De la loi de Bernoulli à la loi normale en suivant le programme de Statistique de Terminale

IREM Marseille / Groupe "Statistique et Probabilités"

Février 2013

Loi de Bernoulli

$\hookrightarrow$ C'est la variable de comptage la plus simple.

X variable aléatoire à valeurs dans $\{0, 1\}$ telle que

$$\begin{aligned} p &= \mathbb{P}(X = 1), \\ 1 - p &= \mathbb{P}(X = 0). \end{aligned}$$

Une autre écriture

- $\mathbb{P}(X = x) = p^x(1-p)^{1-x}$ avec $x \in \{0, 1\}$,
- $\mathbb{P}(X = x) = p^x(1-p)^{1-x}\mu(x)$,

  avec $\mu = \delta_0 + \delta_1$.

Propriétés :

- $\mathbb{E}(X) = p$, $\mathbb{V}(X) = p(1-p)$ (minimum en 1/2),

  $\hookrightarrow$ Lorsque $p = 0$ ou $p = 1$, variable constante (Dirac).

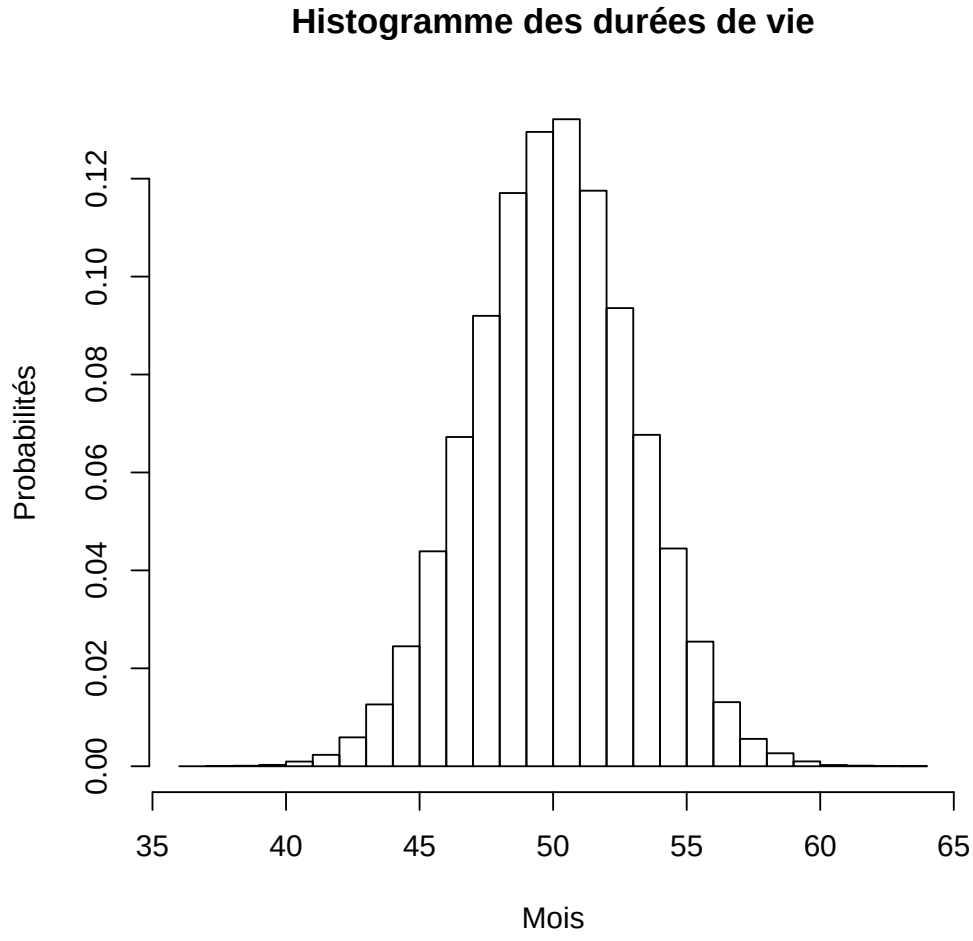
- si X et Y sont deux variables de Bernoulli indiquant chacune la présence d'une maladie différente alors
  - XY est de Bernoulli
    $\hookrightarrow$ indique la présence des deux maladies,
  - X^Y est de Bernoulli,
  - $X + Y$ est une binomiale (si indépendance des maladies)
    $\hookrightarrow$ indique le nombre de maladies.

Autres exemples

- Présence d'une anomalie génétique chez un individu.
- Etre favorable à un candidat.
- Réussite d'une greffe.

Remarques : $\hookrightarrow$ On peut aussi construire une Bernoulli à partir de n'importe quelle variable aléatoire comme le montre l'exemple suivant

On s'intéresse à des durées de vie après rechute d'une maladie.



On peut associer une variable de Bernoulli à chaque classe de l'histogramme.

Loi binomiale

Soient $X_1, \cdots, X_n$ des variables aléatoires i.i.d. (identiquement et indépendamment distribuées) de Bernoulli $B(p)$.

On pose $S = X_1 + \cdots + X_n$.

S suit une loi binomiale $B(n, p)$ définie par

$$\mathbb{P}(S = s) = \frac{n!}{(n-s)!s!} p^x (1-p)^{n-s},$$

pour $s = 0, 1, \cdots, n$

Propriétés :

- Moyenne et variance :

$$\mathbb{E}(S) = np, \qquad \mathbb{V}(S) = np(1-p).$$

- Si S_1 et S_2 sont deux binomiales $B(n_1, p)$ et $B(n_2, p)$ indépendantes alors $S_1 + S_2$ est une binomiale $B(n_1 + n_2, p)$.
- $\hookrightarrow$ faux s'il n'y a plus indépendance, ou si les probabilités p sont différentes.

Exemples

La loi binomiale apparaît comme un compteur (une somme de Bernoulli), elle apparaît aussi assez naturellement dans les "systèmes en parallèle" :

- Nombre de réacteurs en panne (parmi 4 réacteurs mutuellement indépendants) $\hookrightarrow B(4, p)$.
- Plus généralement, le nombre de réacteurs qui ont une durée de fonctionnement supérieure à 1000 heures $\hookrightarrow B(4, p(1000))$.
- Dans les familles de quatre enfants, combien de filles ?

