

Remise à niveau - Statistique

Vecteurs aléatoires, vecteurs gaussiens et théorème de Cochran

Claire Boyer

1. Vecteurs aléatoires
2. Vecteurs aléatoires gaussiens
3. Convergences
4. Théorèmes asymptotiques
5. Théorème de Cochran et applications

Définition. Un vecteur aléatoire $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ est un vecteur dont toutes les composantes sont des variables aléatoires réelles.

Fonction de répartition F_X de X : pour tout $t = (t_1, \dots, t_d)$,

$$F_X(t_1, \dots, t_d) = \mathbb{P}[X_1 \leq t_1, \dots, X_d \leq t_d].$$

Densité Si X est **continu**, i.e. X admet une **densité** f_X par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$, on a

$$F_X(t) = \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_d} f_X(u_1, \dots, u_d) du_1 \cdots du_d,$$

avec pour presque tout t

$$f_X(t) = \frac{\partial^d}{\partial t_1 \cdots \partial t_d} F_X(t).$$

- ▶ Toute densité f_X vérifie $f_X(t) \geq 0$ et

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_X(t_1, \dots, t_d) dt_1 \cdots dt_d = 1.$$

- ▶ Si $\tilde{X} = (X_1, \dots, X_k)^\top$ avec $k \leq d$, alors la densité marginale $\tilde{f}$ de $\tilde{X}$ est donnée par

$$f_{\tilde{X}}(t_1, \dots, t_k) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_X(t_1, \dots, t_k, u_{k+1}, \dots, u_d) du_{k+1} \cdots du_d.$$

Achtung ! La connaissance des lois marginales de toutes les composantes de X ne suffit pas à déterminer sa loi.

Soit $X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \in \mathbb{R}^d$ tel que $\mathbb{E}[X_j^2] < +\infty$. On définit pour tout i, j :

$$\text{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])].$$

Soit $X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \in \mathbb{R}^d$ tel que $\mathbb{E}[X_j^2] < +\infty$

Matrice de variance-covariance $\Sigma \in \mathbb{R}^{d \times d}$ définie par

$$\Sigma_{i,j} = \text{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]$$

Il s'agit d'une matrice **symétrique, semi-définie positive**

Forme matricielle. On définit

$$\mathbb{E}[X] = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_d] \end{pmatrix} \in \mathbb{R}^d \quad \text{et} \quad \text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top] \in \mathbb{R}^{d \times d}$$

$$\text{càd } \text{Var}[X] = \begin{pmatrix} \text{Var}[X_1] & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_2, X_1) & \text{Var}[X_2] & \cdots & \text{Cov}(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_d, X_1) & \text{Cov}(X_d, X_2) & \cdots & \text{Var}[X_d] \end{pmatrix} \in \mathbb{R}^{d \times d}$$

Achtung ! Si X_i et X_j sont indépendantes, alors $\text{Cov}(X_i, X_j) = 0$, mais la réciproque est en général fausse (sauf dans le cas gaussien)

Transformation linéaire. Si X est un vecteur aléatoire de moyenne μ et de covariance Σ , et A une matrice réelle $p \times d$, alors le vecteur aléatoire AX est de moyenne $A\mu$ et de matrice de covariance $A\Sigma A^\top$.

Fonction caractéristique. Soit X un vecteur aléatoire de $\mathbb{R}^d$. Sa fonction caractéristique est donnée pour tout $u \in \mathbb{R}^d$ par

$$\Phi_X(u) = \mathbb{E} \left[e^{iu^\top X} \right] = \mathbb{E} \left[e^{i(u_1 X_1 + \dots + u_d X_d)} \right]$$

Indépendance. X et Y sont indépendants ssi l'une des conditions suivantes est vérifiée

- ▶ $\mathbb{P}[(X, Y) \in A \times B] = \mathbb{P}[X \in A] \mathbb{P}[Y \in B]$ pour tout borélien A et B
- ▶ $\Phi_{(X, Y)}(u) = \Phi_X(u_1) \Phi_Y(u_2)$ pour $u = (u_1, u_2)$
- ▶ $f_{(X, Y)}(x, y) = f_X(x) f_Y(y)$

1. Vecteurs aléatoires
2. Vecteurs aléatoires gaussiens
3. Convergences
4. Théorèmes asymptotiques
5. Théorème de Cochran et applications

Définition. Soit X un vecteur aléatoire de $\mathbb{R}^d$. On dit que X est un *vecteur gaussien* de $\mathbb{R}^d$ si et seulement si toute combinaison linéaire de ses composantes est une variable aléatoire réelle gaussienne, i.e.

$$\forall u \in \mathbb{R}^d, \quad u^\top X \sim \mathcal{N}(u^\top m, \sigma_u^2).$$

Exemple : un vecteur X dont les composantes X_i sont indépendantes et suivent une loi normale est un vecteur gaussien.

