PC 1 (Rappels de Probabilités)

1 Lemme de Slutsky

1. Donner un exemple de suites (X_n) et (Y_n) telles que $X_n \overset{\text{loi}}{\to} X$ et $Y_n \overset{\text{loi}}{\to} Y$, mais $X_n + Y_n$ ne converge pas en loi vers $X + Y$.
2. Soient (X_n), (Y_n) deux suites de variables aléatoires réelles, X et Y des variables aléatoires réelles, telles que
$$X_n \overset{\text{loi}}{\to} X, \qquad Y_n \overset{\mathbb{P}}{\to} c,$$
où c est une constante. Montrer que le couple (X_n, Y_n) converge en loi vers (X, c).
3. En déduire que $(X_n + Y_n)$ converge en loi vers $X + c$ et $X_n Y_n$ converge en loi vers cX.

Corrigé :

1. Soit $X_n = X + 1/n$ pour tout n où X est une loi symétrique (X et $-X$ ont même loi, e.g. $X \sim \mathcal{N}(0, \sigma^2)$). On pose $Y_n = -X_n$. Alors $X_n + Y_n = X_n - X_n = 0$ presque-sûrement. D'autre part, X_n converge en loi vers X tandis que $Y_n = -X_n$ converge en loi vers $-X$, et donc également vers X.
   Si la proposition était vraie, $X_n + Y_n$ convergerait en loi, à la fois vers 0 et vers $2X$, ce qui n'est pas possible dès que X prend des valeurs non nulles.
2. Soient s, t deux réels
$$\begin{aligned}
\left|\mathbb{E}(e^{itX_n + isY_n}) - \mathbb{E}(e^{itX+isc})\right| &= \left|\mathbb{E}(e^{itX_n+isY_n}) - \mathbb{E}(e^{itX_n+isc}) + \mathbb{E}(e^{itX_n+isc}) - \mathbb{E}(e^{itX+isc})\right| \\
&\leq \left|\mathbb{E}(e^{itX_n+isY_n} - e^{itX_n+isc})\right| + \left|\mathbb{E}(e^{itX_n+isc} - \mathbb{E}(e^{itX+isc})\right| \\
&\leq \mathbb{E}(|e^{itX_n}||e^{isY_n} - e^{isc}|) + |e^{isc}| \left|\mathbb{E}(e^{itX_n} - e^{itX})\right| \\
&= \mathbb{E}(|e^{isY_n} - e^{isc}|) + \left|\mathbb{E}(e^{itX_n} - e^{itX})\right|.
\end{aligned}$$
   - La convergence en proba de Y_n vers c implique que e^{isY_n} converge en proba vers e^{isc}. D'autre part, comme e^{isY_n} est une v.a (complexe) bornée, alors e^{isY_n} converge vers e^{isc} dans L^1. On en conclut que le premier terme tend vers zéro.
   - Le deuxième terme tend vers zéro par hypothèse.
3. On vient de voir que (X_n, Y_n) converge en loi vers (X, c). Les applications $(x, y) \mapsto x + y$ et $(x, y) \mapsto xy$ sont continues. Puisque l'application d'une fonction continue conserve la convergence en loi, on en déduit que $X_n + Y_n \to X + c$ en loi et que $X_n Y_n \to cX$ en loi.

2 TCL pour la médiane

Soit $(X_1, X_2, \ldots, X_{2n+1})$ un échantillon de $2n+1$ v.a. i.i.d. uniformes sur $[0, 1]$, on note Y_n la médiane de l'échantillon. On s'attend à ce que (Y_n) tende vers $1/2$, nous allons montrer que $(Y_n - 1/2)$ a des fluctuations gaussiennes.

1. Se convaincre que Y_n a pour densité
$$(2n+1)\binom{2n}{n} x^n (1-x)^n \mathbf{1}_{x \in [0,1]}.$$

2. Déterminer la densité g_n de

$$Z_n = 2\sqrt{2n}\,(Y_n - 1/2)$$

et en déduire que Z_n converge en loi vers une $\mathcal{N}(0,1)$.

