# Aléatoire

# MAP 311 - X2013

# Leçon 7

*Mots clés :*

1. Applications du Théorème de la Limite Centrale
   - aux sondages
   - aux méthodes de Monte-Carlo
2. Etude des échantillons gaussiens
   - Transformation affine
   - Projection, rotation, théorème de Cochran
3. A propos de la corrélation
4. Statistique gaussienne : estimation de la moyenne d'un échantillon
   - Variance connue ou inconnue
   - Estimateur du maximum de vraisemblance

# Applications du Théorème de la Limite Centrale

**Théorème (rappel leçon 6).** Soit $(X_n)_n$ une suite de v.a. réelles indépendantes, de même loi, de carré intégrable : $m = \mathbb{E}(X_1)$ et $\sigma^2 = \mathbb{V}\mathrm{ar}(X_1)$. Alors

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - m\right)$$

**converge en loi (quand $n \to +\infty$) vers une variable aléatoire de loi normale $\mathcal{N}(0, \sigma^2)$.**

**Corollaire (rappel leçon 6).** Pour tout $a \leq b$,

$$\lim_{n\to+\infty} \mathbb{P}\Big(a \leq \frac{\sqrt{n}}{\sigma}\big(\frac{1}{n}\sum_{k=1}^{n} X_k - m\big) \leq b\Big) = \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx.$$

## Application du TLC en statistique : les sondages

**Objectif :** estimer un paramètre inconnu à partir de données observées et produire une fourchette d'estimation (intervalles de confiance)

# Modélisation probabiliste du sondage

- Population :
  - **N** individus dont $\mathbf{N_1}$ sont favorables (OUI).
  - Objectif : **estimer** $p = \frac{N_1}{N}$**.**
- Sous-échantillon de taille **n** individus tirés au hasard parmi $N$.
  - Chaque réponse notée $X_i = 1$ (si OUI) et $X_i = 0$ (si NON).
  - Tirage sans remise : $\mathbb{P}(\sum_{i=1}^{n} X_i = n_1) = \frac{\binom{N_1}{n_1}\binom{N-N_1}{n-n_1}}{\binom{N}{n}}$.
  - Mais $n \ll N$ ➟ $\mathbb{P}(\sum_{i=1}^{n} X_i = n_1) \approx \binom{n}{n_1} p^{n_1}(1-p)^{n-n_1}$ (p.28, *Aléatoire*).
  - "Tout se passe comme si" **les v.a.r.** $(X_i)_i$ **étaient i.i.d. de loi de Bernoulli** $\mathcal{B}(p)$, avec $p = \mathbb{E}(X_1)$ et $\mathbb{V}\mathrm{ar}(X_1) = p(1-p)$.

## Résultats asymptotiques de fluctuations d'échantillon pour $n$ grand :

- LGN : $\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \overset{p.s.}{\to} p$.

- TLC : pour tout $a < b$

$$\mathbb{P}\Big(\frac{\sqrt{\mathbf{n}}}{\sqrt{\mathbf{p}(\mathbf{1}-\mathbf{p})}}(\hat{\mathbf{p}}_\mathbf{n} - \mathbf{p}) \in [\mathbf{a}, \mathbf{b}]\Big) \approx \int_\mathbf{a}^\mathbf{b} \frac{\mathbf{e}^{-\mathbf{x}^2/2}}{\sqrt{\mathbf{2}\pi}}\mathbf{dx} = \boldsymbol{\Pi}(\mathbf{b}) - \boldsymbol{\Pi}(\mathbf{a}).$$

➠ **Intervalles de confiance (avec un point de vue asymptotique) :**

**avec probabilité proche de $\Pi(b) - \Pi(a)$,**

$$\hat{\mathbf{p}}_\mathbf{n} - \mathbf{b}\frac{\sqrt{\mathbf{p}(\mathbf{1}-\mathbf{p})}}{\sqrt{\mathbf{n}}} \leq \mathbf{p} \leq \hat{\mathbf{p}}_\mathbf{n} - \mathbf{a}\frac{\sqrt{\mathbf{p}(\mathbf{1}-\mathbf{p})}}{\sqrt{\mathbf{n}}}.$$

Si $b = c_\alpha$ et $a = -c_\alpha$ avec $\Pi(c_\alpha) - \Pi(-c_\alpha) = \alpha \approx 1$, **intervalles de confiance (asymptotiques) de $p$ au niveau $\alpha$.** Calcul de $\Pi(.)$ p.107.

