Théorème central limite : (énoncé)

Soit $(X_1, \dots, X_n)$ un échantillon i.i.d (indépendant et identiquement distribué) qui admet une espérance $E[X_1]$ et une variance $Var(X_1)$.

$$\boxed{\text{Alors} \quad \frac{\sqrt{n}\left(\overline{X}_n - E[X_1]\right)}{\sqrt{Var(X_1)}} \xrightarrow{\text{loi}} \mathcal{N}(0, 1)}$$

$\overline{X}_n - E[X_1]$ → on centre $\overline{X}_n$

$\dfrac{\overline{X}_n - E[X_1]}{\sqrt{Var(\overline{X}_n)}}$ → on réduit

En particulier, si $Z_n = \dfrac{\sqrt{n}\,(\bar{X}_n - E[X_1])}{\sqrt{var(X_1)}}$

- $F_{Z_n}(t) \xrightarrow[n\to+\infty]{(\approx)} F_{\mathcal{N}(0,1)}(t)$ pour tout $t \in \mathbb{R}$.

- $\mathbb{P}(Z_n \in [a,b]) \xrightarrow[n\to+\infty]{(\approx)} \underbrace{\mathbb{P}(\mathcal{N}(0,1) \in [a,b])}_{\text{calculable}}$

Exercice 1 :

1) ok

2) p représente la fraction de la population qui va effectivement être impacté par la campagne.

3) $$E[\overline{X}_n] = E\left[\frac{X_1 + \cdots + X_n}{n}\right]$$
$$= \frac{1}{n}\left(E[X_1] + \cdots + E[X_n]\right)$$
$$= \frac{1}{n} \times np = p.$$

On retrouve que $\overline{X}_n$ est un estimateur sans biais de $E[X_1]$ (donc de p)

$$\mathrm{Var}(\overline{X}_n) = \mathrm{Var}\left(\frac{X_1 + \dots + X_n}{n}\right)$$

$$= \frac{1}{n^2}\left(\mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_n)\right)$$ car les variables sont indépendantes

$$= \frac{1}{n^2}\, n \times p(1-p)$$

$$\underline{\mathrm{Var}(\overline{X}_n) = \frac{p(1-p)}{n}}$$

(Rq : $\left.\begin{array}{l} * \text{ sans biais} \\ * \ \mathrm{Var}(\,) \xrightarrow[n \to +\infty]{} 0 \end{array}\right\}$ estimateur convergent)

4) $\mathbb{P}(|\overline{X_n} - p| \geq \varepsilon)$?

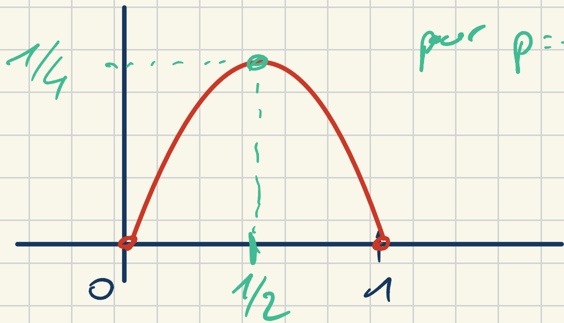
inégalité de B-T : $P(|X - E[X]| \geq \varepsilon) \leq \dfrac{\sigma^2}{\varepsilon^2}$

D'après l'inégalité de B-T :

$$\mathbb{P}(|\overline{X_n} - \underbrace{E[\overline{X_n}]}_{p}| \geq \varepsilon) \leq \frac{Var(\overline{X_n})}{\varepsilon^2}$$

Donc $\mathbb{P}(|\overline{X_n} - p| \geq \varepsilon) \leq \dfrac{p(1-p)}{n\varepsilon^2} \leq \dfrac{1}{4n\varepsilon^2}$

or $p(1-p) \leq 1$ et même $p(1-p) \leq 1/4$



pour $p = \frac{1}{2}$: $\frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}$

5) On cherche à avoir :

$$P\left(|\bar{X}_n - p| < 0{,}05\right) \geq 0{,}95$$

on sait que $P\left(|\bar{X}_n - p| \geq 0{,}05\right) \leq \dfrac{1}{4n\, 0{,}05^2}$

donc

$$1 - P\left(|\bar{x}_n - p| < 0{,}05\right) \leq \frac{1}{4n\, 0{,}05^2}$$

$$1 - \frac{1}{4n\,0{,}05^2} \leq \mathbb{P}\left(|\bar{X}_n - p| < 0{,}05\right)$$