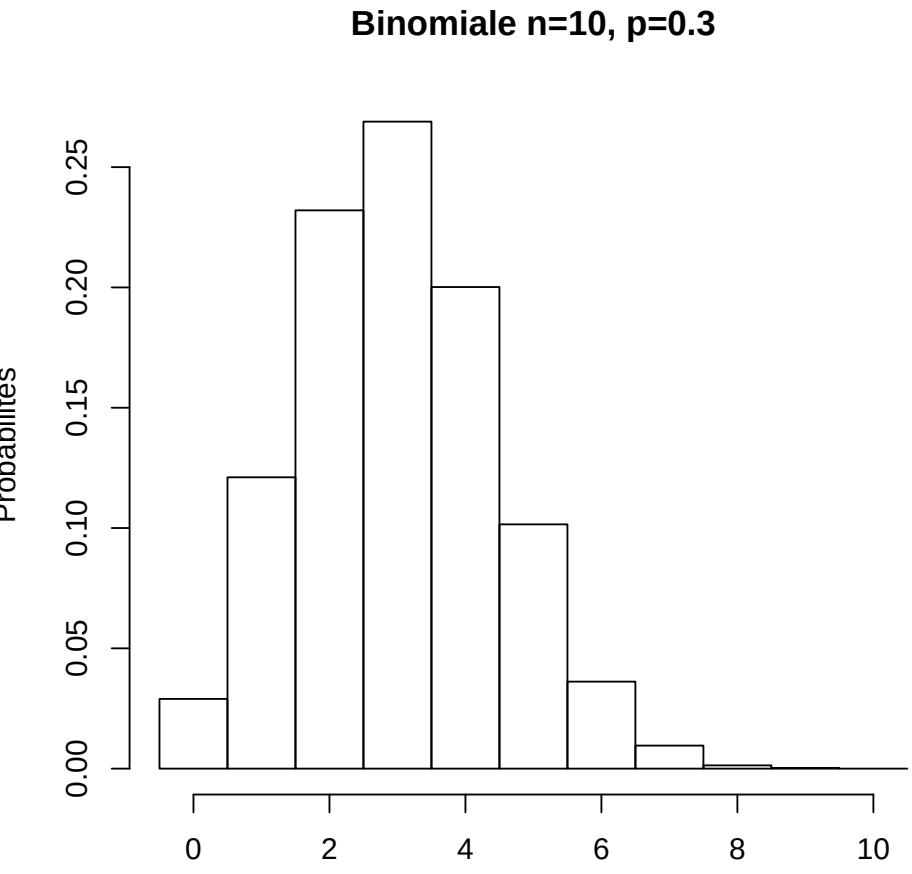
Contre-exemple

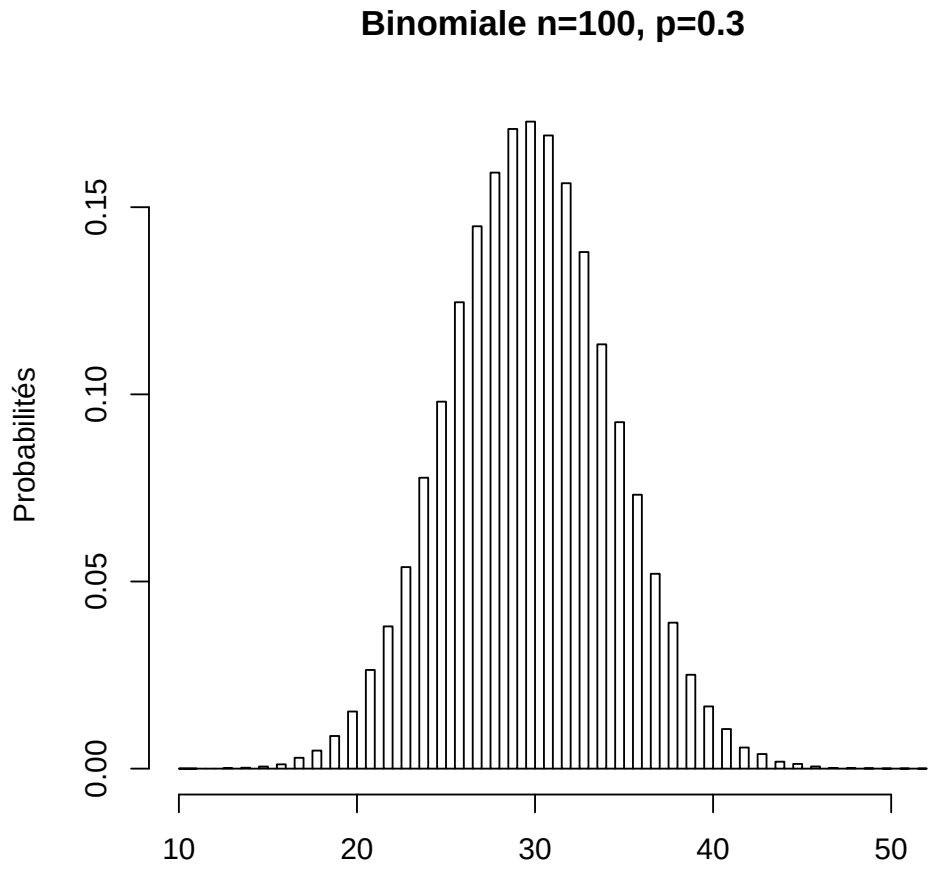
Deux maladies : la première contractée avec une probabilité $p_1 = 1/4$, la deuxième contractée avec une probabilité $p_{2|0} = 1/6$ si on n'a pas la première et $p_{2|1} = 1/2$ si on a déjà la première. On observe sur un même individu le nombre S de maladies (0,1 ou 2). Alors

$$\begin{aligned}
\mathbb{P}(S=0) &= (1-p_1)(1-p_{2|0}) = 15/24, \\
\mathbb{P}(S=1) &= p_1(1-p_{2|1}) + (1-p_1)p_{2|0} = 1/4, \\
\mathbb{P}(S=2) &= p_1 p_{2|1} = 1/8,
\end{aligned}$$

ce n'est pas une binomiale, pourtant les deux maladies sont des Bernoulli de même paramètre $1/4$.

Convergence vers une loi normale ?





On a $S \longrightarrow \infty$...

La loi normale

On approxime souvent la loi binomiale par une loi normale, qui pourtant est une loi à densité.