(On pourra utiliser le Théorème de Scheffé : si chaque Z_n a pour densité g_n et que la suite (g_n) converge simplement vers une densité g, alors (Z_n) converge en loi vers la loi de densité g.)

3. Comment généraliser ce résultat et cette preuve à des variables continues non-uniformes ?

Corrigé :

1. On s'autorise à manipuler les dx comme des vrais nombres, calculons $\mathbb{P}(Y_n \in (x, x+dx))$. Il faut choisir la médiane parmi les $2n+1$ et ensuite les n qui seront plus petites que x parmi les $2n$ variables restantes. Les autres n variables sont plus grandes que $x+dx$. En considérant dx suffisament petit pour que la probabilité que deux variables prennent leurs valeurs dans le même intervalle de longueur dx soit négligeable, on a donc

$$\mathbb{P}(Y_n \in (x, x+dx)) = dx \times (2n+1)\binom{2n}{n} x^n (1-x-dx)^n.$$

On définit la densité de Y_n comme la limite de $\mathbb{P}(Y_n \in (x, x+dx))/dx$ quand $dx \to 0$.

2. On effectue le changement de variable $Z_n = Z_n = 2\sqrt{2n}\,(Y_n - 1/2) = h(Y_n)$. Donc, [1]

$$\begin{aligned} g_n(z) &= \frac{1}{2\sqrt{2n}}(2n+1)\binom{2n}{n}\left(\frac{1}{2} + \frac{z}{\sqrt{2n}}\right)^n \left(\frac{1}{2} - \frac{z}{\sqrt{2n}}\right)^n \mathbf{1}_{-\sqrt{2n} \le z \le \sqrt{2n}} \\ &= \frac{1}{2\sqrt{2n}}(2n+1)\binom{2n}{n}\frac{1}{2^{2n}}\left(1 - \tfrac{z^2}{2n}\right)^n \mathbf{1}_{|z| \le \sqrt{2n}} \end{aligned}$$

avec Stirling ($n! \approx n^n e^{-n}\sqrt{2\pi n}$) on trouve que

$$\begin{aligned} \binom{2n}{n} &= \frac{2n!}{n!n!} \\ &\approx \frac{(2n)^{2n} e^{-2n}\sqrt{4\pi n}}{n^(2n) e^{-2n}(2\pi n)} \\ &\approx \frac{2^{2n}}{\sqrt{\pi n}} \end{aligned}$$

et donc,

$$\frac{1}{2\sqrt{2n}}(2n+1)\binom{2n}{n}\frac{1}{2^{2n}} \approx \frac{1}{\sqrt{2\pi}}$$

D'autre part, $\lim n \to \infty \left(1 - \frac{z^2}{2n}\right)^n = e^{-z^2/2}$. Donc, pour tout z

$$g_n(z) \to \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

On en déduit alors d'après le théorème de Scheffé que $Z_n \xrightarrow{\text{loi}} \mathcal{N}(0,1)$ et donc que $2\sqrt{2n}(Y_n - 0.5) \xrightarrow{\text{loi}} \mathcal{N}(0,1)$. On trouve finalement que

$$\sqrt{n}(Y_n - 0.5) \xrightarrow{\text{loi}} \mathcal{N}(0, 1/8).$$

1. Soit Y une v.a. réelle continue et $Z = h(Y)$ où h est une fonction dérivable, strictement croissante. Pour tout t,

$$F_Z(t) = \mathbb{P}(Z \le t) = \mathbb{P}(h^{-1}(Z) \le h^{-1}(t)) = F_Y(h^{-1}(t)).$$

Donc, $f_Z(t) = (h^{-1})'(t) f_Y(h^{-1}(t)) = f_Y(h^{-1}(t))/h'(h^{-1}(t))$.

