Série 01 - Processus Empirique et Applications

Université Mohamed Khider Biskra, Département de Mathématiques

2025/2026

Exercice 01

1. Calcul de $E(Y^4)$ pour $Y \sim N(0,1)$

Soit $Y \sim N(0,1)$, une variable aléatoire suivant une loi normale centrée réduite. Nous souhaitons calculer $E(Y^4)$, c'est-à-dire l'espérance du quatrième moment de Y.

Pour une variable aléatoire $Y \sim N(0,1)$, on sait que les moments pairs sont donnés par :

$$E(Y^{2k}) = (2k-1)!! = 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k-1)$$

Ainsi, pour $k = 2$, nous avons :

$$E(Y^4) = 3$$

2. Limite de la moyenne des X_k^4

Soit $X_k \sim N(m, \sigma^2)$ une suite de variables aléatoires indépendantes de même loi normale de moyenne m et de variance σ^2. Nous cherchons à démontrer que :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k^4 = m^4 + 6m^2\sigma^2 + 3\sigma^4 \quad \text{presque sûrement.}$$

Par la **loi des grands nombres**, la moyenne des X_k^4 converge presque sûrement vers l'espérance de X_k^4, soit :

$$E(X_k^4) = m^4 + 6m^2\sigma^2 + 3\sigma^4$$

Cela signifie que la moyenne des X_k^4 converge presque sûrement vers cette valeur à mesure que n devient grand.

Exercice 02: Propriétés élémentaires de la fonction de répartition empirique

Soit $F_n(x)$ la fonction de répartition empirique associée à un échantillon de taille n.

1. Biais de $F_n(x)$

La fonction de répartition empirique $F_n(x)$ est un estimateur sans biais de $F(x)$, la fonction de répartition théorique. Cela signifie que :

$$E[F_n(x)] = F(x)$$

En effet, $F_n(x)$ est la proportion d'observations inférieures ou égales à x, et en moyenne, cette proportion correspond à $F(x)$.

2. Consistance de $F_n(x)$

La fonction de répartition empirique $F_n(x)$ converge en probabilité vers la fonction de répartition théorique $F(x)$ lorsque n devient grand :

$$F_n(x) \xrightarrow{P} F(x) \quad \text{lorsque} \quad n \to \infty$$

Cela découle du **théorème de Glivenko-Cantelli**, qui garantit que $F_n(x)$ converge uniformément vers $F(x)$ à mesure que n augmente.

3. Consistance forte de $F_n(x)$

La fonction de répartition empirique $F_n(x)$ converge presque sûrement vers $F(x)$ lorsque $n \to \infty$:

$$F_n(x) \xrightarrow{p.s.} F(x) \quad \text{lorsque} \quad n \to \infty$$

Cela signifie que $F_n(x)$ converge vers $F(x)$ pour presque tous les x, avec probabilité 1.

4. Théorème de Glivenko-Cantelli

Le **théorème de Glivenko-Cantelli** stipule que :

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{p.s.} 0 \quad \text{lorsque} \quad n \to \infty$$

Cela signifie que la différence maximale entre $F_n(x)$ et $F(x)$ converge presque sûrement vers 0. En d'autres termes, $F_n(x)$ devient de plus en plus proche de $F(x)$ de manière uniforme sur tout $\mathbb{R}$.

5. Normalité asymptotique

Lorsque $n \to \infty$, la différence entre $F_n(x)$ et $F(x)$, multipliée par $\sqrt{n}$, suit une distribution normale :

$$\sqrt{n}(F_n(x) - F(x)) \xrightarrow{d} N(0, F(x)(1 - F(x)))$$

Cela montre que, pour de grandes valeurs de n, la différence $F_n(x) - F(x)$ se distribue normalement, ce qui permet des approximations normales pour des tests statistiques.

Exercice 03: Statistique d'ordre

Soit $X_1, X_2, \ldots, X_n$ une suite de variables aléatoires i.i.d. suivant une fonction de répartition F. Nous définissons la statistique d'ordre $(X(1), X(2), \ldots, X(n))$, où les $X(i)$ sont les valeurs de $X_1, X_2, \ldots, X_n$ triées dans l'ordre croissant.

1. Fonction de répartition de $X(1)$ et $X(n)$

Les fonctions de répartition des variables $X(1)$ (le minimum) et $X(n)$ (le maximum) sont données par :

$$F_{X(1)}(x) = P(X(1) \leq x) = 1 - (1 - F(x))^n$$

$$F_{X(n)}(x) = P(X(n) \leq x) = F(x)^n$$

Cela montre que $X(1)$ a une fonction de répartition croissante, et $X(n)$ est également croissante mais converge plus lentement vers 1.

2. Loi du couple $(X(1), X(n))$ et indépendance

Le couple $(X(1), X(n))$ n'est **pas indépendant**. La statistique du minimum et du maximum dans un échantillon sont liées, car elles dépendent des mêmes variables. La loi jointe de $X(1)$ et $X(n)$ peut être obtenue par le produit de leurs distributions, mais elles ne sont pas indépendantes.

3. Loi de la statistique $W = X(n) - X(1)$

La statistique W, appelée **étendue**, est la différence entre le maximum et le minimum dans un échantillon. Sa loi peut être obtenue à partir des lois de $X(1)$ et $X(n)$, et elle dépend de la fonction de répartition F.

4. Calcul de $P(X(k) < x)$

La fonction de répartition de la k-ème statistique d'ordre $X(k)$ est donnée par :

$$P(X(k) < x) = \sum_{i=k}^{n} \binom{n}{i} F(x)^i (1 - F(x))^{n-i}$$

Cela donne la probabilité que la k-ème plus petite valeur dans l'échantillon soit inférieure à x.

5. Densité de $X(k)$

La fonction de densité de $X(k)$ est :

$$f_k(x) = n \binom{n-1}{k-1} F(x)^{k-1} (1 - F(x))^{n-k} f(x)$$

Cela permet de calculer la densité de la k-ème statistique d'ordre dans un échantillon i.i.d.

Exercice 04: Convergence presque sûre de F_n

Soit $X_1, X_2, \ldots, X_n$ une suite de variables aléatoires i.i.d. et F la fonction de répartition de X_1. La fonction de répartition empirique est définie par :

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq t)$$

où $I(X_i \leq t)$ est la fonction indicatrice qui est 1 si $X_i \leq t$ et 0 sinon.

1. Loi de $nF_n(t)$ et loi limite de $\sqrt{n}(F_n(t) - F(t))$

La loi de $nF_n(t)$ suit une loi binomiale, et la loi limite de $\sqrt{n}(F_n(t) - F(t))$ est une loi normale centrée :

$$\sqrt{n}(F_n(t) - F(t)) \xrightarrow{d} N(0, F(t)(1 - F(t)))$$

2. Calcul de $E[(F_n(t) - F(t))^2]$

On montre que :

$$E[(F_n(t) - F(t))^2] \to 0 \quad \text{lorsque} \quad n \to \infty$$

Cela montre que $F_n(t)$ converge en moyenne quadratique vers $F(t)$.

3. Convergence presque sûre de $F_n(t)$

En utilisant le théorème de Borel-Cantelli, on montre que $F_n(t)$ converge presque sûrement vers $F(t)$.

4. Calcul de D_n

La statistique de Kolmogorov-Smirnov D_n est donnée par :

$$D_n = \max_{0 \leq i \leq n} \max \left(\frac{i}{n} - F(X_{n;i}), \frac{i+1}{n} - F(X_{n;i+1}) \right)$$