**La règle du $\pm 1/\sqrt{n}$ :** pour $\mathbf{c}_\alpha = \mathbf{2}$, on a $\mathbf{c}_\alpha\sqrt{\mathbf{p}(\mathbf{1}-\mathbf{p})} \leq \mathbf{2}\sqrt{\frac{\mathbf{1}}{\mathbf{2}}(\mathbf{1}-\frac{\mathbf{1}}{\mathbf{2}})} = \mathbf{1}$.
Alors pour "$n$ grand", l'évènement

$$\big|\hat{\mathbf{p}}_\mathbf{n} - \mathbf{p}\big| \leq \frac{\mathbf{1}}{\sqrt{\mathbf{n}}}$$

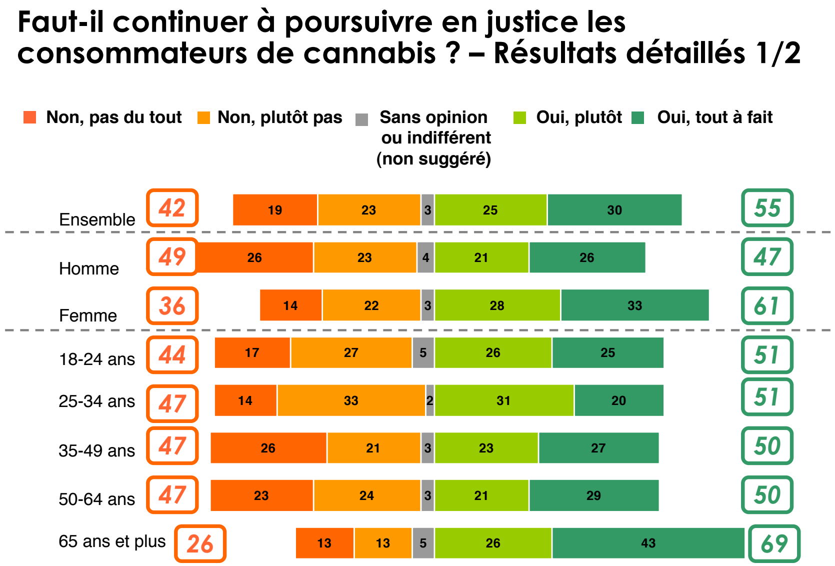
se produit environ avec probabilité au moins égale à

$$\Pi(c_\alpha) - \Pi(-c_\alpha) = 0.9792 - (1 - 0.9792) = 95.84\%.$$

**Exemple.** Sondage d'opinion *TNS Sofres* : LES FRANÇAIS ET LE CANNABIS (thèmes de la dépénalisation et de la légalisation). Réalisé les 17-18 juin 2011. Source : http://www.tns-sofres.com/points-de-vue/21C11AD137A44C9294196E235AAD9EBD.aspx

▷ Echantillon de $n = 956$ personnes représentatif de l'ensemble de la population âgée de 18 ans et plus.

▷ QUESTION : **Faut-il continuer à poursuivre en justice les consommateurs de cannabis ?** 55% des interrogés répondent **OUI**.




**Cet échantillon représente-t-il correctement l'opinion des français ?**

**Oui, très probablement : avec environ 99%,** $p \geq \hat{p}_n - 2.33\sqrt{\frac{1}{4}}\sqrt{\frac{1}{n}} = 51.32\%$

Attention aux conclusions hâtives sans intervalle de confiance et sans niveau de risque.

**Exemple.** Sondage d'opinion *TNS Sofres* : Les desserts préférés des Français . Réalisé entre le 28 septembre et le 1er octobre 2012.
Source : http://www.tns-sofres.com/points-de-vue/3CD4D2A3646648DDA91F8129138BEED7.aspx

▷ Echantillon de $n = 1000$ personnes représentatif de l'ensemble de la population âgée de 18 ans et plus.

▷ Question : **Quel est votre dessert préféré ?**

Les desserts préférés selon le sexe.

| Homme | | |
|---|---|---|
| 1 | Crêpes, Mousse au chocolat | 24% |
| 3 | Île flottante | 23% |
| 4 | Fondant au chocolat, Tarte aux pommes | 22% |

| Femme | | |
|---|---|---|
| 1 | Fondant au chocolat | 27% |
| 2 | Crêpes, Mousse au chocolat, Île flottante | 22% |
| 5 | Tarte aux pommes, Tiramisù | 21% |

Fiabilité des résultats :

- plus complexe car réponses multiples
- régions de confiance multidimensionnelles

# Application du TLC aux méthodes de Monte-Carlo

Rappel dans le cas de variables aléatoires de loi uniforme.