3. Si l'on suppose que F (fonction de répartition) est un C^1-difféomorphisme (bijection dérivable), alors $\hat{m} = F^{-1}(Y_n)$ a même loi que la médiane empirique d'un échantillon de densité $f = F'$. En effet, $\xi_i = F^{-1}(X_i) \sim F$ puisque X_i suit une loi uniforme sur $[0,1]$. D'autre part, F étant strictement croissante, l'ordre est conservé et $\xi_{(i)} = F^{-1}(X_{(i)})$. En particulier, $\hat{m}$, la médiane empirique des ξ_i est la transformée par F^{-1} de Y_n, la médiane empirique des X_i. D'autre part, $F(m) = 0.5$ par définition et donc $m = F^{-1}(0.5)$.
   Du coup on utilise la méthode delta, sachant que $(F^{-1})' = 1/f(F^{-1})$ et on trouve
$$\sqrt{n}\left(F^{-1}(Y_n) - F^{-1}(0.5)\right) \to \mathcal{N}\left(0, \frac{1}{8}(F^{-1})'(0.5)\right)$$
   C'est à dire
$$\sqrt{n}\left(\hat{m} - m\right) \to \mathcal{N}\left(0, \frac{1}{8f(m)^2}\right)$$
   On vérifie donc que la variance de l'estimateur de la médiane est d'autant plus petite que la loi F est concentrée autour de la médiane m.

3 Estimateur de la variance

Soient $X_1, \ldots, X_n$ des variables aléatoires i.i.d., $X_i \sim f(\cdot - \theta)$, où f est une densité de probabilité sur $\mathbb{R}$ symétrique dont on note $\mu_k = \int_{\mathbb{R}} x^k f(x)\,dx$ les moments d'ordre $k = 2$ et $k = 4$. On note $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Montrer que l'estimateur $\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ de la variance des X_i vérifie un théorème central limite.

Indication : on montrera d'abord que l'on peut se ramener au cas où $\theta = 0$, puis on exprimera l'estimateur comme une transformation de $S_n = \frac{1}{n}\sum_{i=1}^n X_i^2$ et de $\bar{X}_n$.

Corrigé : On n'utilise pas le fait que f est symétrique mais on montre tout d'abord que l'on peut se ramener au cas $\mathbb{E}(X) = 0$. En effet, posons $Y_i = X_i - \theta$. Alors pour toute fonction test h
$$\mathbb{E}\left(h(Y_i)\right) = \mathbb{E}\left(h(X_i - \theta)\right) = \int h(x-\theta)\ f(x-\theta)dx = \int h(y)f(y)dy.$$
Donc Y_i a pour densité f. Remarquons aussi que l'on a
$$\begin{aligned} \mathbb{E}(Y_i) &= \mathbb{E}(X_i) - \theta \\ \mathrm{Var}(Y_i) &= \mathrm{Var}(X_i) \\ \hat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \end{aligned}$$
de sorte que l'on peut considérer dans la suite que $\mathbb{E}(X_i) = 0$.
$$\begin{aligned} \hat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^n \left(X_i^2 + (\bar{X}_n)^2 - 2X_i\bar{X}_n\right) \\ &= \frac{1}{n}\sum_{i=1}^n X_i^2 + (\bar{X}_n)^2 - 2\bar{X}_n\frac{1}{n}\sum_{i=1}^n X_i \\ &= \frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2 \end{aligned}$$
Etudions alors
$$\sqrt{n}\left(\hat{\sigma}_n^2 - \sigma^2\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma^2\right) - \sqrt{n}(\bar{X}_n)^2$$
$\mathbb{E}(X_i^2) = \sigma^2$ puisque $\mathbb{E}(X_i) = 0$. D'autre part, X_i^2 a bien une variance finie puisque μ_4, le moment d'ordre 4, existe. Le TCL pour des v.a. i.i.d. prouve la convergence en loi de la moyenne des carrés :
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma^2\right) \xrightarrow{\text{loi}} \mathcal{N}\left(0, \mathbb{E}((X_1^2 - \mathbb{E}(X_1^2))^2)\right)$$

D'autre part, $\bar{X}_n = \mathcal{O}_P(1/\sqrt(n)$. Donc $\sqrt{n}(\bar{X}_n)^2$ converge vers 0 en probabilité. On peut utiliser Slutzky pour conclure :

$$\sqrt{n}\left(\hat{\sigma}_n^2 - \sigma^2\right) \xrightarrow{\text{loi}} \mathcal{N}\left(0, \mu_4 - \mu_2^2\right)$$

4 Taux de défaillance

Une chaîne de production doit garantir une qualité minimale de ses produits. En particulier, elle doit garantir que la proportion θ des produits défaillants reste inférieure à un taux fixé par le client. Un échantillon de n produits est prélevé et analysé. On note $\hat{\theta}_n$ la proportion de produits défectueux dans l'échantillon.