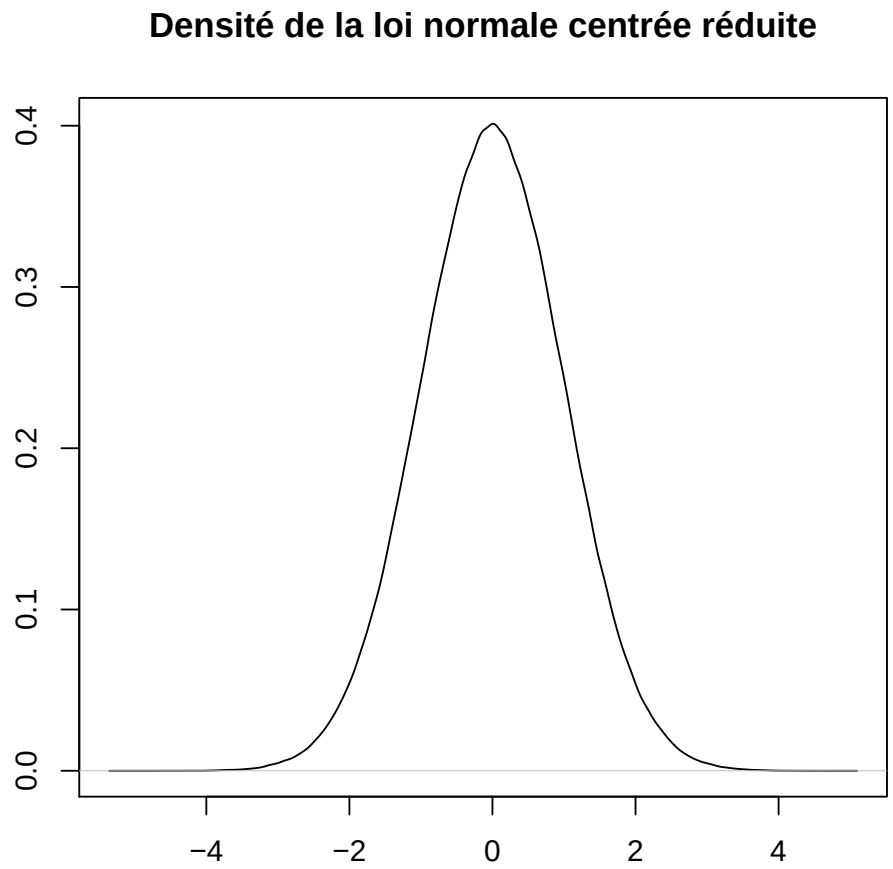
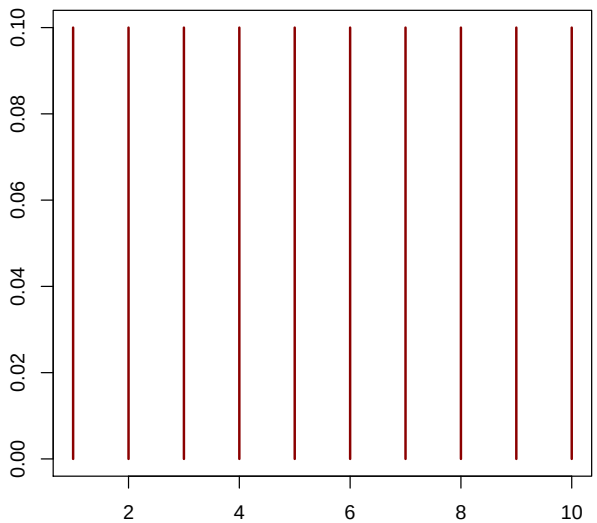
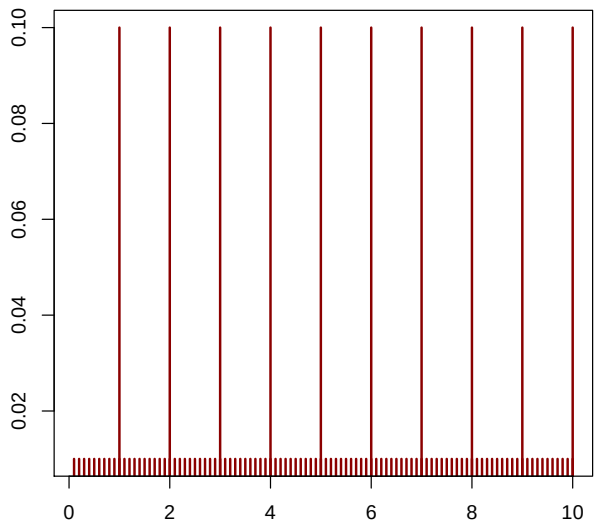