**Proposition.** Soit $(X_n)_n$ **une suite de v.a. indépendantes de loi uniforme sur** $[0,1]^d$ et $g : [0,1]^d \mapsto \mathbb{R}$ une fonction intégrable. Alors, quand $n \to \infty$,

$$\mathbf{I_n} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \mathbf{g(X_j)} \overset{\mathbf{p.s.}}{\to} \mathbf{I} = \int_{\mathbf{[0,1]^d}} \mathbf{g(x)dx_1 \dots dx_d}.$$

Si de plus $g$ est de carré intégrable et $\sigma^2 = \int_{[0,1]^d} g^2(x)dx_1 \dots dx_d - I^2$, alors

$$\mathbb{P}\Big(\frac{\sqrt{\mathbf{n}}}{\sigma}|\mathbf{I_n} - \mathbf{I}| \leq \mathbf{a}\Big) \approx \int_{-\mathbf{a}}^{\mathbf{a}} \frac{\mathbf{e^{-x^2/2}}}{\sqrt{\mathbf{2}\pi}}\mathbf{dx}.$$

**Vitesse de convergence $\sqrt{n}$ indépendante de la dimension $d$.**

A la différence des sondages, on ne sait pas majorer $\sigma^2$ (en général).

Estimation de $\sigma^2$ avec le même échantillon :

$$\sigma_n^2 = \frac{1}{n}\sum_{j=1}^{n} g^2(X_j) - \Big(\frac{1}{n}\sum_{j=1}^{n} g(X_j)\Big)^2 \overset{p.s.}{\to} \mathbb{E}(g^2(X)) - (\mathbb{E}(g(X))^2 = \sigma^2.$$

# Les intervalles de confiance restent-ils valides ?

Validité du TLC avec $\sigma_n$ au lieu de $\sigma$ : $\mathbb{P}\Big(\frac{\sqrt{\mathbf{n}}}{\sigma_\mathbf{n}}|\mathbf{I_n} - \mathbf{I}| \leq \mathbf{a}\Big) \approx \int_{-\mathbf{a}}^{\mathbf{a}} \frac{\mathbf{e}^{-\mathbf{x}^2/2}}{\sqrt{2\pi}} \mathbf{dx}$?

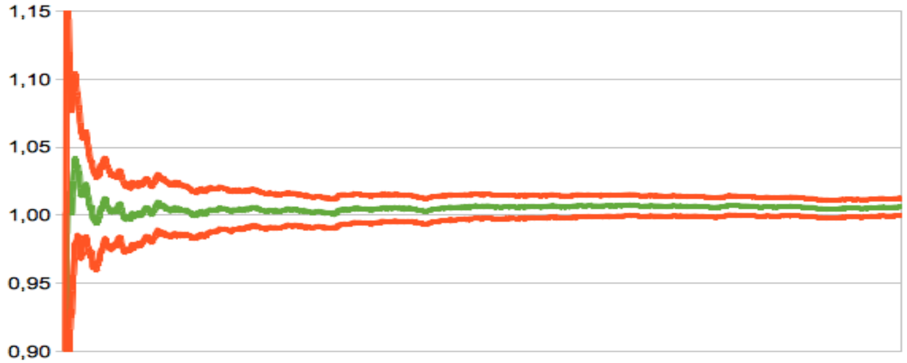
**Lemme (de Slutsky).** Si $r_n \xrightarrow{\mathbb{P}} 1$ et $Y_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$, alors $r_n Y_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$.

**Corollaire.** Avec probabilité 95% (asymptotiquement quand $n \to \infty$), on a

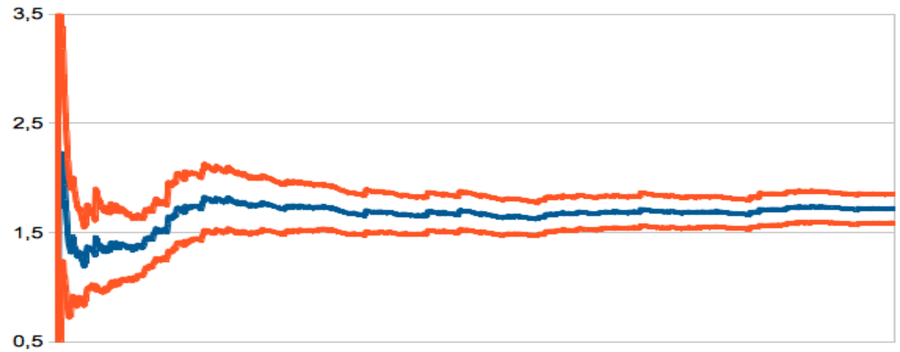
$$\mathbf{\hat{I}_n} - \mathbf{1.96}\frac{\sigma_\mathbf{n}}{\sqrt{\mathbf{n}}} \leq \mathbf{I} \leq \mathbf{\hat{I}_n} + \mathbf{1.96}\frac{\sigma_\mathbf{n}}{\sqrt{\mathbf{n}}}.$$

# De l'importance (bis) de l'intervalle de confiance

Tracé jusqu'à 1000 simulations des moyennes empiriques et intervalles de confiance à 95%, dont la largeur dépend fortement de l'exemple considéré.




Calcul Monte Carlo de $\mathbb{E}(e^{X/10}) = 1.005\ldots$ où $X$ de loi $\mathcal{N}(0,1)$.




Calcul Monte Carlo de $\mathbb{E}(e^{X}) = 1.649\ldots$.