1. Proposer un modèle statistique pour ce problème. Quelle est la loi de $n\hat{\theta}_n$?
2. Quelle information donne la loi des grands nombres et le théorème central limite sur le comportement asymptotique de $\hat{\theta}_n$?
3. On donne $\mathbb{P}(N > 1.64) = 5\%$ pour $N \sim \mathcal{N}(0, 1)$. En déduire ε_n (dépendant de n et θ) tel que $\mathbb{P}(\theta \geq \hat{\theta}_n + \varepsilon_n) \overset{n\to\infty}{\to} 5\%$.
4. La valeur ε_n précédente dépend de θ. A l'aide du lemme de Slutsky, donner ε'_n ne dépendant que de n et $\hat{\theta}_n$ tel que $\mathbb{P}(\theta \geq \hat{\theta}_n + \varepsilon'_n) \overset{n\to\infty}{\to} 5\%$.

Corrigé :

1. Soit X_i une v.a. qui vaut 1 (défaillance) ou 0 (bon état) selon l'état du produit i ; c'est une v.a. de Bernoulli de paramètre θ. Alors la proportion de produits défaillants est $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^n X_i$. On supposera de plus que les produits sont indépendants.
   Par suite, $n\hat{\theta}_n = \sum_{i=1}^n X_i$ représente le nombre de produit défectueux et suit une loi Binomiale de paramètres (n, p).
2. Par la LGN, $\hat{\theta}_n$ converge p.s. vers $\mathbb{E}(X_1) = \theta$. On dit que l'estimateur est fortement consistant. Le TCL dit que $\sqrt{n}\left(\hat{\theta}_n - \theta\right)$ converge en loi vers une gaussienne centrée et de variance $\theta(1-\theta)$. Autrement dit, $\hat{\theta}_n = \theta + \frac{\sqrt{\theta(1-\theta)}}{\sqrt{n}}\mathcal{N}(0, 1)$ (en loi).
3. On cherche un seuil tel que la probabilité pour que le vrai taux de défaillance soit supérieur à ce seuil est faible (ici 5%). On a lorsque n est grand

$$\begin{aligned}\mathbb{P}\left(\theta \geq \hat{\theta}_n + \epsilon_n\right) &= \mathbb{P}\left(\sqrt{n}\left(\hat{\theta}_n - \theta\right)/\sqrt{\theta(1-\theta)} \leq -\sqrt{n}\epsilon_n/\sqrt{\theta(1-\theta)}\right) \\ &\approx \Phi(-\sqrt{n}\epsilon_n/\sqrt{\theta(1-\theta)}\end{aligned}$$

   où Φ est la fonction de répartition d'une $\mathcal{N}(0, 1)$.
   On veut que cette intégrale soit de l'ordre de 5% : il suffit de prendre $-\sqrt{n}\epsilon_n/\sqrt{\theta(1-\theta)} = -1.64$. Soit

$$\epsilon_n = 1.64\sqrt{\frac{\theta(1-\theta)}{n}}$$

4. Slutsky dit que le TCL reste encore vrai si on remplace la variance $\theta(1-\theta)$ par $\hat{\theta}_n(1-\hat{\theta}_n)$ (qui converge p.s. et donc en proba vers $\theta(1-\theta)$). Donc en procédant de même, on a

$$\epsilon_n = 1.64\sqrt{\frac{\hat{\theta}_n(1-\hat{\theta}_n)}{n}}$$

   On mesure une probabilité de défaillance $\hat{\theta}_n$. En vue du contrat avec le client, on cherche une valeur seuil qui est dépassée rarement. On cherche donc une zone de confiance de la forme $\{\theta > \hat{\theta}_n + \epsilon_n\}$.