$= 0{,}95$

$\geq 0{,}95$

or $$1 - \frac{1}{4n\,0{,}05^2} \geq 0{,}95$$

$$\iff 0{,}05 \geq \frac{1}{4n\,0{,}05^2}$$

$$\iff n \geq \frac{1}{4 \times 0{,}05^3} = 2000$$

Donc il faut interroger au moins 2000 personnes

6) On cherche toujours à avoir

$$\mathbb{P}\left(|\bar{X}_n - p| < 0{,}05 \right) \geq 0{,}95 .$$

i.e $$\mathbb{P}\left(-0{,}05 \leq \bar{X}_n - p \leq 0{,}05 \right) \geq 0{,}95$$

D'après le TCL :

$$\frac{\sqrt{n}\,(\bar{X}_n - p)}{\sqrt{p(1-p)}} \approx \mathcal{N}(0,1)$$

Donc $$\mathbb{P}\left(a \leq \frac{\sqrt{n}\,(\bar{X}_n - p)}{\sqrt{p(1-p)}} \leq b \right) \approx \mathbb{P}\left(a \leq \mathcal{N}(0,1) \leq b \right)$$

Or, $\mathbb{P}\left(-0{,}05 \le \bar{X}_n - p \le 0{,}05\right)$

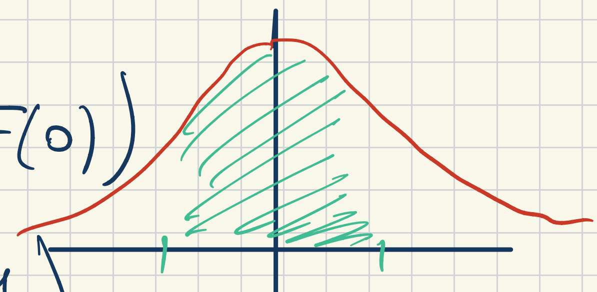
$$= \mathbb{P}\left(-0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}} \le \frac{\sqrt{n}\,(\bar{X}_n - p)}{\sqrt{p(1-p)}} \le 0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)$$

$$\approx \mathbb{P}\left(-0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}} \le \mathcal{N}(0,1) \le 0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right)$$

$$= 2 \times \int_0^{0{,}05\frac{\sqrt{n}}{\sqrt{p(1-p)}}} f(x)\,dx$$

$$= 2 \times \left(F\left(0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - F(0)\right)$$

$$= 2 \times \left(F\left(0{,}05\,\frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - \frac{1}{2}\right)$$



on veut que

$$2 \times \left(F\left(0{,}05 \frac{\sqrt{n}}{\sqrt{p(1-p)}}\right) - \frac{1}{2} \right) \geqslant 0{,}95$$

or, $\sqrt{p(1-p)} \leq \sqrt{\frac{1}{4}} = \frac{1}{2}$

$$\mathbb{P}\left(-0{,}05 \frac{\sqrt{n}}{\sqrt{p(1-p)}} \leq \mathcal{N}(0,1) \leq 0{,}05 \frac{\sqrt{n}}{\sqrt{p(1-p)}} \right)$$

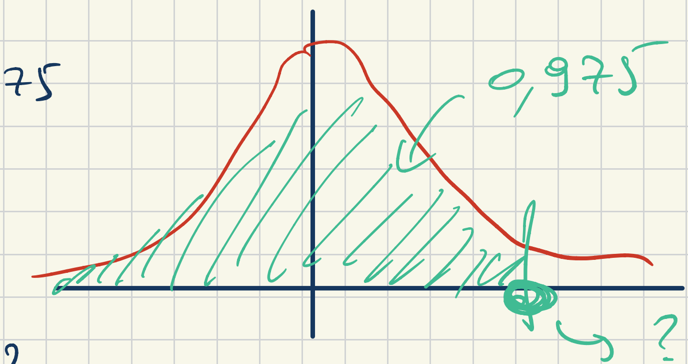
$$\geqslant \mathbb{P}\left(-0{,}05 \frac{\sqrt{n}}{1/2} \leq \mathcal{N}(0,1) \leq 0{,}05 \frac{\sqrt{n}}{1/2} \right)$$

on veut donc que $2\left(F\left(\frac{0,05\sqrt{n}}{1/2}\right)-\frac{1}{2}\right) \geq 0,95$

$$2F\left(\frac{0,05\sqrt{n}}{1/2}\right) \geq 1,95$$

donc $F\left(\frac{0,05\sqrt{n}}{1/2}\right) \geq 0,975$

Donc $\frac{0,05\sqrt{n}}{1/2} = q_{0,975}$



et $n = \left(\frac{q_{0,975}}{2\times 0,05}\right)^2$

$$n = \left(\frac{1,96}{0,1}\right)^2 \gtrsim 385$$

Exercice 2

On pose :

$$X_i = \begin{cases} 1 & \text{si l'abonné } i \text{ est connecté} \\ 0 & \text{sinon} \end{cases}$$

indépendant ("les comportements des abonnés sont indépendants")

on dispose d'un échantillon $X_1, \dots, X_{5000}$ de loi $\mathcal{B}(p = 0{,}2)$