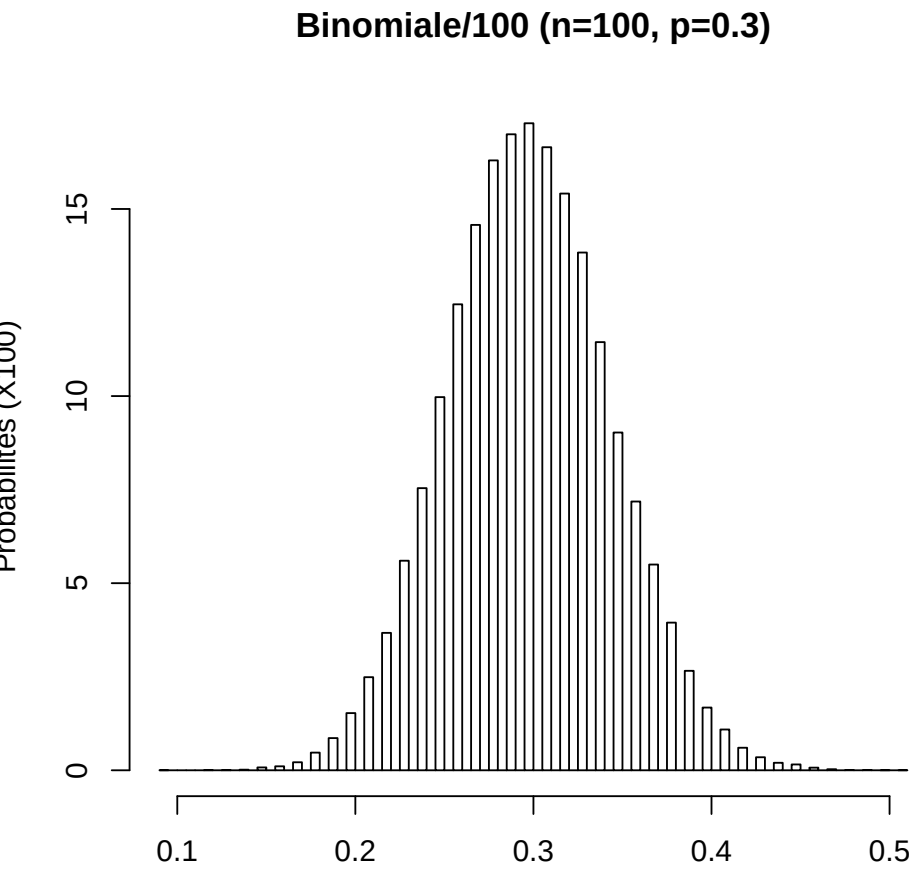
Illustration de la densité

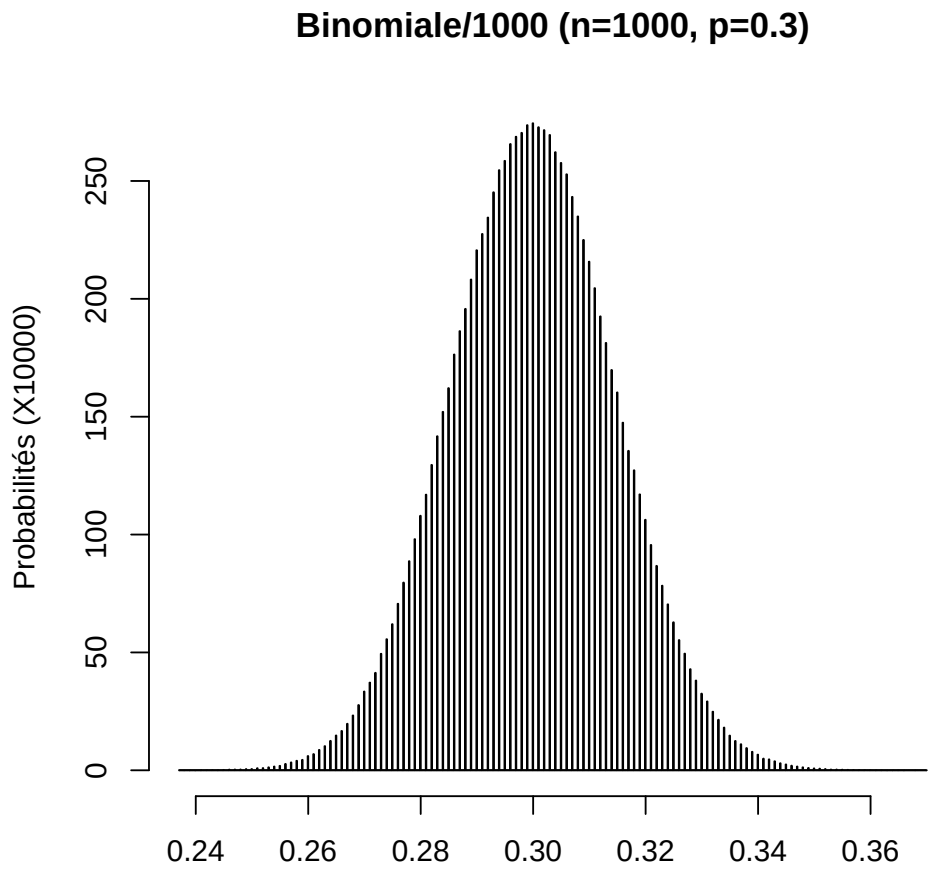
Sur une règle de dix centimètres, chaque centimètre a une probabilité uniforme d'être choisi. Puis on coupe en dix : chaque millimètre, etc...





Distributions de S/n





S/n reste entre 0 et 1 avec des probabilités de plus en plus petites.

On a $S/n \longrightarrow p$...

Que représente $\bar{X}$ (ou S/n) ?

$S/n = (X_1 + \cdots + X_n)/n.$

$\hookrightarrow$ est une variable aléatoire,

$\hookrightarrow$ est la moyenne empirique.

Dans le cas de Bernoulli

- $\mathbb{E}(S/n) = p$,
- $\mathbb{V}(S/n) = p(1-p)/n$.

Dans le cas général

- $\mathbb{E}(S/n) = \mathbb{E}(X) = m$,
- $\mathbb{V}(S/n) = \mathbb{V}(X)/n = \sigma^2/n$.

En moyenne S/n donne la bonne valeur de la moyenne ou de p.
$\hookrightarrow$ Estimateur sans biais

La variance de S/n tend vers zéro
$\hookrightarrow$ Estimateur convergent

Le Théorème de la Limite Centrale (TLC ou TCL)

On approxime souvent la loi binomiale par une loi normale On peut généraliser ce résultat grâce au théorème suivant.

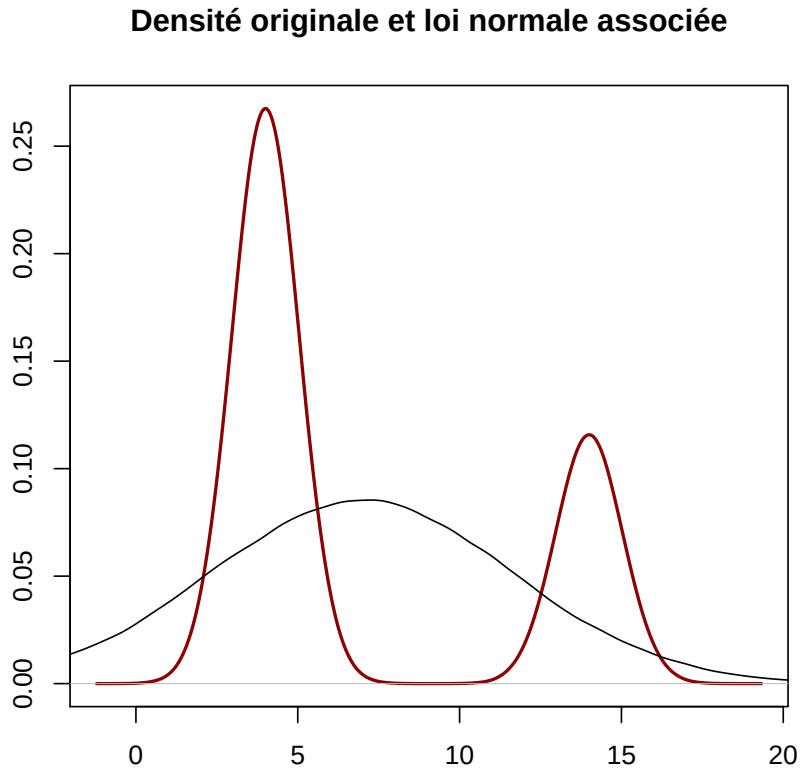
Théorème

Si $X_1, \cdots, X_n$ est une suite de variables aléatoires indépendantes de même loi (donc de même moyenne m et de même variance σ^2 supposée finie). Alors

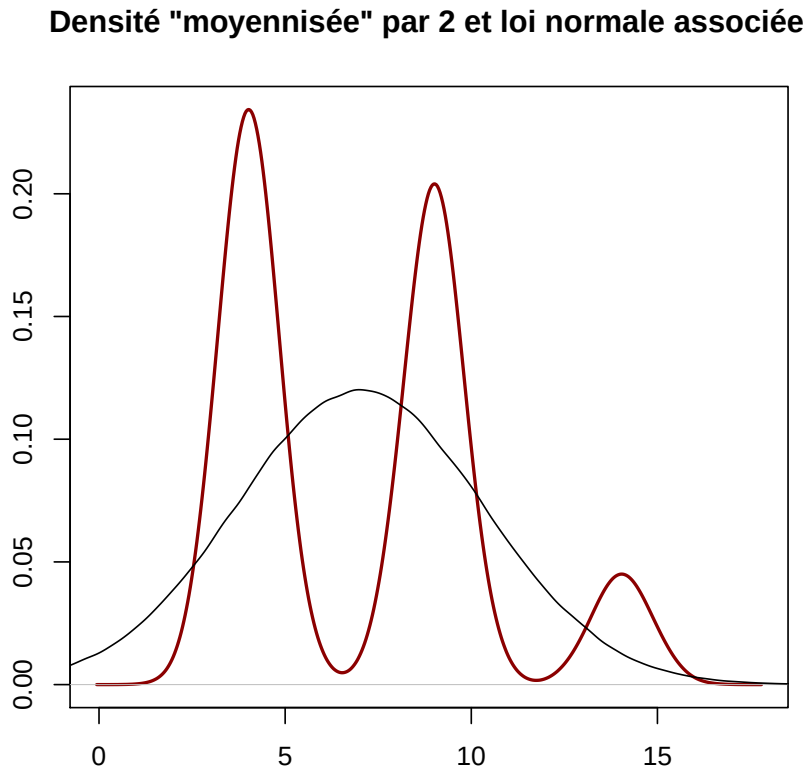
$$\sqrt{n}\,\frac{\bar{X} - m}{\sigma} \quad \xrightarrow{L} \quad \mathcal{N}(0,1),$$

où $\bar{X} = (X_1 + \cdots + X_n)/n = S/n$.

