^2} \leq \frac{1}{4n\delta^2}.$$

2. (a) *D'après le théorème de transfert,*

$$C_n(f)(x) = E(f(X_n)) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{n,k}(x) = P_n(f)(x)$$

Donc $C_n(f)$ est polynomiale sur [0,1].

(b) i. *Puisque, pour tout $k \in [\![0\ ,\ n]\!]$, on a $P\left(X = \dfrac{k}{n}\right) = P(S_n = k) = B_{n,k}(x)$, alors d'après la question I.4.c.i on a le résultat.*

ii. *Pour tout $k \in [\![0\ ,\ n]\!]$ tel que $\left|\dfrac{k}{n} - x\right| > \beta$, on a $\left(X_n = \dfrac{k}{n}\right) = (|X_n - x| > \beta)$ et donc, d'après II.1.b, $P\left(X_n = \dfrac{k}{n}\right) \leq \dfrac{1}{4n\beta^2}$. D'où*

$$\sum_{\left|\frac{k}{n} - x\right| > \beta} \left(f(x) - f\left(\frac{k}{n}\right)\right) P\left(X_n = \frac{k}{n}\right) \leq \frac{M}{2n\beta^2}.$$

I Inégalité de Bienaymé-Tchebychev

Théorème 1
Soit X une VAR admettant un moment d'ordre 2. Alors :

$$\forall \varepsilon > 0, \quad P\left(|X - E(X)| \geqslant \varepsilon\right) \leqslant \frac{V(X)}{\varepsilon^2}$$

Démonstration :

Comme X admet un moment d'ordre 2, X admet bien une espérance et une variance.
• Si X est une variable discrète :
On pose $X(\Omega) = \{x_i, i \in I\}$ et $p_i = P(X = x_i)$.
Par définition on sait que $V(X) = E\left((X - E(X))^2\right)$. Donc d'après le théorème de transfert :

$$V(X) = E\left((X - E(X))^2\right) = \sum_{i \in I} (x_i - E(X))^2 p_i$$

Et de plus :

$$P(|X - E(X)| \geqslant \varepsilon) = \sum_{j \in J} p_j \text{ où } J = \{j \in I / |x_j - E(X)| \geqslant \varepsilon\}$$

On peut donc écrire :

$$\begin{aligned}
V(X) &= \sum_{j \in J} (x_j - E(X))^2 p_j + \sum_{i \notin J} (x_i - E(X))^2 p_i \\
&\geqslant \sum_{j \in J} (x_j - E(X))^2 p_j \\
&\geqslant \varepsilon^2 \sum_{j \in J} p_j \\
&\geqslant \varepsilon^2 P(|X - E(X)| \geqslant \varepsilon)
\end{aligned}$$

$$\Leftrightarrow P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{V(X)}{\varepsilon^2}$$

II Convergence d'une suite de VAR

1 Loi faible des grands nombres

Théorème 2
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de VAR indépendantes, admettant une même espérance m et une même variance σ^2. On pose $\overline{X}_n = \dfrac{X_1 + \cdots + X_n}{n}$. Alors on a :

$$\forall\varepsilon > 0, \quad \lim_{n\to+\infty} P\left(|\overline{X}_n - m| \geqslant \varepsilon\right) = 0$$

On dit que $\overline{X}_n$ converge en probabilité vers la variable constante égale à m.

Démonstration :

On a $E(\overline{X}_n) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} E(X_i) = \dfrac{1}{n}nm = m$ et comme les variables sont indépendantes, $V(\overline{X}_n) = \dfrac{1}{n^2}\displaystyle\sum_{i=1}^{n} V(X_i) = \dfrac{\sigma^2}{n}$.
D'après l'inégalité de Bienaymé-Tchebychev,

$$0 \leqslant P\left(|\overline{X}_n - m| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

Donc par encadrement de limites, $\lim\limits_{n\to+\infty} P\left(|\overline{X}_n - m| \geqslant \varepsilon\right) = 0$

$\square$

Propriété 1
Soit (X_n) une suit de VAR convergeant en loi vers la VAR X. Pour tous points a et b de continuité de F_X tels que $a < b$, on a :
$$\lim_{n\to+\infty} P(a < X_n \leqslant b) = P(a < X \leqslant b)$$

Théorème 3
Soient $(X_n)_{n\in\mathbb{N}}$ une suite de VAR **discrètes** et X une VAR **discrète** telles que pour tout entier n, $X_n(\Omega) \subset \mathbb{N}$ et $X(\Omega) \subset \mathbb{N}$.
Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X ssi :
$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(X_n = k) = P(X = k)$$

3 Théorème de la limite centrée

Théorème 4
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires réelles définies sur une même espace probabilisé et **indépendantes**. On suppose que les X_n ont toutes la même loi et qu'elles admettent une variance non nulle.
On note :
$$E(X_1) = m \quad V(X_1) = \sigma^2$$
$$\forall n \in \mathbb{N}^*, \quad S_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad Y_n = \frac{S_n}{n}$$
On a donc $S_n^* = \dfrac{S_n - nm}{\sigma\sqrt{n}} = \dfrac{\sqrt{n}(Y_n - m)}{\sigma} = Y_n^*$
Alors la suite (S_n^*) (et donc la suite (Y_n^*)) converge en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$.