1) $X = \sum_{i=1}^{5000} X_i$ donc $X \sim \mathcal{B}(5000, 0.2)$

(somme de 5000 Bernoulli indépendantes).

Donc $E[X] = 5000 \times 0.2 = 1000$

et $\sigma(X) = \sqrt{5000 \times 0.2 \times 0.8} = \sqrt{800}$

c) On pose $Y = \dfrac{X - 1000}{\sqrt{800}}$

"approché par une loi $\mathcal{N}(0,1)$" $\rightarrow$ TCL.

On applique le TCL à $X_1, \dots, X_{5000}$.

$$\frac{\sqrt{5000}\left(\overline{X}_{5000} - \overset{p=0.2}{E[X_1]}\right)}{\sqrt{Var(X_1)}} \approx \mathcal{N}(0,1),$$

$= \sqrt{0.2 \times 0.8}$

$= \sqrt{0,16}$

Donc $\dfrac{\sqrt{5000}\left(\bar{X}_{5000} - 0.2\right)}{\sqrt{0.16}} = \dfrac{\sqrt{5000}\left(\frac{1}{5000}\sum_{i=1}^{5000} X_i - 0.2\right)}{\sqrt{0.16}}$

$$= \frac{\frac{1}{\sqrt{5000}} X - \sqrt{5000} \times 0.2}{\sqrt{0.16}}$$

$$= \frac{X - 5000 \times 0.2}{\sqrt{5000} \times \sqrt{0.16}} = \frac{X - 1000}{\sqrt{800}}$$

3) Posons Δ le nombre de connexions simultanées que le fournisseur peut gérer.

X représente le nombre de connexions simultanées réelles.

être saturé : $X > s$

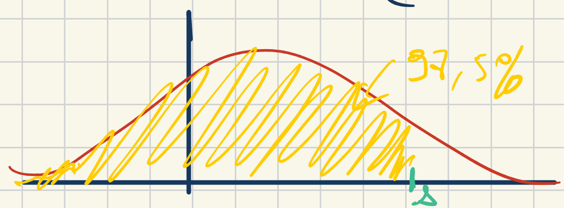
On veut que $\mathbb{P}(X > s) < 2,5\%$.

on cherche s tel que

cela revient à $1 - \mathbb{P}(X \leq s) < 2,5\%$

ou encore $\mathbb{P}(X \leq \underbrace{s}_{}) \geq 97,5\%$

quantile d'ordre 97,5%.



or $\mathbb{P}(X \leq \delta)$

$$= \mathbb{P}\left(\frac{X - 1000}{\sqrt{800}} \leq \frac{\delta - 1000}{\sqrt{800}}\right)$$

$$= \mathbb{P}\left(Y \leq \boxed{\frac{\delta - 1000}{\sqrt{800}}}\right)$$

quantile d'ordre 0,975
de la loi $\mathcal{N}(0,1)$

or $q_{0,975} = 1,96$

donc $\dfrac{\delta - 1000}{\sqrt{800}} = 1.96$

et donc $s = 1.96 \times \sqrt{800} + 1000$

$\simeq 1056$

Donc il suffit que le fournisseur puisse gérer 1056 connexions simultanément pour avoir 97,5% de chance de ne pas avoir de problème.

Exercice 3 :

1) $X_i \sim \mathcal{E}(\lambda)$

Donc $E[X_i] = \frac{1}{\lambda}$ et $Var(X_i) = \underbrace{\frac{1}{\lambda^2}}_{= \theta^2}$

