TD 5 PROBABILITÉS - VECTEURS GAUSSIENS ET CONVERGENCE - 1SN

Vecteurs gaussiens

Exercice 1 : Indépendance de la moyenne et de la variance empiriques pour un vecteur gaussien.
Soit V un vecteur aléatoire gaussien à valeurs dans $\mathbb{R}^3$. Soient X_1, X_2 et X_3 les composantes de V suivant la base canonique. On suppose que X_1, X_2 et X_3 sont des variables aléatoires indépendantes de loi normale $\mathcal{N}(0,1)$.
1) Quelle est la densité du triplet (X_1, X_2, X_3) ?
2) Soit P la matrice de changement de base orthonormée telle que :

$$P^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{pmatrix}$$

P est orthonormée donc inversible.

$P \times P^T = I_n$

$P^T = P^{-1}$

On note X le vecteur colonne de composantes $(X_i)_{i=1,2,3}$ et Y le vecteur colonne $Y = P^T X$ de composantes $(Y_i)_{i=1,2,3}$.
2.1) Quelle est la loi du triplet (Y_1, Y_2, Y_3) ?
2.2) Déterminer les lois de Y_1, Y_2 et Y_3.
3) Soit $\overline{X} = \frac{1}{3}\sum_{i=1}^{3} X_i$. Vérifier que $\sum_{i=1}^{3} X_i^2 = \sum_{i=1}^{3} Y_i^2$ et $Y_2^2 + Y_3^2 = \sum_{i=1}^{3} X_i^2 - 3\overline{X}^2$. Exprimer $\overline{X}$ et $S^2 = \frac{1}{2}\sum_{i=1}^{3}\left(X_i - \overline{X}\right)^2$ en fonction des variables aléatoires Y_i. En déduire que $\overline{X}$ et S^2 sont des variables aléatoires indépendantes.
4) Donner la loi de $\overline{X}$ et de $2S^2$.

Exercice 2 : changement de variables et indépendance pour vecteurs gaussiens
Soient X, Y et Z trois variables aléatoires réelles indépendantes de loi $\mathcal{N}(0,1)$.
1) Déterminer la loi de $U = X + Y + Z$.
2) Montrer que $X - Y$ est une variable aléatoire indépendante de U.

Exercice 3 : comment générer un vecteur gaussien de vecteur moyenne $\boldsymbol{m}$ et de matrice de covariance $\boldsymbol{\Sigma}$ donnés ?
Soient X et Y deux variables aléatoires indépendantes de loi normale de moyenne $\mu = 0$ et de variance $\sigma^2 = 1$. Soit $\boldsymbol{m}$ un vecteur de $\mathbb{R}^2$ et $\boldsymbol{\Sigma}$ une matrice symétrique définie positive de $\mathcal{M}_2(\mathbb{R})$. Déterminer une matrice M et un vecteur $\boldsymbol{n}$ tels que $\boldsymbol{v} = M\begin{pmatrix} X \\ Y \end{pmatrix} + \boldsymbol{n}$ soit un vecteur Gaussien de $\mathbb{R}^2$ de moyenne $\boldsymbol{m}$ et de matrice de covariance $\boldsymbol{\Sigma}$. Ce résultat permet de générer des vecteurs gaussiens à partir de variables aléatoires normales centrées réduites.

Convergence

Exercice 4 : convergence vers 0.
Soit la suite de va X_n définie pour $n \in \mathbb{N}$ par :

$$P[X_n = 0] = 1 - \frac{1}{n}$$
$$P[X_n = n] = \frac{1}{n}$$

Montrer que la suite X_n converge en loi et en probabilité vers $X = 0$ mais que X_n ne converge pas en moyenne quadratique vers $X = 0$.

Exercice 5 : convergence vers une loi uniforme

1) Soit Y une variable aléatoire de loi uniforme sur $[-1, +1]$. Quelle est la fonction caractéristique de Y ?

2) On considère une suite $(X_j)_{j\in\mathbb{N}^*}$ de variables aléatoires indépendantes de loi :

$$P\left[X_j = \frac{1}{2^j}\right] = \frac{1}{2} \text{ et } P\left[X_j = -\frac{1}{2^j}\right] = \frac{1}{2}$$

On pose $S_n = \sum_{j=1}^{n} X_j$. Déterminer la fonction caractéristique de X_j, puis celle de S_n notée $\varphi_n(t)$.

3) En utilisant la formule $\sin t = 2\sin(t/2)\cos(t/2)$, vérifier que :

$$\varphi_n(t) = \frac{\sin t}{t} \frac{\frac{t}{2^n}}{\sin\frac{t}{2^n}} \qquad t \neq 0, \forall n \in \mathbb{N}^*$$

4) En déduire que S_n converge en loi vers une variable aléatoire que l'on précisera.