# Etude des échantillons gaussiens

Préparatoire aux applications en statistique.

Modèle : $(Y_1, \ldots, Y_n)$ v.a. i.i.d. de loi $\mathcal{N}(m, \sigma^2)$.

## Transformation affine de gaussiennes indépendantes

**Proposition.** Si $(Y_1, \ldots, Y_n)$ v.a. i.i.d. de loi $\mathcal{N}(m, \sigma^2)$, alors $(\alpha + \beta Y_1, \ldots, \alpha + \beta Y_n)$ définit une suite i.i.d. de loi normale $\mathcal{N}(\alpha + \beta m, \beta^2 \sigma^2)$.

**Exemple.** Si $X_i \overset{\text{loi}}{=} \mathcal{N}(0,1)$ (centré réduit), alors $Y_i := m + \sigma X_i \overset{\text{loi}}{=} \mathcal{N}(m, \sigma^2)$.

**Proposition.** Si $(Y_i)_i$ suite i.i.d. de loi normale $\mathcal{N}(m, \sigma^2)$, alors

$$\sum_{i=1}^n \alpha_i Y_i \overset{\text{loi}}{=} \mathcal{N}\Big(\sum_{i=1}^n \alpha_i m, \sum_{i=1}^n \alpha_i^2 \sigma^2\Big).$$

## Transformation affine plus générale

Soit $\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ avec composantes i.i.d. gaussiennes standards (centrées réduites), on définit $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} := m + K\mathbf{X}$ pour

- une matrice $K$ inversible (de taille $n \times n$),
- un vecteur $m$ (de taille $n$).

**Proposition.** $\mathbf{Y}$ est un vecteur aléatoire de loi gaussienne de densité

$$\mathbf{f(y)} = \frac{\mathbf{1}}{(\mathbf{2}\pi)^{\mathbf{n/2}}\sqrt{\det(\mathbf{KK}^\top)}}\mathbf{e}^{-\frac{1}{2}(\mathbf{y}-\mathbf{m})^\top(\mathbf{KK}^\top)^{-\mathbf{1}}(\mathbf{y}-\mathbf{m})}$$

(voir diapositive 13 du cours 4).

Preuve. Appliquer la formule de changement de variable multidimensionnelle et se ramener à la quantité canonique $KK^\top$. □

## Calculs des deux premiers moments

- Espérance : $\mathbb{E}(Y_i) = \mathbb{E}(m_i + \sum_k K_{ik} X_k) = m_i$
- Covariance :

$$\begin{aligned}
\mathbb{C}\mathrm{ov}(Y_i, Y_j) &= \mathbb{C}\mathrm{ov}(m_i + \sum_k K_{ik} X_k, m_j + \sum_l K_{jl} X_l) \\
&= \sum_{k,l} K_{ik} K_{jl} \mathbb{C}\mathrm{ov}(X_k, X_l) \qquad \text{(bilinéarité)} \\
&= \sum_k K_{ik} K_{jk} \qquad \text{(indépendance des } X_k\text{)} \\
&= (KK^\top)_{ij}.
\end{aligned}$$

**Propriétés.**

1. $m$ est le vecteur d'espérance de $Y$.
2. $KK^\top$ est la matrice de variance-covariance de $Y$.

## Fonctions caractéristiques

**Théorème.** $\Phi_Y(u) = \mathbb{E}(e^{i\langle u,Y\rangle}) = e^{i\langle u,m\rangle - \frac{1}{2}u\cdot KK^\top u}$.

**Corollaire.** Les composantes de $\mathbf{Y} = m + K\mathbf{X}$ sont indépendantes *si et seulement si* $KK^\top$ est diagonale.

Preuve. Composantes indépendantes ssi $\Phi_Y(u) = \prod_{j=1}^d \Phi_{Y_j}(u_j)$, $\forall u$ (voir Leçon 6).

Mais $\Phi_{Y_j}(u_j) = e^{iu_jm_j - \frac{1}{2}(KK^\top)_{jj}u_j^2}$, d'où

$$\prod_{j=1}^d \Phi_{Y_j}(u_j) = \prod_{j=1}^d e^{iu_jm_j - \frac{1}{2}(KK^\top)_{jj}u_j^2}.$$

Egal à $\Phi_Y(u)$ ssi $(KK^\top)_{ij} = 0$ pour $i \neq j$.

$\square$

**Remarque.** Il existe un théorème un peu plus général : "un vecteur aléatoire de loi gaussienne a des composantes indépendantes ssi sa matrice de covariance est diagonale".