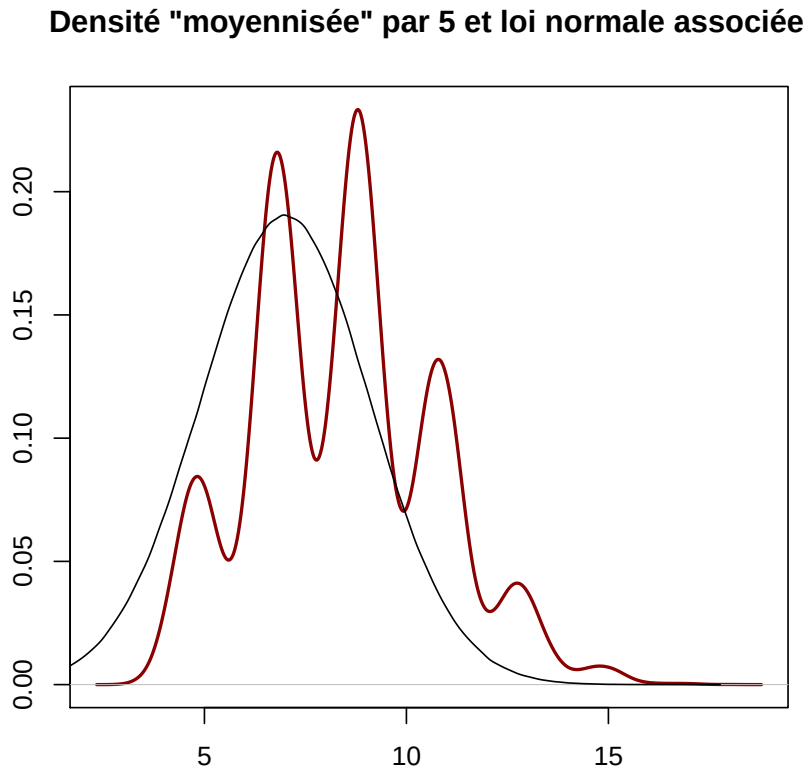
Illustrations du TCL



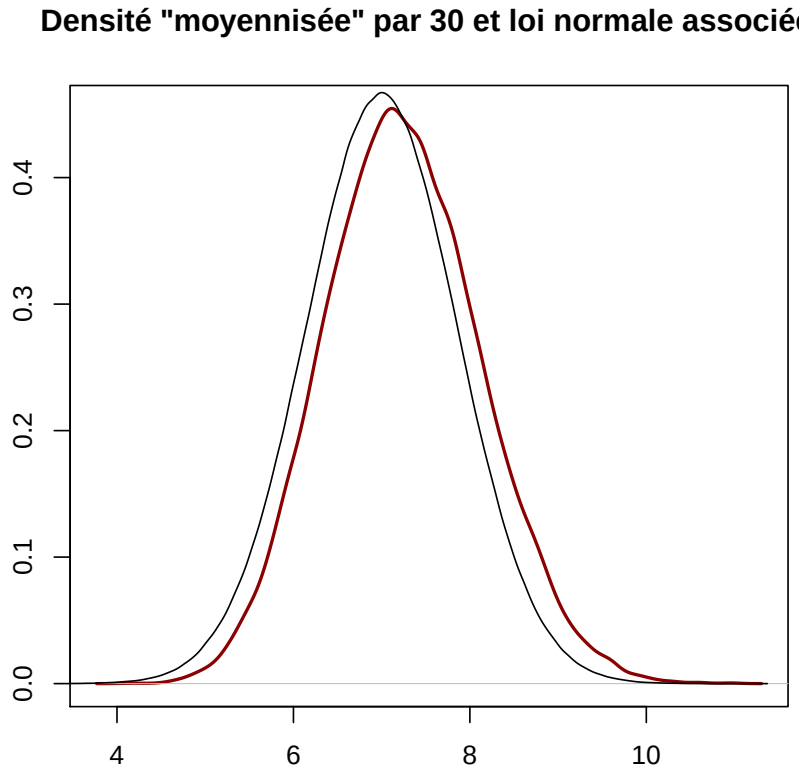
Illustrations du TCL



Illustrations du TCL



Illustrations du TCL



Approximation par la loi normale

D'après le TCL on a la cvce en loi suivante :

$$U = \sqrt{n}\,\frac{S/n - p}{\sqrt{p(1-p)}} \quad \to \quad \mathcal{N}(0,1),$$

ce qui signifie que la fonction de répartition $\mathbb{P}(U \leq u)$ converge vers la fonction de répartition d'une loi normale $\mathbb{P}(\mathcal{N}(0,1) \leq u)$.
D'où l'approximation pour "n grand" :

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

Revenons à

$$U = \sqrt{n}\,\frac{S/n - p}{\sqrt{p(1-p)}} \quad \to \quad \mathcal{N}(0,1).$$

Pour "n grand"

$$\begin{aligned} S/n &\approx \mathcal{N}(p, p(1-p)/n), \\ S &\approx \mathcal{N}(np, np(1-p)). \end{aligned}$$

Remarque : la qualité de l'approximation dépend de la valeur (inconnue) de p. Plus p est proche de 0.5 et plus on s'approche rapidement de la loi normale. On impose généralement $n > 30$, $np > 5$ et $np(1-p) > 5$, ce qui revient à vérifier que $S > 5$ et $S(1 - S/n) > 5$.