III Approximation de variables aléatoires

1 Approximation d'une loi hypergéométrique par une loi binomiale

Théorème 5
Soient n un entier fixé et $p \in]0;1[$. On pose $I = \{N \in \mathbb{N}/Np \in \mathbb{N}\}$. Soit $(X_N)_{N\in I}$ une suite de VAR telle que pour tout N, X_N suit la loi hypergéométrique $\mathcal{H}(N, n, p)$. Alors (X_N) converge en loi vers une VAR X qui suit la loi binomiale $\mathcal{B}(n, p)$.

Démonstration : Hors programme

On a :
$$\begin{aligned}
P(X_N = k) &= \frac{\dbinom{Np}{k}\dbinom{Nq}{n-k}}{\dbinom{N}{n}} \\
&= \frac{(Np)!(Nq)!n!(N-n)!}{k!(Np-k)!(n-k)!(Nq-n+k)!N!} \\
&= \frac{n!}{k!(n-k)!}\frac{(Np)(Np-1)\cdots(Np-k+1)\times(Nq)\cdots(Nq-n+k+1)}{N(N-1)\cdots(N-n+1)}
\end{aligned}$$
Le produit $(Np)(Np-1)\cdots(Np-k+1)\times(Nq)\cdots(Nq-n+k+1)$ possède n facteurs chacun équivalent à Np ou Nq lorsque $N \to +\infty$ donc est équivalent à $(Np)^k(Nq)^{n-k}$ lorsque $N \to +\infty$. De même $N(N-1)\cdots(N-n+1) \sim N^n$ donc on en déduit que
$$\lim_{N\to+\infty} P(X_N = k) = \frac{n!}{k!(n-k)!}p^k q^{n-k}$$
X_N converge bien en loi vers une VAR X qui suit la loi $\mathcal{B}(n, p)$.

□

2 *Approximation d'une loi binomiale par une loi de Poisson*

Théorème 6
Soit λ un réel strictement positif et $(X_n)_{n\in\mathbb{N}}$ une suite de VAR discrètes telles que X_n suit la loi binomiale de paramètre $\left(n, \dfrac{\lambda}{n}\right)$. Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une VAR X qui suit une loi de Poisson de paramètre λ.

Démonstration :

On a
$$\begin{aligned} P(X_n = k) &= \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}e^{(n-k)\ln(1-\lambda/n)} \end{aligned}$$
On a lorsque $n \to +\infty$, $n(n-1)\cdots(n-k+1) \sim n^k$
De plus comme $\dfrac{\lambda}{n} \to 0$, $\ln(1-\lambda/n) \sim -\dfrac{\lambda}{n}$. Donc on en déduit :
$$\lim_{n\to+\infty} P(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda}$$
Ainsi X_n converge bien en loi vers une variable qui suit la loi $\mathcal{P}(\lambda)$.

□

3 *Approximation d'une loi binomiale par une loi normale*

Théorème 7
Soit $p \in]0;1[$, $q = 1-p$ et $(S_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires telle que S_n suit la loi binomiale $\mathcal{B}(n,p)$. Alors la suite de variable aléatoire $S_n^* = \dfrac{S_n - np}{\sqrt{npq}}$ converge en loi vers la loi normale centrée réduite.

Démonstration : Hors programme

Toute variable binomiale de paramètre (n,p) peut être considérée comme la somme de n variables aléatoires de Bernoulli de paramètre p mutuellement indépendantes. On a donc $S_n = \displaystyle\sum_{k=1}^{n} X_k$ où X_k est une variable de Bernoulli de paramètre p. Comme les variables de Bernoulli admettent une espérance p et une variance pq, le théorème de la limite centrée s'applique bien et il nous donne le résultat demandé.

□

4 *Approximation d'une loi de Poisson par une loi normale*

Théorème 8
Soit α un réel strictement positif et $(S_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires telle que S_n suit la loi de Poisson $\mathcal{P}(n\alpha)$. Alors la suite de variable aléatoire $S_n^* = \dfrac{S_n - n\alpha}{\sqrt{n\alpha}}$ converge en loi vers la loi normale centrée réduite.

Démonstration : Hors programme

Toute variable de Poisson de paramètre $n\alpha$ peut être considérée comme la somme de n variables aléatoires de Poisson de paramètre α mutuellement indépendantes. On a donc $S_n = \displaystyle\sum_{k=1}^{n} X_k$ où X_k est une variable de Poisson de paramètre α. Comme les variables de Poisson admettent une espérance α et une variance α, le théorème de la limite centrée s'applique bien et il nous donne le résultat demandé.

□

En pratique :
On considère que pour $\lambda \geqslant 18$ on peut approcher la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{N}(\lambda, \lambda)$.

Correction de continuité :
Tout comme précédemment il ne faut pas oublier d'appliquer, si nécessaire, la correction de continuité.

Exemple 8:
Ici encore pour de grandes valeurs de λ, la calcul de $P(X \leqslant a)$ pourra nécessiter l'utilisation d'un ordinateur.

Si on considère X qui suit une loi de Poisson de paramètre 64 que l'on cherche à calculer $P(X \leqslant 74)$, on a intérêt à approcher la loi de X par la loi $\mathcal{N}(64, 64)$ et donc $\dfrac{X-64}{8}$ suit la loi normale centrée réduite. On a donc

$$P(X \leqslant 74) = P\left(\frac{X-64}{8} \leqslant 1,25\right) \approx \Phi(1,25) \approx 0,8944$$