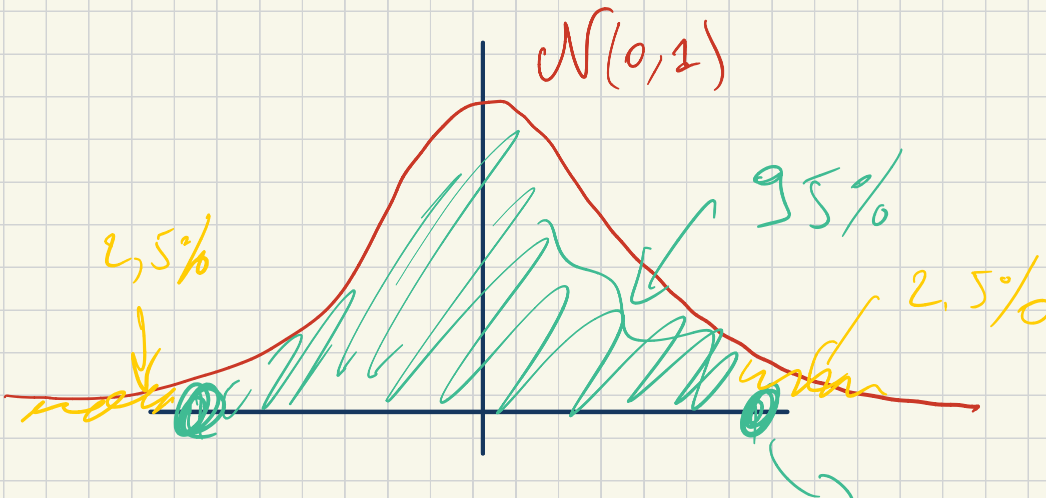
On note $\theta = E[X_i] = \frac{1}{\lambda}$.

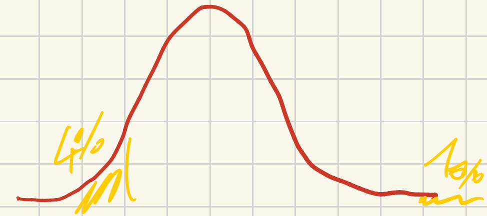
2) D'après le thm centrale limite :

$$\frac{\sqrt{n}(\overline{X}_n - \theta)}{\theta} \approx \mathcal{N}(0, 1)$$

3) On cherche a, b tels que

$$\mathbb{P}(a \leq \theta \leq b) \geq 0,95$$





$$q_{0,975} = 1,96$$

Donc on sait que

$$\mathbb{P}\left(-1.96 \leq \frac{\sqrt{n}(\overline{X}_n - \theta)}{\theta} \leq 1,96\right) = 0,95$$

$\approx \mathcal{N}(0,1)$

$\theta > 0$ donc $\mathbb{P}\left(-1.96\theta \leq \sqrt{n}(\overline{X}_n - \theta) \leq 1.96\theta\right) = 0,95$

$$\mathbb{P}\left(-\frac{1,96\theta}{\sqrt{n}} - \overline{X}_n \leq -\theta \leq \frac{1,96\theta}{\sqrt{n}} - \overline{X}_n\right)$$

$$= \mathbb{P}\left(\underbrace{\overline{X}_n - \frac{1,96\theta}{\sqrt{n}}}_{a} \leq \theta \leq \underbrace{\overline{X}_n + \frac{1.96\theta}{\sqrt{n}}}_{b}\right) = 0,95$$

a et b dépendent de θ

4) Dans les bornes de l'intervalle, on remplace θ par une estimation, c'est-à-dire $\overline{X_n}$.

Donc $\theta \in \left[\overline{X_n} - \dfrac{1{,}96\,\overline{X_n}}{\sqrt{n}} \,,\, \overline{X_n} + \dfrac{1{,}96\,\overline{X_n}}{\sqrt{n}} \right]$

avec proba 95%

$$\theta \in \left[8 - \frac{1.96 \times 8}{10} \, , \, 8 + \frac{1.96 \times 8}{10} \right]$$

$$\approx \left[6{,}4 \; ; \; 9.6 \right]$$

$$X_i = \begin{cases} 1 & \text{si l'étudiant } i \text{ vient} \\ 0 & \text{sinon} \end{cases} \rightsquigarrow \mathcal{B}\left(\underbrace{\tfrac{4}{5}}_{0.8}\right)$$

On a un échantillon $X_1, \ldots, X_N$ et on cherche N (le nombre d'étudiants pouvant être inscrits au cours) tq

$$\mathbb{P}\left(\sum_{i=1}^{N} X_i \leq 500\right) \geq 0{,}99$$

$$\frac{\sqrt{n}\left(\overline{X}_n - E[X_1]\right)}{\sqrt{Var(X_1)}} \simeq \mathcal{N}(0,1)$$

$$X_1 \sim \mathcal{B}(0.8)$$

$$E[X_1] = p = 0.8$$

$$Var(X_1) = p(1-p) = 0.8 \times 0.2$$