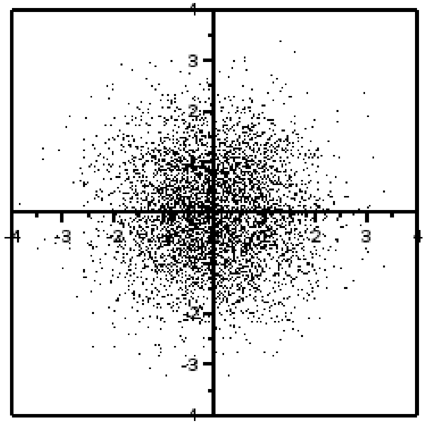
## Application : invariance de la loi gaussienne par rotation

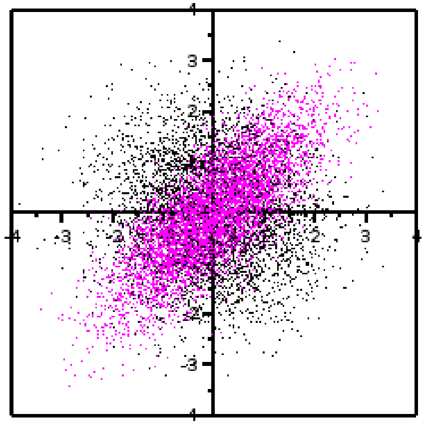
**Proposition.** Si $K$ est une matrice de rotation ($K^{-1} = K^{\top}$) et si $\mathbf{Y} = K\mathbf{X}$ avec $\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ vecteur aléatoire à composantes indépendantes gaussiennes centrées réduites, alors $\mathbf{Y}$ est un vecteur aléatoire avec composantes indépendantes, chacune de loi gaussienne centrée réduite.

**Exemple (démonstration).** $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_2) & \cos(\theta_2) \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

avec deux angles $\theta_1, \theta_2$ de rotation des axes 1 et 2.

➠ $KK^{\top} = \begin{pmatrix} 1 & \sin(\theta_2 - \theta_1) \\ \sin(\theta_2 - \theta_1) & 1 \end{pmatrix}$





## Deux nouvelles lois usuelles

Préparatoire à la statistique sur échantillon gaussien.

**Définition (loi du $\chi^2$).** La loi du chi 2 à $d$ degrés de liberté (notée $\chi^2(d)$) est la loi de la v.a. égale à la somme de $d$ carrés de v.a. i.i.d. de loi $\mathcal{N}(0,1)$.

**Propriétés.**

- $\chi^2(d) \overset{\text{loi}}{=} \Gamma(\frac{d}{2}, \frac{1}{2})$.
- Sa densité est $c_{\chi^2(d)}\ e^{-y/2} y^{d/2-1} \mathbf{1}_{y>0}$.
- Si $Y \overset{\text{loi}}{=} \chi^2(d)$, $\mathbb{E}(Y) = d$ et $\mathbb{V}\text{ar}(Y) = 2d$.

**Définition (loi de Student).** La loi de Student de paramètre $d$ (notée $t(d)$) est la loi de $\frac{G}{\sqrt{Y/d}}$ où $G \overset{\text{loi}}{=} \mathcal{N}(0,1)$, $Y \overset{\text{loi}}{=} \chi^2(d)$, $G$ et $Y$ sont indépendants.

**Propriété.** Sa densité est $\frac{c_{t(d)}}{\left(1+\frac{y^2}{d}\right)^{\frac{d+1}{2}}} \mathbf{1}_{y \in \mathbb{R}}$.

## Théorème de Cochran

**Théorème.** Soit $\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ un vecteur à composantes indépendantes gaussiennes centrées réduites.

Soit $(E_1, \ldots, E_p)$ une décomposition de $\mathbb{R}^n$ en $p$ sous-espaces orthogonaux de dimensions respectives $d_1, \ldots, d_p$ $(d_1 + \cdots + d_p = n)$. Alors

- les projections orthogonales $(Y_{E_j}, 1 \leq j \leq p)$ de $Y$ sur ces sous-espaces sont des vecteurs aléatoires indépendants,
- le carré de la norme $|Y_{E_j}|^2$ a la loi d'un $\chi^2(d_j)$ (càd la loi de la somme de $d_j$ carrés de v.a. normales centrées réduites).

*Application statistique vue plus tard...*

# A propos de la corrélation

**Définition.** La corrélation entre deux v.a.r. $X$ et $Y$ de carré intégrable (et non constantes) est donnée par

$$\mathbb{C}\text{or}(X, Y) := \frac{\mathbb{C}\text{ov}(X, Y)}{\sqrt{\mathbb{V}\text{ar}(X)}\sqrt{\mathbb{V}\text{ar}(Y)}}.$$

**Propriétés.**

- Mesure la dépendance entre $X$ et $Y$.
- $\rho \in [-1, 1]$ par inégalité de Cauchy-Schwarz.
- $\rho = \pm 1$ si et seulement il existe $(a \neq 0, b)$ tq $\mathbb{P}(Y = aX + b) = 1$.
- Si $X$ et $Y$ indépendants, alors $\mathbb{C}\text{or}(X, Y) = 0$.
- Si $(X, Y)$ est un vecteur aléatoire de loi gaussienne :
$$\text{indépendance} \Leftrightarrow \mathbb{C}\text{or}(X, Y) = 0.$$
- Si $\mathbb{C}\text{or}(X, Y) = 0$ ➠ ? ? en général pas de conclusion sur l'indépendance.