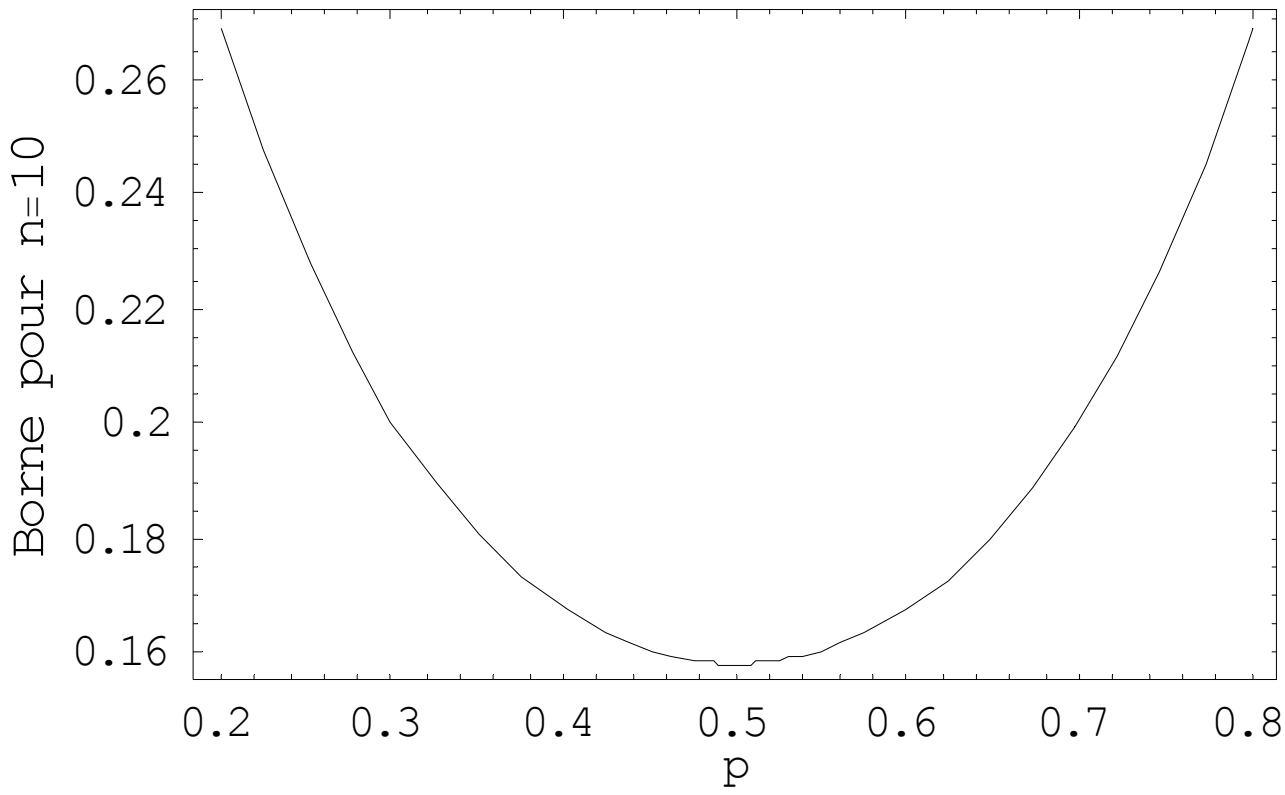
Erreur d'approximation

On a une borne de type Berry-Esseen :

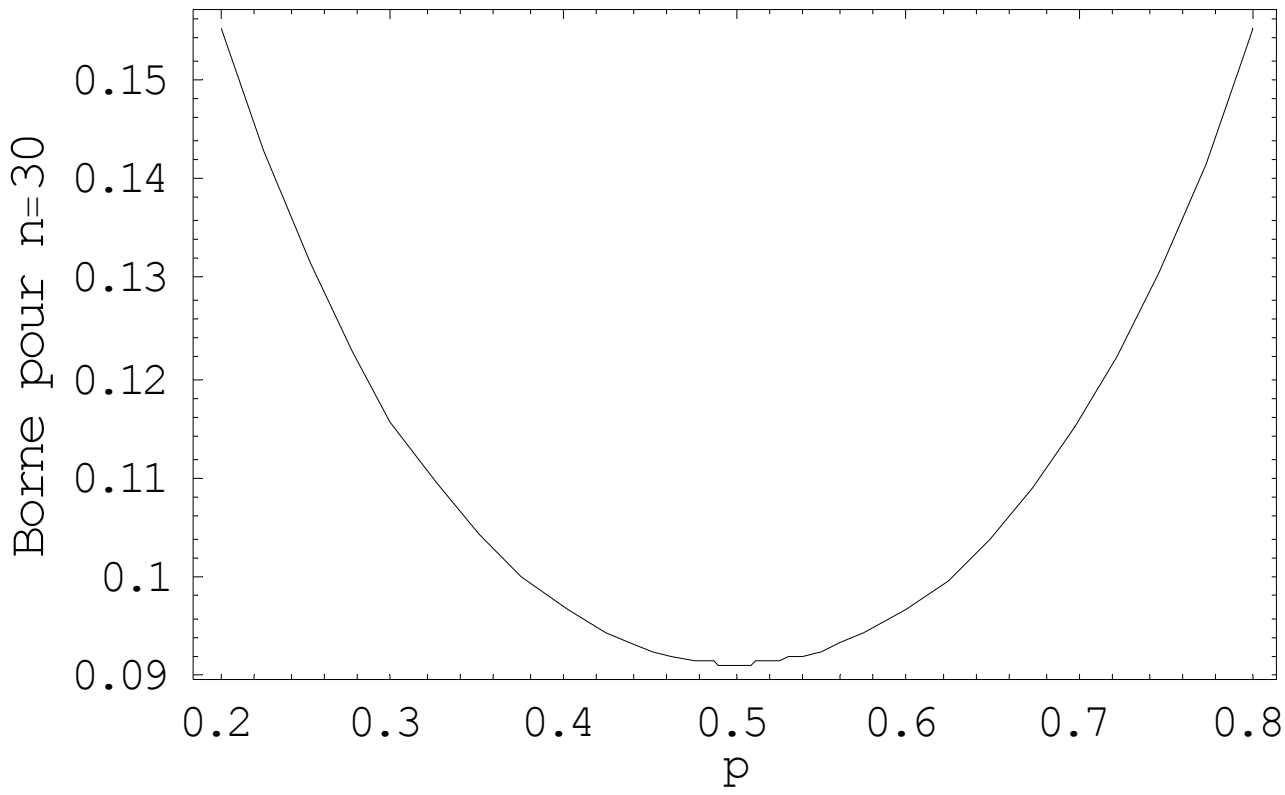
$$|\mathbb{P}(U \leq u) - \mathbb{P}(\mathcal{N}(0,1) \leq u)| \quad \leq \quad \frac{(1-p)^2 + p^2}{2\sqrt{np(1-p)}},$$

où $U = \sqrt{n}\,\dfrac{S/n - p}{\sqrt{p(1-p)}}$.

