

2.4 Théorèmes fondamentaux de probabilité

Approche expérimentale

- Considérons l'expérience de jeter une pièce de monnaie 10'000 fois et observons le nombre de "face" obtenues
- Soient X_1, \dots, X_n les variables aléatoires indépendantes

$$X_i = \begin{cases} 1, & \text{si le } i\text{ème jet donne "face",} \\ 0, & \text{si le } i\text{ème jet donne "pile"} \end{cases} \quad \sim B(1, p)$$

- Donc $S_n = X_1 + \dots + X_n$ représente le nombre de "face" sur n essais et

$$S_n \sim \mathcal{B}(n, p)$$

- La proportion de "face" sur n jets est $\bar{X}_n := S_n/n$ et

$$\begin{aligned} \mathbb{E}(\bar{X}_n) &= n^{-1} \mathbb{E}(S_n) = n^{-1} np = p, \\ \text{var}(\bar{X}_n) &= n^{-2} \text{var}(S_n) = n^{-2} np(1-p) = p(1-p)/n \rightarrow 0 \end{aligned}$$

quand $n \rightarrow \infty$

- Donc \bar{X}_n se concentre de plus en plus autour de p

Théorème (loi (faible) des grands nombres) Soient X_1, X_2, \dots des variables aléatoires indépendantes et identiquement distribuées d'espérance $\mu = \mathbb{E}(X_1)$ et variance $\sigma^2 = \text{var}(X_1)$ finies. Alors pour tout $\epsilon > 0$

$$\Pr(|\bar{X}_n - \mu| \geq \epsilon) \rightarrow 0, \quad n \rightarrow \infty. \quad (1)$$

***Théorème (loi forte des grands nombres)** Soient X_1, X_2, \dots des variables aléatoires indépendantes et identiquement distribuées d'espérance $\mu = \mathbb{E}(X_1)$ finie. Alors

$$\Pr\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1 \quad (2)$$

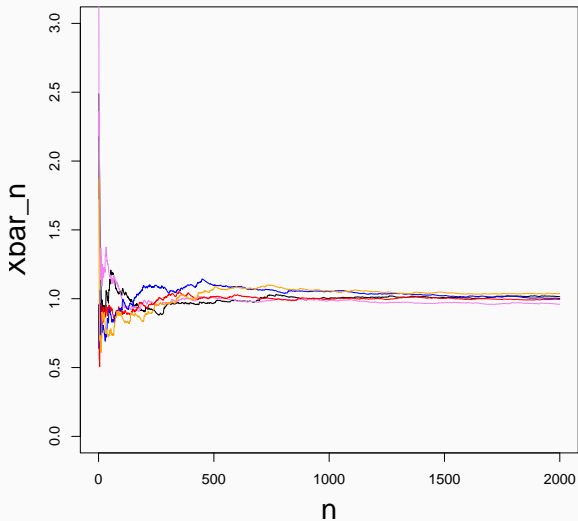
*Il est donc certain que \bar{X}_n soit proche de μ pour n grand

- *La loi forte est plus forte parce que (2) implique (1) et la variance peut être infinie
- *La loi faible utilise seulement $\text{cov}(X_i, X_j) = 0$ pour $i \neq j$

Pour l'examen il suffit de connaître la loi faible

Illustration de la loi des grands nombres : $\exp(1)$

5 replications de \bar{X}_n



Vitesse de convergence : Théorème central limite

- $\bar{X}_n \rightarrow \mu$ quand $n \rightarrow \infty$, mais à quelle vitesse ?
- Comme $\mathbb{E}(\bar{X}_n) = \mu$ et $\text{var}(\bar{X}_n) = \sigma^2/n \in (0, \infty)$, pour tout n