## Exemple : $\mathbb{C}\mathrm{or}(X, Y) = 0$ et $(X, Y)$ dépendants

$X$ de loi $\mathcal{N}(0, 1)$, $Y = \varepsilon X$ avec

- $\mathbb{P}(\varepsilon = \pm 1) = \frac{1}{2}$ ;
- $\varepsilon$ et $X$ indépendants.
  - ▷ $\mathbb{E}(\mathbf{X}) = \mathbb{E}(\mathbf{Y}) = \mathbf{0}$.
  - ▷ $\mathbb{C}\mathrm{ov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E}(\varepsilon X^2) = \mathbb{E}(\varepsilon)\mathbb{E}(X^2) = \mathbf{0}$.
  - ▷ $\varepsilon X$ **est de loi** $\mathcal{N}(0, 1)$ car

$$\mathbb{E}(e^{iu\varepsilon X}) = \mathbb{E}(e^{-iuX}\mathbf{1}_{\varepsilon=-1}) + \mathbb{E}(e^{iuX}\mathbf{1}_{\varepsilon=1}) = \frac{1}{2}e^{-\frac{(-u)^2}{2}} + \frac{1}{2}e^{-\frac{u^2}{2}} = \Phi_X(u).$$

  - ▷ $X$ **et** $Y$ **non indépendantes :**

$$\mathbb{P}(Y > 1, |X| \leq 1) = 0 \neq \mathbb{P}(Y > 1)\mathbb{P}(|X| \leq 1).$$

➠ Corrélation $=$ *une* mesure de la dépendance qui ne décrit pas toute la dépendance.

# Pour en savoir plus : bornes sur la corrélation

**Théorème (de Lancaster).** Soit $(X, Y)$ un vecteur aléatoire de loi gaussienne, avec composante centrée réduite et corrélation $\rho \in [-1, 1]$. Alors pour toutes fonctions $(f, g)$ telles que $\mathbb{E}(f^2(X) < +\infty)$ et $\mathbb{E}(g^2(Y) < +\infty)$, on a

$$|\mathbb{C}\mathrm{or}(f(X), g(Y))| \leq \rho.$$

Attention à la signification de la corrélation pour mesurer la dépendance entre données.

La corrélation extrêmale peut être différente de $\pm\rho$ (voir après).

Preuve. Transformée de Laplace de $G \overset{\text{loi}}{=} \mathcal{N}(m, \sigma^2)$ : $\mathbb{E}(e^{uG}) = e^{um + \frac{1}{2}u^2\sigma^2}$.

1. $\mathbb{E}(e^{uX - \frac{1}{2}u^2}) = 1$.
2. Polynômes d'Hermite : $P_k(x) = \partial_u^k [e^{ux - \frac{1}{2}u^2}]|_{u=0}$.
   **Exemple.** $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = x^2 - 1$, ...
3. $\mathbb{E}(P_k(X)) = 0$ pour $k \geq 1$.

4. $\mathbb{E}(P_k(X)P_l(Y)) = \begin{cases} 0 & \text{si } k \neq l \\ \rho^k k! & \text{si } k = l \end{cases}$ ➠ $(P_k)_k$ est une famille orthogonale pour la loi gaussienne : $\langle P_k, P_l \rangle = \mathbb{E}(P_k(X)P_l(X)) = \delta_{k,l}$
5. On suppose d'abord $f(x) = \sum_{k=1}^n a_k P_k(x)$ et $g(y) = \sum_{k=1}^n b_k P_k(y)$. Alors
   - $\mathbb{E}(f(X)) = \mathbb{E}(g(Y)) = 0$,
   - $$|\mathbb{C}\text{ov}(f(X), g(Y))| = |\mathbb{E}(f(X)g(Y))| = |\sum_{k=1}^n a_k b_k \rho^k k!|$$
     $$\leq |\rho| \sum_{k=1}^n |a_k||b_k|k!$$
     $$\leq |\rho| \sqrt{\sum_{k=1}^n a_k^2 k!} \sqrt{\sum_{k=1}^n b_k^2 k!} = |\rho| \sqrt{\mathbb{V}\text{ar}(f(X))} \sqrt{\mathbb{V}\text{ar}(g(Y))}.$$
6. Le cas général :
   - On peut toujours recentrer $f(X)$ et $g(Y)$ avec $a_0 = b_0 = 0$.
   - Toute v.a. $f(X)$ de carré intégrable (centrée) peut se décomposer sur la base des polynômes d'Hermite : $f(X) = \sum_{k=1}^{+\infty} a_k P_k(X)$.

