CHAPITRE 2

THEOREMES LIMITES, METHODES D'APPROXIMATION

I Approximation par la loi normale

Rappel :
On considère une variable somme de variables indépendantes, dont on suppose qu'elles suivent toutes la même loi de probabilité. On ne sait pas forcément déterminer la loi que suit la somme, alors qu'on peut facilement calculer son espérance et sa variance.

Concrètement, si $S = X_1 + X_2 \cdots + X_n$, avec les X_i, X_j deux à deux indépendantes et suivant toutes la même loi d'espérance $E(X_i) = \mu$ et de variance $V(X_i) = \sigma^2$, on a :
$E(S) = \sum_{i=1}^{n} E(X_i) = n * \mu$ et $V(S) = \sum_{i=1}^{n} V(X_i) = n * \sigma^2$.

En ce qui concerne la loi de S, le théorème suivant permet d'en obtenir une approximation.

Théorème central limite :
Soit $X_1, X_2 \cdots, X_n$ une famille de variables aléatoires indépendantes et obéissant à la même loi de probabilité. On suppose que $E(X_i) = \mu$ et $V(X_i) = \sigma^2$, avec $\sigma > 0$. Si $S_n = X_1 + X_2 \cdots + X_n$, alors la loi de probabilité de la somme réduite $S_n^* = \frac{1}{\sigma\sqrt{n}}(S_n - n * \mu)$ converge vers la loi normale réduite $\mathcal{N}(0,1)$, c'est-à-dire que, pour tous réels a et b tels que $a < b$, on a :

$$\lim_{n \to +\infty} P(a < S_n^* < b) = \Pi(b) - \Pi(a).$$

On peut donc, pour n assez grand, faire l'approximation de S_n^* par $\mathcal{N}(0,1)$ et de la loi de S_n par la loi $\mathcal{N}(n\mu, \sigma\sqrt{n})$.

Exemple d'application :

Dans une file d'attente, 10 personnes attendent d'être servies. On appelle T_k la durée de service pour la personne n°k et on suppose que pour tout k, T_k suit une loi exponentielle de moyenne 1 minute.
Le temps d'attente de la 11ème personne qui se présente sera donc $T = T_1 + T_2 \cdots + T_{10}$. On cherche la probabilité que cette personne attende plus de 15 minutes.

On sait que $E(T_k) = \frac{1}{\lambda} = 1$ et $V(T_k) = \frac{1}{\lambda^2} = 1$. Donc T suit approximativement la loi $\mathcal{N}(10, \sqrt{10})$.
Donc $P(T > 15) = P\left(T^* > \frac{15-10}{\sqrt{10}}\right) = P(T^* > 1.5811)$

$$= 1 - \Pi(1.5811) \approx 0.057$$

II Approximations de la loi binomiale

Dans la pratique, on remplace souvent une loi de probabilité donnée par une loi plus facile à calculer. Ici on étudie la loi binomiale.

1) Approximation par la loi normale

Exemple : on sait qu'en moyenne, 5% des articles produits par un certain procédé de fabrication sont défectueux. Quelle est la probabilité qu'il y ait au plus 60 pièces défectueuses dans un lot de 1000 pièces choisies au hasard ?

Si l'on admet que la taille du lot est petite par rapport à la production totale, on peut utiliser le modèle du tirage avec remise des pièces prélevées. Si X est la variable aléatoire "nombre de pièces défectueuses dans le lot", X suit la loi $\mathcal{B}(1000, 0.05)$.
Donc $P(X \leq 60) = \sum_{k=0}^{60} \binom{1000}{k} * 0.05^k * 0.95^{1000-k}$, calcul très lourd.
Or, X est également la somme de 1000 variables indépendantes qui suivent la loi de Bernoulli $\mathcal{B}(1, 0.05)$. Donc, d'après l'approximation du I, on peut considérer que X suit une loi normale de paramètres $E(X)$ et $\sigma(X)$.
On sait que, pour une loi binomiale, $E(X) = n * p$ et $\sigma(X) = \sqrt{n * p * q}$ avec $q = 1 - p$.
Donc ici, $E(X) = 1000 * 0.05 = 50$ et $\sigma(X) = \sqrt{1000 * 0.05 * 0.95} = \sqrt{47.5} \approx 6.9$.