Théorème (fonction caractéristique). Soit X un vecteur aléatoire de $\mathbb{R}^d$. Les assertions suivantes sont équivalentes :

1. X est un vecteur gaussien de moyenne m et de matrice de covariance Γ .
2. La fonction caractéristique de X est donnée, pour tout $u \in \mathbb{R}^d$, par

$$\Phi_X(u) = \mathbb{E}[e^{iu^\top X}] = \exp\left(iu^\top m - \frac{1}{2}u^\top \Gamma u\right).$$

Corollaires.

- ▶ Si $X \sim \mathcal{N}(m, \Gamma)$ à valeurs dans $\mathbb{R}^d$, si A est une matrice $p \times d$ et $b \in \mathbb{R}^p$, alors

$$AX + b \sim \mathcal{N}(Am + b, A\Gamma A^\top).$$

- ▶ Si $X \sim \mathcal{N}(m, \Gamma)$ et que Γ est inversible, alors

$$\Gamma^{-1/2}(X - m) \sim \mathcal{N}(0, I_d).$$

Théorème (densité d'un vecteur gaussien). Soit $X \in \mathbb{R}^d$ tel que $X \sim \mathcal{N}(m, \Gamma)$. Si Γ est inversible, la densité de X est donnée par

$$f(x) = \frac{1}{\sqrt{(2\pi)^d |\det(\Gamma)|}} \exp\left(-\frac{1}{2}(x - m)^\top \Gamma^{-1}(x - m)\right).$$

Théorème (indépendance). Pour tout vecteur gaussien $X = (X_1, \dots, X_d)$ de $\mathbb{R}^d$, les propriétés suivantes sont équivalentes :

- ▶ Les composantes $X_1, \dots, X_d$ sont indépendantes.
- ▶ La matrice de covariance Γ de X est diagonale.

Corollaire : $\forall i \neq j$, X_i et X_j sont indépendants ssi $\text{Cov}(X_i, X_j) = 0$.

1. Vecteurs aléatoires
2. Vecteurs aléatoires gaussiens
3. Convergences
4. Théorèmes asymptotiques
5. Théorème de Cochran et applications

Extension directe au cas vectoriel

Convergence en loi. Soit (X_n) une suite de vecteurs aléatoires. On dit que X_n converge en loi vers X si, pour tout point de continuité t de F_X , on a

$$F_{X_n}(t) \xrightarrow{n \rightarrow +\infty} F_X(t)$$

où $F_X(t) = \mathbb{P}(X_1 \leq t_1, \dots, X_d \leq t_d)$

Proposition : (X_n) converge en loi vers X si et seulement si, pour toute fonction continue bornée h , $\mathbb{E}[h(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[h(X)]$.

Convergence en probabilité. Soit (X_n) une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^d$ que l'on munit d'une norme $\|\cdot\|$.

On dit que X_n converge en probabilité vers X si, pour tout $\varepsilon > 0$,

$$\mathbb{P}[\|X_n - X\| > \varepsilon] \xrightarrow{n \rightarrow +\infty} 0.$$

Remarque : comme toutes les normes sont équivalentes dans $\mathbb{R}^d$, si (X_n) converge pour une norme, il converge pour toute autre norme.

Convergence presque sûre. On dit que (X_n) converge presque sûrement vers X si

$$\mathbb{P}\left[\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}\right] = 1.$$

Remarque (critère de Borel–Cantelli) : si, pour tout $\varepsilon > 0$, on a

$$\sum_{n \geq 1} \mathbb{P}[\|X_n - X\| > \varepsilon] < +\infty,$$

alors $X_n \rightarrow X$ presque sûrement.

Lien entre les modes de convergence.

$$X_n \xrightarrow{\text{p.s.}} X \implies X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{\mathcal{L}} X,$$

$$X_n \xrightarrow{L^2} X \implies X_n \xrightarrow{\mathbb{P}} X \implies X_n \xrightarrow{\mathcal{L}} X.$$

1. Vecteurs aléatoires
2. Vecteurs aléatoires gaussiens
3. Convergences
4. Théorèmes asymptotiques
5. Théorème de Cochran et applications

Théorème de continuité. Soit $g : D \rightarrow \mathbb{R}^{d'}$ une fonction continue définie sur un ouvert $D \subset \mathbb{R}^d$. Soit X un vecteur aléatoire à valeurs dans D et (X_n) une suite de vecteurs aléatoires dans $\mathbb{R}^d$.