$\square$

# Les corrélations minimale et maximale

## Exemple (variables lognormales corrélées).

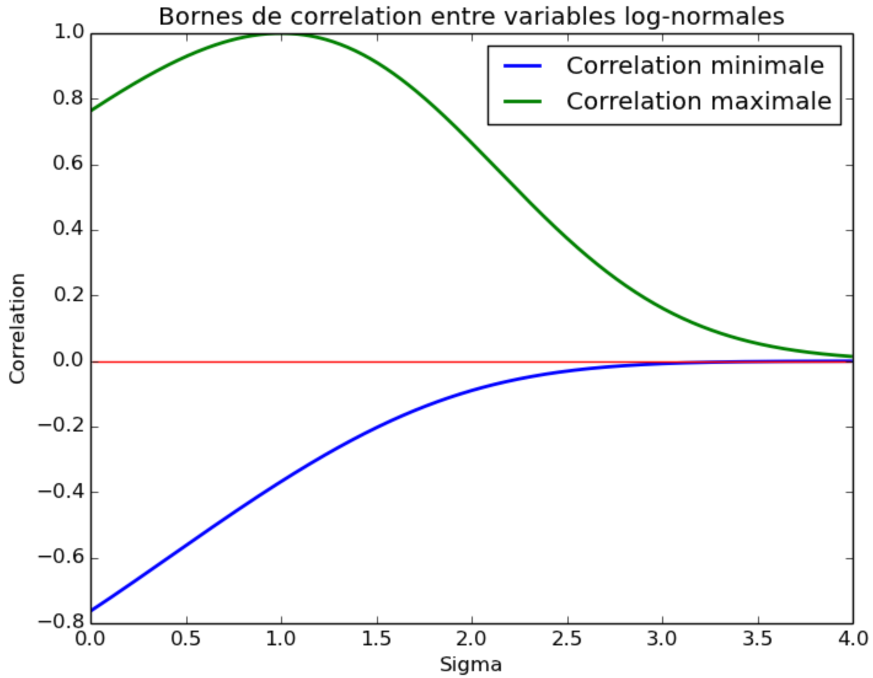
$$F = e^X, \qquad G_\rho = e^{\sigma Y}, \qquad G^+ = e^{\sigma X}, \qquad G^- = e^{-\sigma X}.$$

avec $\mathbb{C}\mathrm{or}(X, Y) = \rho$.

Par le calcul gaussien direct sur les transformées de Laplace, on peut montrer

$$\mathbb{C}\mathrm{or}(F, G^-) \leq \mathbb{C}\mathrm{or}(F, G_\rho) \leq \mathbb{C}\mathrm{or}(F, G^+).$$

Même avec des v.a. très dépendantes, la corrélation peut être très loin de $\pm 1$.



# Statistique sur échantillon gaussien, estimation de moyenne

Modèle : $(Y_1, \ldots, Y_n)$ v.a. i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$.

## 1er cas : moyenne $\mu$ inconnue, variance $\sigma^2$ connue

Moyenne empirique :

$$\bar{Y}_n := \frac{1}{n} \sum_{i=1}^{n} Y_i \overset{\text{loi}}{=} \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

**Théorème.** $\bar{Y}_n$ est un estimateur convergeant de $\mu$.

Avec probabilité 95%, on a (pour chaque $n$)

$$\bar{Y}_n - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + 1.96 \frac{\sigma}{\sqrt{n}}.$$

Intervalles de confiance non asymptotiques ($\neq$ cadre Monte-Carlo)

## 2ème cas : moyenne $\mu$ et variance $\sigma^2$ inconnues

**Proposition.** $\bar{Y}_n \overset{p.s.}{\to} \mu$ quand $n \to +\infty$.

On définit $\sigma_n^2 := \frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ variance empirique.

**Proposition.** $\sigma_n^2$ est un estimateur convergeant et sans biais de $\sigma^2$.

Preuve. On veut démontrer

- $\sigma_n^2 \overset{p.s.}{\to} \sigma^2$,
- $\mathbb{E}(\sigma_n^2) = \sigma^2$.

On a $\sigma_n^2 = \dfrac{n}{n-1}\Big(\dfrac{1}{n}\displaystyle\sum_{i=1}^n Y_i^2 - \bar{Y}_n^2\Big)$.