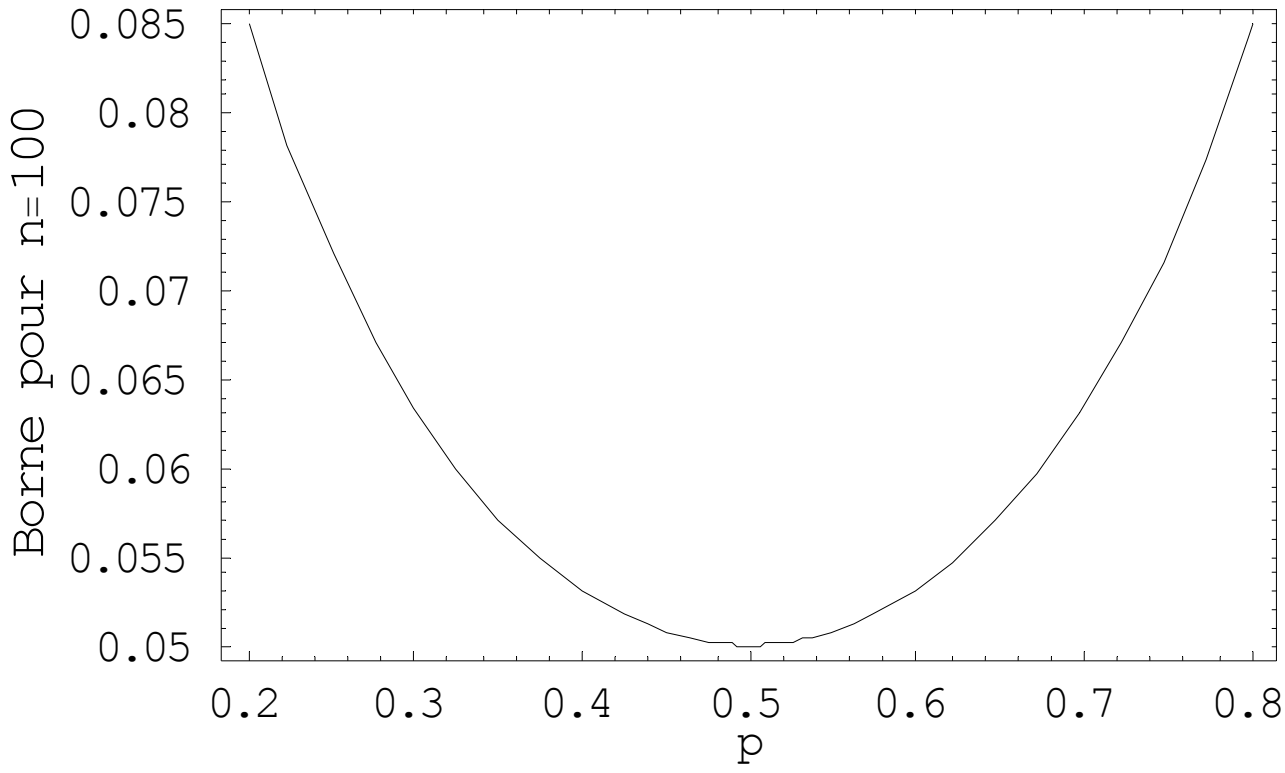
Borne pour $n = 10$ en fonction de p



Borne pour $n = 30$ en fonction de p



Borne pour $n = 100$ en fonction de p



Propriétés de la loi normale

Si $X \sim \mathcal{N}(m, \sigma^2)$ alors

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

Si $X \sim \mathcal{N}(m_1, \sigma_1^2)$ et $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ sont indépendantes alors

- $X + Y \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$,
- $X - Y \sim \mathcal{N}(m_1 - m_2, \sigma_1^2 + \sigma_2^2)$.

A partir de la loi normale

Si $X_1, \cdots, X_d$ sont i.i.d. $\sim \mathcal{N}(0,1)$ alors

- $T = X_1^2 + \cdots + X_d^2 \sim \chi^2_d$,
- Si on a une autre variable indépendante $X \sim \mathcal{N}(0,1)$, alors $\dfrac{X}{\sqrt{T/d}} \sim Student_d$.
- Si $U \sim \chi^2_k$ et $V \sim \chi^2_p$ sont indépendantes alors, $\dfrac{U/k}{V/p} \sim Fisher_{k,p}$.

Intervalle de confiance

On observe $X_1, \cdots, X_n$ i.i.d. de loi $B(p)$.
On s'intéresse à la valeur de $p \in]0,1[$ inconnue.
Par exemple pour chaque individu : 1 = satisfait, 0 = non satisfait.
$\hookrightarrow$ Quel est le pourcentage d'individus satisfaits dans la population ?

- On peut estimer ponctuellement p par S/n.
- On peut aussi proposer un intervalle de confiance lorsque "n est grand". C'est-à-dire a, b tels que $P(a \leq p \leq b) = 0.95$ par exemple.

Remarque : c'est a et b qui sont aléatoires (ils vont dépendre des $X_1, \cdots, X_n$)

On utilise l'approximation précédente combinée à la convergence (en probabilité) de S/n vers p :

$$\begin{aligned} \sqrt{n}\,\frac{S/n-p}{\sqrt{p(1-p)}} &\to \mathcal{N}(0,1) \text{ (en loi)}, \\ S/n &\to p \text{ (en probabilité)}, \end{aligned}$$

pour conclure à l'approximation pour "n grand" (Théorème de Slutsky)

$$T = \sqrt{n}\,\frac{S/n-p}{\sqrt{S/n(1-S/n)}} \approx \mathcal{N}(0,1).$$

Prenons ensuite u fractile de la loi normale tel que (par ex.) :

$$\mathbb{P}(-u \leq \mathcal{N}(0,1) \leq u) = 0.95$$

On en déduit l'intervalle de confiance de niveau 0.95 pour p (avec "n grand")

$$\mathbb{P}(-u \leq \sqrt{n}\,\frac{S/n - p}{\sqrt{S/n(1-S/n)}} \leq u) \quad = \quad 0.95$$

$$\mathbb{P}(\frac{S}{n} - 1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}} \leq p \leq \frac{S}{n} + 1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}}) \quad = \quad 0.95$$

On en déduit l'intervalle de confiance de niveau 0.95 pour p (avec "n grand")

$$\mathbb{P}(-u \leq \sqrt{n}\,\frac{S/n - p}{\sqrt{S/n(1-S/n)}} \leq u) \quad = \quad 0.95$$

$$\mathbb{P}(\frac{S}{n} - 1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}} \leq p \leq \frac{S}{n} + 1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}}) \quad = \quad 0.95$$

$$|\underbrace{\qquad\qquad}_{-1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}}}\underbrace{|}_{\frac{S}{n}}\underbrace{\qquad\qquad}_{+1.96\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}}}|$$

En résumé :

- S/n est l'estimation ponctuelle.
- $\frac{\sqrt{\frac{S}{n}(1-\frac{S}{n})}}{\sqrt{n}}$ est (une estimation de) l'écart-type de S/n
- 1.96 est associée au niveau

$$IC(p, 0.95) \quad = \quad [S/n \pm 1.96 \frac{\sqrt{S/n(1-S/n)}}{\sqrt{n}}]$$

L'amplitude de l'intervalle vaut $2\,u\,\dfrac{\sqrt{S/n(1-S/n)}}{\sqrt{n}}$.

- Plus n est grand et plus cette amplitude va diminuer,
- Plus le niveau de confiance est grand et plus l'amplitude va augmenter.

Remarque : lorsque S/n est entre 0.2 et 0.8, alors $\sqrt{S/n(1-S/n)} \in [0.4, 0.5]$ et on fait l'approximation suivante :

$$IC(p, 0.95) \quad \approx \quad [S/n \pm \frac{1}{\sqrt{n}}]$$

Test sur une valeur

Après une opération (appendicite) on veut s'assurer qu'un patient ne développe pas une infection (pouvant être causée par la présence d'un abcès). On réalise pour cela une simple prise de température toutes les 12h. En cas de non infection le risque habituel de poussée de température est de 1/4. En cas d'abcès ce risque passe à 1/2. Après 12h que peut-on décider après la première prise de température ? Après 24h, 36h, ... ?

Test sur une valeur

Est-ce qu'une décision doit être privilégiée ?

Risque d'être faux positif ?

Risque d'être faux négatif ?