Si $X_n \rightarrow X$ (dans un certain mode de convergence), alors $g(X_n) \rightarrow g(X)$ dans le même mode, i.e.

$$X_n \xrightarrow{\mathcal{L}} X \implies g(X_n) \xrightarrow{\mathcal{L}} g(X),$$

$$X_n \xrightarrow{\mathbb{P}} X \implies g(X_n) \xrightarrow{\mathbb{P}} g(X),$$

$$X_n \xrightarrow{\text{p.s.}} X \implies g(X_n) \xrightarrow{\text{p.s.}} g(X).$$

Théorème de Slutsky. Soit (X_n) une suite de vecteurs aléatoires convergeant en loi vers X et (Y_n) une suite convergeant en probabilité vers une constante c . Alors

$$X_n + Y_n \xrightarrow{\mathcal{L}} X + c,$$

$$X_n Y_n \xrightarrow{\mathcal{L}} Xc, \quad Y_n X_n \xrightarrow{\mathcal{L}} cX.$$

Remarque : les dimensions de X et c doivent être compatibles.

- **Loi des grands nombres :** Soit $(X_i)_{i \geq 1}$ une suite de v.a. i.i.d. de $\mathbb{R}^d$ telle que $\mathbb{E}[\|X_1\|] < +\infty$. Alors

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\text{p.s.}} \mathbb{E}[X_1].$$

- **Théorème central limite :** Si de plus $\mathbb{E}[\|X_1\|^2] < +\infty$, alors

$$\sqrt{n}(\bar{X}_n - \mathbb{E}[X_1]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \text{Var}[X_1]).$$

Cas général. Soit (X_n) une suite de v.a., X un vecteur aléatoire, (v_n) une suite de réels telle que $v_n \rightarrow +\infty$, et $m \in D \subset \mathbb{R}^d$ tel que

$$v_n(X_n - m) \xrightarrow{\mathcal{L}} X.$$

Si $g : D \rightarrow \mathbb{R}^{d'}$ est continûment différentiable, alors

$$v_n(g(X_n) - g(m)) \xrightarrow{\mathcal{L}} J_g(m)X,$$

où $J_g(m)$ est la matrice Jacobienne de g en m .

Cas particulier. Si $\sqrt{n}(X_n - m) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma)$ et $g : D \rightarrow \mathbb{R}^{d'}$ est C^1 , alors

$$\sqrt{n}(g(X_n) - g(m)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, J_g(m)\Sigma J_g(m)^\top).$$

1. Vecteurs aléatoires
2. Vecteurs aléatoires gaussiens
3. Convergences
4. Théorèmes asymptotiques
5. Théorème de Cochran et applications

Théorème. Soit $X \in \mathbb{R}^d$ un vecteur aléatoire suivant une loi

$$X \sim \mathcal{N}(m, \sigma^2 I_d), \quad \sigma^2 > 0.$$

Soit $\mathbb{R}^d = E_1 \oplus E_2 \oplus \dots \oplus E_p$ une décomposition en somme directe de p sous-espaces vectoriels orthogonaux, de dimensions respectives $d_1, \dots, d_p$. On note P_k le projecteur orthogonal sur E_k et $Y_k = P_k X$ la projection de X sur E_k . Alors :

1. Les $Y_1, \dots, Y_p$ sont des vecteurs gaussiens indépendants, et

$$Y_k \sim \mathcal{N}(P_k m, \sigma^2 P_k).$$

2. Les variables $\|Y_1 - P_1 m\|^2, \dots, \|Y_p - P_p m\|^2$ sont indépendantes, et

$$\frac{\|Y_k - P_k m\|^2}{\sigma^2} \sim \chi_{d_k}^2.$$

Corollaire. Soient $X_1, \dots, X_n$ i.i.d. avec $X_1 \sim \mathcal{N}(m, \sigma^2)$. On a alors :

1. $\bar{X}_n \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$
2. $\frac{n-1}{\sigma^2} S_n^2 \sim \chi_{n-1}^2$ pour $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$
3. $\bar{X}_n$ et S_n^2 sont indépendantes

Proposition. Soient Z, Z_p, Z_q trois v.a. telles que :

1. $Z = Z_p + Z_q$,
2. $Z_p \sim \chi_p^2$ et $Z_q \sim \chi_q^2$,
3. Z_p et Z_q indépendantes,

alors

$$Z \sim \chi_{p+q}^2.$$

Corollaire (Test du χ^2 d'ajustement). Soit X une v.a. à valeurs dans $\{a_1, \dots, a_K\}$ avec $\mathbb{P}[X = a_k] = p_k > 0$. Soient $X_1, \dots, X_n$ i.i.d. de même loi que X , et $E_k = \sum_{i=1}^n \mathbb{1}_{\{X_i = a_k\}}$.

Alors

$$\sum_{k=1}^K \frac{(E_k - np_k)^2}{np_k} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \chi_{K-1}^2.$$