- Par la LGN, $\sigma_n^2 \overset{p.s.}{\to} \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \sigma^2$.
-

$$
\begin{aligned}
\mathbb{E}(\sigma_n^2) &= \frac{n}{n-1}\mathbb{E}(Y^2) - \frac{n}{n-1}\Big[\frac{\mathbb{V}\mathrm{ar}(Y)}{n} + (\mathbb{E}(Y))^2\Big] \\
&= \frac{n}{n-1}\mathbb{V}\mathrm{ar}(Y) - \frac{n}{n-1}\frac{\mathbb{V}\mathrm{ar}(Y)}{n} = \mathbb{V}\mathrm{ar}(Y).
\end{aligned}
$$

□

## Vitesse de convergence

**Théorème.** $\sqrt{n}\frac{(\bar{Y}_n-\mu)}{\sigma}$ et $(n-1)\frac{\sigma_n^2}{\sigma^2}$ sont indépendants, de loi $\mathcal{N}(0,1)$ et de loi $\chi^2(n-1)$.

**Corollaire.** $\sqrt{n}\frac{(\bar{Y}_n-\mu)}{\sigma_n} \overset{\text{loi}}{=} \frac{\mathcal{N}(0,1)}{\sqrt{\frac{\chi^2(n-1)}{n-1}}}$ avec numérateur et dénominateur indépendants, c.à-d. une v.a. de loi de Student avec paramètre $n-1$.

**Propriété (intervalles de confiance).** Avec probabilité 95%, on a (pour chaque $n$)

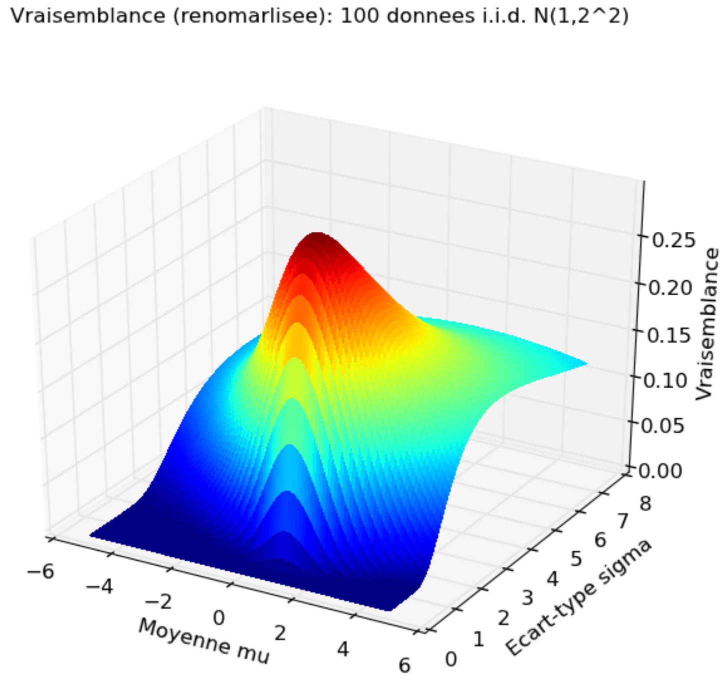
$$\bar{Y}_n - t_{n-1}(97.5\%)\frac{\sigma_n}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + t_{n-1}(97.5\%)\frac{\sigma_n}{\sqrt{n}}.$$

| $n$ | 10 | 20 | 100 | $+\infty$ |
|---|---|---|---|---|
| $t_{n-1}(97.5\%)$ | 2.262 | 2.093 | 1.984 | 1.96 |

# Estimateur du maximum de vraisemblance (e.m.v.)

Etant donnés des paramètres $(\mu, \sigma^2)$, le vecteur aléatoire $(Y_1, \ldots, Y_n)$ i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$ a une loi de densité

$$p_{\mu,\sigma^2}(y_1, \ldots, y_n) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_j-\mu)^2}{2\sigma^2}}.$$




**Définition.** L'e.m.v. est défini par

$$(\hat{\mu}_{\mathbf{n}}, \hat{\sigma}^{\mathbf{2}}_{\mathbf{n}}) = \arg\max_{\mu,\sigma^{\mathbf{2}}} \mathbf{p}_{\mu,\sigma^{\mathbf{2}}}(\mathbf{Y_1}, \ldots, \mathbf{Y_n})$$

(paramètres qui rendent l'observation la plus "probable").

**Proposition.** Dans ce cas gaussien, on trouve $\hat{\mu}_n = \bar{X}_n$ et $\hat{\sigma}^2_n = \frac{n-1}{n}\sigma^2_n$.

**Remarque.** En général, l'e.m.v. a de très bonnes propriétés de convergence et d'optimalité (voir cours de Statistique de 2A).

# Coupe du monde: "Le tirage au sort est injuste"

Propos recueillis par Olivier Monod, publié le 03/06/2014 à 14:20, mis à jour le 04/06/2014 à 12:34

FOOTBALL - Comment améliorer l'équité dans le plus grand événement sportif du monde? Julien Guyon est mathématicien et amateur de football. Il propose une nouvelle méthode de composition des groupes pour la phase de poules du Mondial de football.




Un mathématicien a trouvé une solution pour rendre le tirage au sort de la coupe du monde plus équitable.

AFP PHOTO / NELSON ALMEIDA