Exercice bonus

5 Pas de convergence en proba dans le TCL

Soit $(X_n)_n$ une suite de variables aléatoires i.i.d. centrées de variance $\sigma^2 > 0$. Soit

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} X_j \,.$$

Par le théorème central limite, cette variable converge en loi vers la loi normale centrée réduite, c'est-à-dire, pour tout $t \in \mathbb{R}$, on a $\lim_{n\to+\infty} \mathbb{E}(e^{itZ_n}) = e^{-\frac{t^2}{2}}$. L'objet de cet exercice est de montrer que la suite (Z_n) ne peut pas converger en probabilité.

1. Calculer la fonction caractéristique de $Z_{2n} - Z_n$ et montrer que cette différence converge en loi.
2. On suppose que (Z_n) converge en probabilité vers une variable Z_∞. Que peut-on dire de la limite en probabilité de $(Z_{2n} - Z_n)$? Conclure.

Corrigé :

Contre-exemple : soit U une v.a. à valeur dans $\{0, 1\}$ telle que $\mathbb{P}(U = 1) = \mathbb{P}(U = 0) = 0.5$. Si n est pair, on pose $X_n = 0$ si $U = 0$ et 1 sinon ; si n est impair, on pose $X_n = 0$ si $U = 1$ et 1 sinon.
Alors X_n converge en loi vers $0.5\delta_0 + 0.5\delta_1$. D'autre part, $\mathbb{P}(|X_{2n} - X_1| = 1) = 1$ donc il n'y a pas convergence en proba.

1. Par définition,

$$\begin{aligned}
\sigma(Z_{2n} - Z_n) &= \frac{1}{\sqrt{2n}} \sum_{k=1}^{2n} X_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \\
&= \frac{1}{\sqrt{2n}} \sum_{k=n+1}^{2n} X_k + \left(\frac{1}{\sqrt{2n}} - \frac{1}{\sqrt{n}}\right) \sum_{k=1}^{n} X_k \\
&= \frac{1}{\sqrt{2n}} \sum_{k=n+1}^{2n} X_k + \left(\frac{1}{\sqrt{2}} - 1\right) \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k
\end{aligned}$$

Soit $t \in \mathbb{R}$. En observant que les deux sommes sont indépendantes, il vient

$$\begin{aligned}
\mathbb{E}\left(\exp(it(Z_{2n} - Z_n))\right) &= \mathbb{E}\left(\exp\left\{it\sigma^{-1}\left(\frac{1}{\sqrt{2n}} \sum_{k=n+1}^{2n} X_k + \left(\frac{1}{\sqrt{2}} - 1\right) \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k\right)\right\}\right) \\
&= \mathbb{E}\left(\exp\left\{it\sigma^{-1}\left(\frac{1}{\sqrt{2n}} \sum_{k=n+1}^{2n} X_k\right)\right\}\right) \mathbb{E}\left(\exp\left\{it\sigma^{-1}\left(\frac{1}{\sqrt{2}} - 1\right) \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k\right\}\right).
\end{aligned}$$

Puisque les v.a. $(X_k)_k$ ont même loi, il vient

$$\begin{aligned}
\mathbb{E}\left(\exp(it(Z_{2n} - Z_n))\right) &= \mathbb{E}\left(\exp\left\{it\sigma^{-1}\left(\frac{1}{\sqrt{2n}} \sum_{k=1}^{n} X_k\right)\right\}\right) \mathbb{E}\left(\exp\left\{it\sigma^{-1}\left(\frac{1}{\sqrt{2}} - 1\right) \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k\right\}\right) \\
&= \phi_{Z_n}(t/\sqrt{2})\ \phi_{Z_n}(t(1/\sqrt{2} - 1))
\end{aligned}$$

Ainsi, la limite de cette quantité est

$$\exp\left(-\frac{t^2}{2}\left(\frac{1}{2} + \left(\frac{1}{\sqrt{2}} - 1\right)^2\right)\right) = \exp\left(-t^2\left(1 - 1/\sqrt{2}\right)\right)$$

On en déduit que la limite en loi de $(Z_{2n} - Z_n)_n$ est une loi gaussienne centrée et de variance $2 - \sqrt{2}$.