Exercice 6 : théorème de la limite centrale

On considère une suite de variables aléatoires indépendantes X_j de même loi de Poisson de paramètre $\theta = 1$.

1) Quelle est la loi de $S_n = \sum_{j=1}^{n} X_j$?

2) Soit $T_n = \frac{S_n - n}{\sqrt{n}}$. En utilisant le théorème de la limite centrale et en considérant les événements $\{T_n \leq 0\}$, montrer que :

$$\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{n-1} \frac{n^k}{k!} = \frac{1}{2}$$

Application aux sciences du numérique

Exercice 7 : estimation d'une probabilité d'erreur

Afin de tester les performances d'un système de communications numériques, il est usuel de simuler le fonctionnement de ce système sur un ordinateur. Un des problèmes consiste alors à estimer la probabilité d'erreur p associé à ce système. On estime généralement cette probabilité comme suit :

$$\widehat{p} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

où X_i est une variable aléatoire binaire telle que

$$\begin{aligned} X_i &= 1 \text{ s'il y a une erreur pour le } i^{\grave{e}me} \text{ symbole (évènement de probabilité } p) \\ X_i &= 0 \text{ s'il n'y a pas d'erreur pour le } i^{\grave{e}me} \text{ symbole (évènement de probabilité } 1-p) \end{aligned}$$

Déterminer la moyenne et la variance de $\widehat{p}$ puis sa loi limite en utilisant le théorème de la limite centrale. On cherche le nombre de points N nécessaire pour que $\widehat{p}$ soit une approximation de p avec une précision relative inférieure à 10%. Pour cela, on se fixe un degré de confiance $\alpha = 95\%$, qui indique la probabilité d'avoir cette précision soit

$$P\left[\left|\frac{\widehat{p} - p}{p}\right| < \varepsilon\right] = \alpha$$

Déterminer N pour que l'égalité précédente soit vérifiée.

Remarque : pour $\varepsilon = 20\%$, on trouve $N \simeq 100/p$ d'où $Np \simeq N\widehat{p} \simeq 100$, d'où la règle pratique suivante : il suffit d'observer une centaine d'erreurs pour pouvoir estimer la probabilité d'erreur p à l'aide de l'estimateur $\widehat{p}$ avec une précision relative $\varepsilon = 20\%$ et un degré de confiance $\alpha = 95\%$.

TD 5.1
Probas

Vecteurs Gaussiens et convergence

⚠ Exo 6 important.

Exercice 6: $X_j \hookrightarrow \mathcal{P}(\theta)\ \forall j \qquad P(X_j = k) = e^{-\theta}\dfrac{\theta^k}{k!}$

1/ $S_n = \sum_{j=1}^{n} X_j$ avec $\forall j \in [\![1,n]\!]\ X_j \hookrightarrow \mathcal{P}(1)$

On a $S_n \hookrightarrow \mathcal{P}(n\theta)$ ie $S_n \hookrightarrow \mathcal{P}(n)$.

Preuve on a $G_X(t) = \mathbb{E}(t^X)$ Donc $G_{X+Y}(t) = \mathbb{E}(t^{X+Y}) = \mathbb{E}(t^X) \times \mathbb{E}(t^Y)$ car X, Y indep

Donc $G_{S_n} = \prod G_{X_j}(t) = \prod_{i=1}^{n} e^{\theta(t-1)} = e^{n\theta(t-1)}$

↳ F.C d'une loi de Poisson de param $n\theta$.

2/ $T_n = \dfrac{S_n - n}{\sqrt{n}}$ Comme $X_n \hookrightarrow \mathcal{P}(\theta)$ Donc $\mathbb{E}(X_n)$ et $V(X_n) = 1$

Donc $\dfrac{S_n - n}{\sqrt{n}} = \dfrac{S_n - nm}{\sqrt{n\sigma^2}} \hookrightarrow \mathcal{N}(0,1)$

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx = \sqrt{\frac{\pi}{\alpha}} = 2\int_{0}^{+\infty} e^{-\alpha x^2}\,dx$$

$\Phi(x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-m)^2}{\sigma^2}}\,du.$

$1 - \Phi(x) = \int_{x}^{+\infty} \_\_\ du.$

$\Phi(x) = P(X < x)$

$1 - \Phi(x) = P(X \geq x)$

$P(T_n \leq 0) = \int_{-\infty}^{0}$

2/ On a $S_n \hookrightarrow \mathcal{P}(n)$

$T_n = \dfrac{S_n - n}{\sqrt{n}} = \dfrac{\sum_{j=1}^{n} X_j - n}{\sqrt{n}}$ → Thm de la limite centrale

$T_n \xrightarrow{\mathcal{L}} T$ avec $T \hookrightarrow \mathcal{N}(0,1)$.