On pourrait donc envisager d'écrire $P(X \leq 60) \approx \Pi\left(\frac{60-50}{6.9}\right) \approx 0.9265$. Mais ce calcul ne tient pas compte du fait que la loi binomiale est une loi discrète et la loi normale une loi continue.
On doit tenir compte de cette différence en introduisant une "correction de continuité" : pour la loi binomiale, $P(X \leq 60) = P(X < 61)$, ce qui n'est pas vrai pour la loi normale. On prendra pour correction : $P(X \leq 60) \approx \Pi\left(\frac{\mathbf{60{,}5}-50}{6.9}\right) \approx 0.9357$.

Plus généralement, si X suit une loi $\mathcal{B}(n, p)$, on prendra pour approximation :

$$P(a \leq X \leq b) = P\left(\frac{a - 0.5 - n * p}{\sqrt{n * p * (1-p)}} \leq X \leq \frac{b + 0.5 - n * p}{\sqrt{n * p * (1-p)}}\right)$$

c'est-à-dire :

$$P(a \leq X \leq b) = \Pi\left(\frac{b + 0.5 - n * p}{\sqrt{n * p * (1-p)}}\right) - \Pi\left(\frac{a - 0.5 - n * p}{\sqrt{n * p * (1-p)}}\right)$$

Attention !! cette approximation par la loi normale donne des résultats corrects si $n * p \geq 5$ et $n * (1-p) \geq 5$. Alors $\underline{\mathcal{B}(n, p) \approx \mathcal{N}\left(n * p, \sqrt{n * p * (1-p)}\right)}$, à la correction de continuité près.

2) Approximation par la loi de Poisson

L'approximation de la loi $\mathcal{B}(n, p)$ ne donne pas des résultats satisfaisants si $n * p < 5$. Dans cette situation, on utilise l'approximation de la loi $\mathcal{B}(n, p)$ par la loi de Poisson.

Théorème :

Soit X une variable aléatoire obéissant à une loi binomiale $\mathcal{B}(n,p)$ telle que $n * p$ soit une constante $\lambda > 0$, donc telle que p varie avec n : $p = \frac{\lambda}{n}$. Alors $P(X = k)$ tend vers $\frac{\lambda^k e^{-\lambda}}{k!}$ lorsque $n \to +\infty$, k prenant ici toutes les valeurs possibles.

Dans la pratique, l'approximation "poissonnienne" $\mathcal{B}(n,p) \approx \mathcal{P}(\lambda)$ est acceptable si $n \geq 30$ et $p \leq 0.1$ avec $\lambda = n * p \leq 10$.

Exemple : l'exemple des pièces défectueuses ci-dessus avec $p = 0.01$ et $n = 100$.

X suit alors la loi $\mathcal{B}(100, 0.01)$: $n * p = 100 * 0.01 = 1$ donc $n * p < 5$. L'approximation normale n'est donc pas satisfaisante. Par contre, en posant $\lambda = n * p = 1$, on a bien $n \geq 30$, $p \leq 0.1$ et $\lambda \leq 10$.

On cherche à calculer la probabilité qu'il y ait au plus 3 pièces défectueuses dans un lot de 100 pièces :

$P(X \leq 3) = \sum_{k=0}^{3} \binom{100}{k} * 0.01^k * 0.99^{100-k} \approx \sum_{k=0}^{3} e^{-1} * 1^k * \frac{1}{k!} = 0.981$

Cette approximation est excellente puisque la valeur exacte de cette probabilité est 0.982.

Attention !! Les deux lois sont discrètes, il n'y a pas de correction de continuité pour cette approximation.

III Approximation normale de la loi de Poisson

Théorème :

Soit X une variable aléatoire poissonnienne de paramètre λ. Alors la loi de la variable réduite $\frac{X-\lambda}{\sqrt{\lambda}}$ converge vers la loi normale réduite lorsque $\lambda \to +\infty$. , donc $\mathcal{P}(\lambda) \approx \mathcal{N}(\lambda, \sqrt{\lambda})$.

Dans la pratique cette approximation est valable lorsque $\lambda \geq 10$. De plus la loi de Poisson est discrète, donc il faut tenir compte de la correction de continuité.

Alors : $P\left(a \leq \frac{X-\lambda}{\sqrt{\lambda}} \leq b\right) = \Pi(b + 0.5) - \Pi(a - 0.5)$.

Exemple :

Si $X \leadsto \mathcal{P}(\lambda = 10)$, alors $P(X \leq 6) \approx \Pi\left(\frac{6.5-10}{\sqrt{10}}\right) \approx 0.1345$