$$Z_n := \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}$$

a espérance 0 et variance 1, suggérant que la vitesse est \sqrt{n}

Théorème central limite Soient X_1, X_2, \dots des variables aléatoires indépendantes et identiquement distribuées d'espérance μ et variance $\sigma^2 \in (0, \infty)$. Alors $Z_n := \sqrt{n}(\bar{X}_n - \mu)/\sigma$ satisfait

$$\Pr(Z_n \leq x) \rightarrow \Phi(x), \quad x \in \mathbb{R}$$

La convergence étant uniforme en x , on déduit

$$\Pr(\bar{X}_n \leq x) = \Pr(Z_n \leq \sqrt{n}(x - \mu)/\sigma) \approx \Phi(\sqrt{n}(x - \mu)/\sigma)$$

donc \bar{X}_n suit **approximativement** une loi $\mathcal{N}(\mu, \sigma^2/n)$

Illustration avec des variables $\exp(1)$

On calcule $\sqrt{n}(\bar{X}_n - \mathbb{E}(X_1))/\sqrt{\text{var}(X_1)}$, $R = 5000$ fois

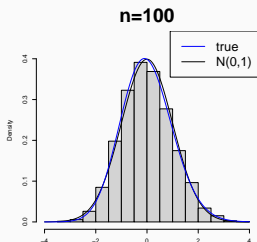
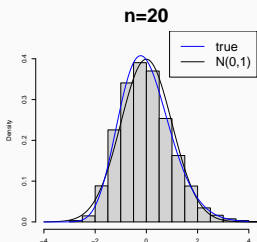
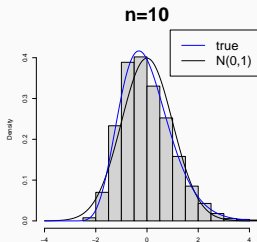
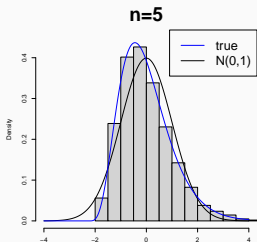


Illustration avec des variables $\exp(1)$

- On s'intéresse à la distribution de

$$\frac{\sqrt{n}(\bar{X}_n - \mathbb{E}(X_1))}{\sqrt{\text{var}(X_1)}}$$

- Fixons $n = 5$ ou 10 ou 20 ou 100 et $R = 5000$
- Générer $z_1^{(1)}, \dots, z_n^{(1)} \stackrel{iid}{\sim} \exp(1)$, et calculer leur moyenne $\bar{z}^{(1)}$
- Générer $z_1^{(2)}, \dots, z_n^{(2)} \stackrel{iid}{\sim} \exp(1)$, et calculer leur moyenne $\bar{z}^{(2)}$
- \vdots
- Générer $z_1^{(R)}, \dots, z_n^{(R)} \stackrel{iid}{\sim} \exp(1)$, et calculer leur moyenne $\bar{z}^{(R)}$
- Les R valeurs

$$\left(\frac{\sqrt{n}(\bar{z}^{(1)} - \mathbb{E}(X_1))}{\sqrt{\text{var}(X_1)}}, \dots, \frac{\sqrt{n}(\bar{z}^{(R)} - \mathbb{E}(X_1))}{\sqrt{\text{var}(X_1)}} \right)$$

sont un échantillon issu de la distribution d'intérêt

Exemple

Exemple Soit $X \sim \mathcal{B}(m, p)$. Donner une approximation de $\Pr(X \leq r)$, pour $r \in \mathbb{R}$.