ie $\forall x \in \mathbb{R}\quad F_{T_n}(x) \xrightarrow[n\to+\infty]{} F_T(x)$

$F_{T_n}(x) = P[T_n < x] = P\left(\dfrac{S_n - n}{\sqrt{n}} < x\right) = P\left(S_n < x\sqrt{n} + n\right)$

donc $F_{T_n}(0) = P(S_n < n)$

$= \sum_{j=0}^{n-1} P(S_n = j)$

$= e^{-n} \sum_{j=0}^{n-1} \dfrac{n^j}{j!}$

et donc $e^{-n} \sum_{j=0}^{n-1} \dfrac{n^k}{k!} \xrightarrow[n\to+\infty]{} F_T(0) = \int_{-\infty}^{0} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$

$= \dfrac{1}{\sqrt{2\pi}} \times \dfrac{1}{2} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx = \dfrac{1}{2}$

Exercice 1 :

1/ $p(x) = \dfrac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} e^{\left(-\frac{1}{2}(x-m)^T \Sigma^{-1}(x-m)\right)}$

Loi Gaussienne de moyenne $m = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ et de matrice de covariance $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$p(x) = \dfrac{1}{(2\pi)^{3/2} \times 1} e^{-\frac{1}{2}x^T x}$

Comme les v.a. sont indep. les Cov sont donc nulles

2/ a $Y = P^T X$ Comme P inversible

$Y \sim \mathcal{N}_3(P^T m + b$

$\mathcal{N}_3(0, P^T I_3 P)$ $\quad 0_{\mathbb{R}^3}$

$Y \hookrightarrow \mathcal{N}_3(0_{\mathbb{R}^3}, I_3)$

b Donc $(Y_1, Y_2, Y_3) \hookrightarrow \mathcal{N}_3(0_{\mathbb{R}^3}, I_3)$

3/ $\overline{X} = \frac{1}{3}\sum_{i=1}^{n} X_i = \cancel{X}$

On a $\sum_{i=1}^{3} X_i^2 = X^T X = Y^T Y = \sum_{i=1}^{3} Y_i^2$ car $\|X\| = \|Y\|$ $\quad Y^T Y = X^T P P^T X = X^T X$

$Y_2^2 + Y_3^2 = Y^T Y - Y_1^2$ $\quad X_1^2 - Y_1^2 =$

$= \sum_{i=1}^{3} X_i^2 - Y_1^2$ $\quad \frac{1}{3}(X_1 + X_2 + X_3)^2$

$= \sum_{i=1}^{3} X_i^2 - \frac{1}{3}(X_1 + X_2 + X_3)^2$ $\quad \frac{1}{3}(X_1^2 + 2X_1X_2 + 2X_2X_3 + 2X_1X_3 + X_2^2 + X_3^2)$

$= \sum_{i=1}^{3} X_i^2 - 3\overline{X}^2$

$Y_1 = P^T X$ Donc $Y_1 = \frac{1}{\sqrt{3}}(X_1 + X_2 + X_3)$

$$\overline{X} = \frac{1}{3}\sum_{i=1}^{3} X_i = \frac{1}{\sqrt{3}} Y_1$$

$S^2 = \frac{1}{2}\sum_{i=1}^{3}(X_i - \overline{X})^2 = \frac{1}{2}\sum_{i=1}^{3} X_i^2 - \sum_{i=1}^{3} X_i \overline{X} + \frac{3}{2}\overline{X}^2$

$= \frac{1}{2}\sum X_i^2 - \overline{X}\sum_{i=1}^{3} X_i + \frac{1}{2} Y_1^2$

$= \frac{1}{2}\sum Y_i^2 - Y_1^2 + \frac{1}{2} Y_1^2$

$= \frac{1}{2}(Y_2^2 + Y_3^2)$

$2S^2 = Y_2^2 + Y_3^2$

$2S^2 \hookrightarrow \chi^2_2$ car $Y_2, Y_3 \hookrightarrow \mathcal{N}(0,1)$

loi du chi 2 à 2 deg de lib

4/ $\overline{X} = \frac{X_1}{\sqrt{3}}$ $\quad E(\overline{X}) = \frac{1}{\sqrt{3}} E(Y_1)$ $\quad \overline{X} \hookrightarrow \mathcal{N}(0, \frac{1}{3})$

$V(\overline{X}) = \frac{1}{3} V(Y_1)$

TD 5.2

Probas

Exercice 2 :

1/ $E(U) = E(X) + E(Y) + E(Z) = 0$

$Var(U) = Var(X) + Var(Y) + Var(Z) = 3$.

$U \hookrightarrow \mathcal{N}(0,3)$.

2/ $\underbrace{\begin{pmatrix} X-Y \\ X+Y+Z \end{pmatrix}}_{S} = \underbrace{\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}}_{\Theta}$. P de rang max 3

$m_S = P m_\Theta = P\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$\Sigma_S = P \Sigma_\Theta P^T = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ Donc Cov(