2. Si $(Z_n)_n$ converge en probabilité vers Z, alors $(Z_{2n})_n$ converge aussi en probabilité vers Z.[2]
En écrivant

$$\begin{aligned}\mathbb{P}\left(|Z_{2n}-Z_n|\geq\delta\right)&=\mathbb{P}\left(|(Z_{2n}-Z)+(Z-Z_n)|\geq\delta\right)\leq\mathbb{P}\left(|Z_{2n}-Z|+|Z-Z_n|\geq\delta\right)\\&\leq\mathbb{P}\left(|Z_{2n}-Z|\geq\delta/2\right)+\mathbb{P}\left(|Z-Z_n|\geq\delta/2\right)\end{aligned}$$

on voit que la convergence en probabilité des suites $(Z_n)_n$ et $(Z_{2n})_n$ vers Z entraine la convergence en probabilité de $(Z_{2n}-Z_n)_n$ vers zero.[3]
Maintenant, si $Z_{2n}-Z_n\to 0$ en probabilité, alors $Z_{2n}-Z_n\to 0$ également en loi,[4] ce qui contredit le résultat de 1).
On ne peut donc avoir $\mathbb{Z}_n\xrightarrow{\mathbb{P}}Z$.

---

2. En effet pour tout $\delta,\epsilon>0$, il existe N tel que pour tout $n\geq N$, $\mathbb{P}(|Z_n-Z|\geq\delta)\leq\epsilon$. On en déduit que pour tout $n\geq N/2$, $\mathbb{P}(|Z_{2n}-Z|\geq\delta)\leq\epsilon$. Ce qui prouve que $\lim_n\mathbb{P}(|Z_{2n}-Z|\geq\delta)=0$.
3. Plus généralement,
  - si $X_n\to X$ et $Y_n\to Y$ en probabilité alors $X_n+Y_n\to X+Y$ en probabilité.
  - si $X_n\to X$ en probabilité, alors cX_n converge vers cX en probabilité.

En effet,

$$\begin{aligned}\mathbb{P}\left(|X_n+Y_n-X-Y|\geq\delta\right)&\leq\mathbb{P}\left(|X_n-X|+|Y_n-Y|\geq\delta\right)\\&\leq\mathbb{P}\left(|X_n-X|\geq\delta/2\right)+\mathbb{P}\left(|Y_n-Y|\geq\delta/2\right)\to 0.\end{aligned}$$

De plus,

$$\mathbb{P}\left(|cX_n-cX|\geq\delta\right)=\mathbb{P}\left(|X_n-X|\geq\delta/|c|\right)\to 0.$$

4. Si X_n converge en probabilité vers X, alors X_n converge en loi vers X.
Démonstration : D'après le théorème porte-manteau, il suffit de montrer que, pour toute fonction f bornée et *uniformément* continue, $\mathbb{E}(f(X_n))\to\mathbb{E}(f(X))$.
Soit donc f bornée et uniformément continue et $\varepsilon>0$. Il existe, f étant uniformément continue, $\alpha>0$ tel que $|f(x)-f(y)|\leq\varepsilon$ si $|x-y|\leq\alpha$. On a alors

$$\begin{aligned}|\mathbb{E}(f(X_n))-\mathbb{E}f(X))|&\leq\mathbb{E}(|f(X_n))-f(X)|1_{|X_n-X|\leq\alpha}+\mathbb{E}(|f(X_n))-f(X)|1_{|X_n-X|>\alpha}\\&\leq\varepsilon+2||f||P(|X_n-X|>\alpha)\end{aligned}$$

d'où $\overline{\lim}_n|\mathbb{E}(f(X_n))-\mathbb{E}f(X))|\leq\varepsilon$ et $\mathbb{E}(f(X_n))\to\mathbb{E}(f(X))$.