Solution Exemple 153 :

On a $X = \sum_{i=1}^m Y_i$, où $Y_1, \dots, Y_m \stackrel{\text{iid}}{\sim} \mathcal{B}(p)$. De plus, $\mathbb{E}(Y_1) = p$ et $\text{Var}(Y_1) = p(1-p)$. Le TCL nous donne donc que $X \stackrel{\text{app}}{\sim} \mathcal{N}(mp, mp(1-p))$ pour m grand. Ainsi, si Z désigne une variable aléatoire de loi $\mathcal{N}(0, 1)$, on a, pour m grand,

$$\begin{aligned}\Pr(X \leq r) &= \Pr\left(\frac{X - mp}{\sqrt{mp(1-p)}} \leq \frac{r - mp}{\sqrt{mp(1-p)}}\right) \\ &\approx \Pr\left(Z \leq \frac{r - mp}{\sqrt{mp(1-p)}}\right) = \Phi\left(\frac{r - mp}{\sqrt{mp(1-p)}}\right).\end{aligned}$$

Utilisation du théorème central limite

- Le théorème central limite est utilisé pour approximer des probabilités impliquant des sommes de variables aléatoires indépendantes
- Sous les conditions précédentes, on a

$$\mathbb{E} \left(\sum_{j=1}^n X_j \right) = n\mu, \quad \text{var} \left(\sum_{j=1}^n X_j \right) = n\sigma^2 \in (0, \infty)$$

- On **standardise** la somme

$$\frac{\sum_{j=1}^n X_j - n\mu}{\sqrt{n\sigma^2}} = \frac{n(\bar{X}_n - \mu)}{\sqrt{n\sigma^2}} = \frac{n^{1/2}(\bar{X}_n - \mu)}{\sigma} = Z_n$$

- Par le théorème central limite Z_n est approximativement $\mathcal{N}(0, 1)$ et donc

$$\Pr \left(\sum_{j=1}^n X_j \leq r \right) = \Pr \left\{ \frac{\sum_{j=1}^n X_j - n\mu}{\sqrt{n\sigma^2}} \leq \frac{r - n\mu}{(n\sigma^2)^{1/2}} \right\} \approx \Phi \left\{ \frac{r - n\mu}{(n\sigma^2)^{1/2}} \right\}.$$

Exemple Un livre de 640 pages a un nombre aléatoire d'erreurs sur chaque page. Si le nombre d'erreurs par page suit une loi de Poisson d'espérance $\lambda = 0.1$, et est indépendant des autres pages, quelle est la probabilité que le livre contienne moins de 50 erreurs ?

Exemple : théorème central limite

Exemple Un livre de 640 pages a un nombre d'erreurs aléatoires à chaque page. Si le nombre d'erreurs par page suit une loi de Poisson d'espérance $\lambda = 0.1$, et est indépendant des autres pages, quelle est la probabilité que le livre contienne moins de 50 erreurs ?

Extensions et remarques

- Le théorème central limite est **remarquable**, car la distribution des X_i **n'a pas d'importance** : seulement l'espérance et la variance apparaissent. $\sum X_i$ a approximativement la même distribution si $X_i \sim \text{Exp}(1)$ ou $X_i \sim \text{Poiss}(1)$
- Méthode delta** Si g est une fonction telle que $g'(\mu)$ existe, alors

$$\sqrt{n} \frac{g(\bar{X}_n) - g(\mu)}{\sigma} = g'(\mu)Z_n + o(Z_n)$$

suit approximativement $\mathcal{N}(0, [g'(\mu)]^2)$ et donc $g(\bar{X}_n) \overset{\text{app}}{\approx} \mathcal{N}\{g(\mu), g'(\mu)^2 \sigma^2 / n\}$

- Version **générale** de la méthode delta : si $\sqrt{n}(Y_n - \theta) \overset{\text{app}}{\approx} Y$ pour une constante $\theta \in \mathbb{R}$, alors $\sqrt{n}(g(Y_n) - g(\theta)) \overset{\text{app}}{\approx} g'(\mu)Y$
- Notation** : $\overset{\text{app}}{\approx}$ indique une distribution **approximative**
- Le théorème central limite dépend d'un effet de moyennement, et échoue quand tout dépend d'une fraction minuscule des variables. Il n'est donc pas valable pour les maxima, les minima, l'étendue, \dots , pour lesquels on a d'autres théorèmes limites