Test sur une valeur

Deux cas sont envisageables :

- L'individu est malade : probabilité $p_0 = 1/2$ d'avoir de la température.
- L'individu est sain : probabilité $p_1 = 1/4$ d'avoir de la température.

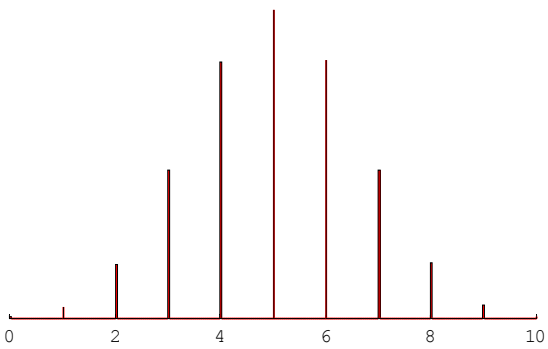
Le suivi d'un patient consiste à observer n températures (supposées indépendantes) et à décider si l'individu est malade ou non.

En notant p la probabilité d'avoir de la température on veut tester

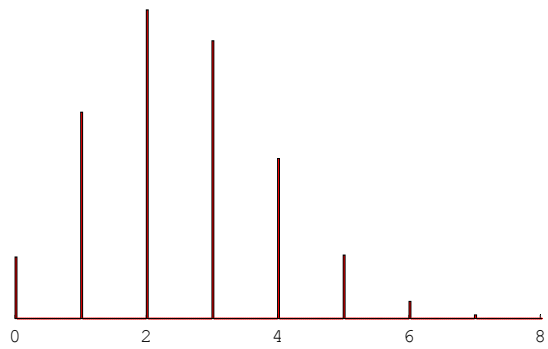
$$H_0 : p = p_0 = 1/2 \text{ (malade)} \quad VS \quad H_1 : p = p_1 = 1/4 \text{ (sain)}$$

- Sous H_0, S provient d'une $B(n, p_0)$.
- Sous H_1, S provient d'une $B(n, p_1)$.

Distributions de S avec $n = 10$ relevés :

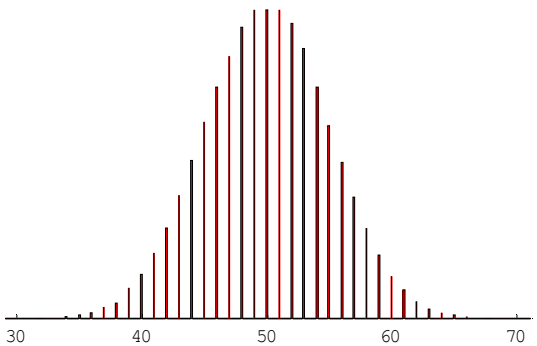


Sous H_0

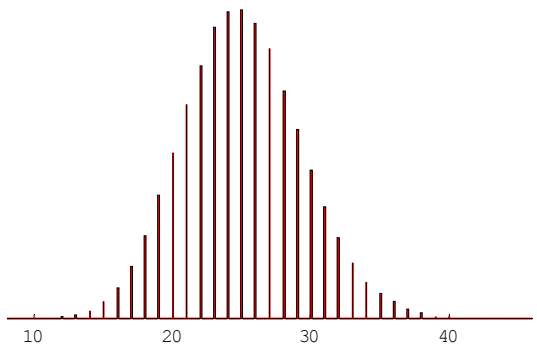


Sous H_1

Distributions de S avec $n = 100$ relevés

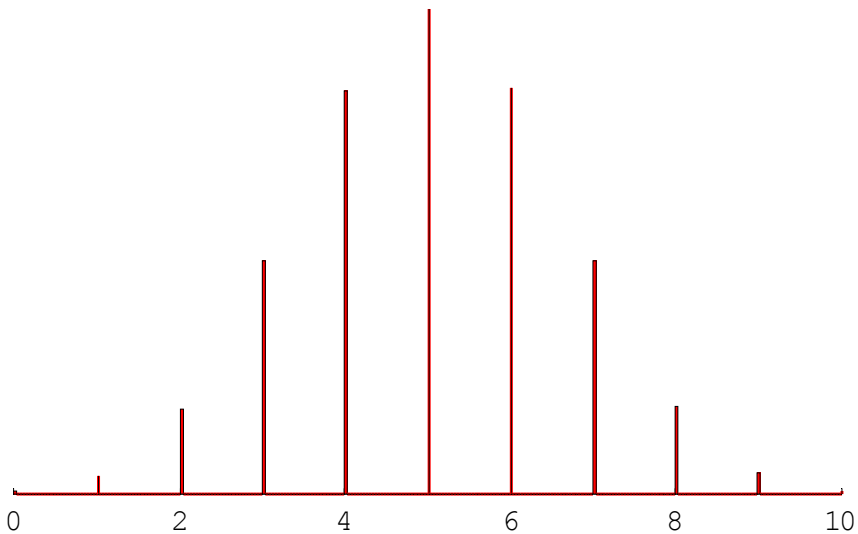


Sous H_0

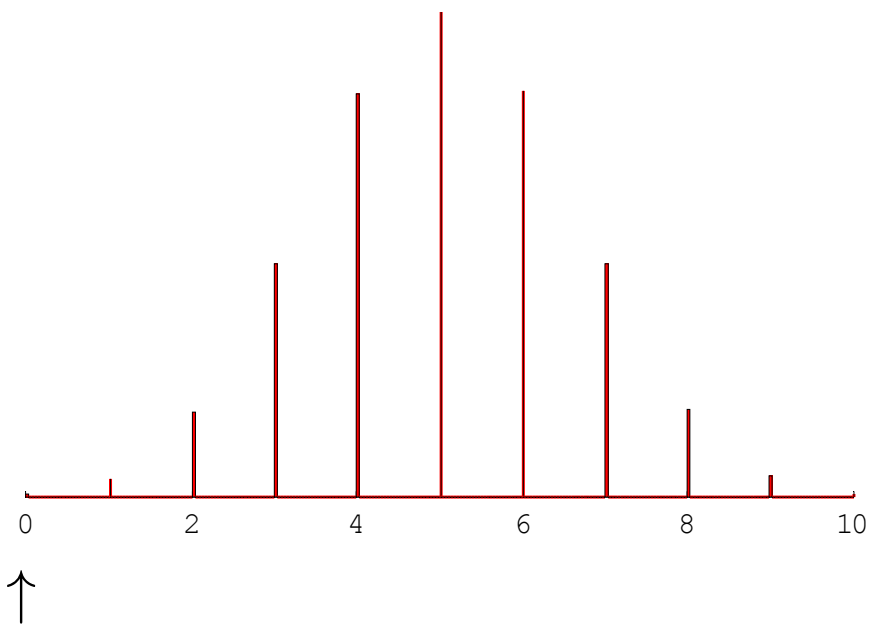


Sous H_1

On doit fixer une règle de décision pour rejeter H_0 avec un risque maîtrisé. Par exemple avec 10 observations on regarde la distribution théorique de S sous H_0 :

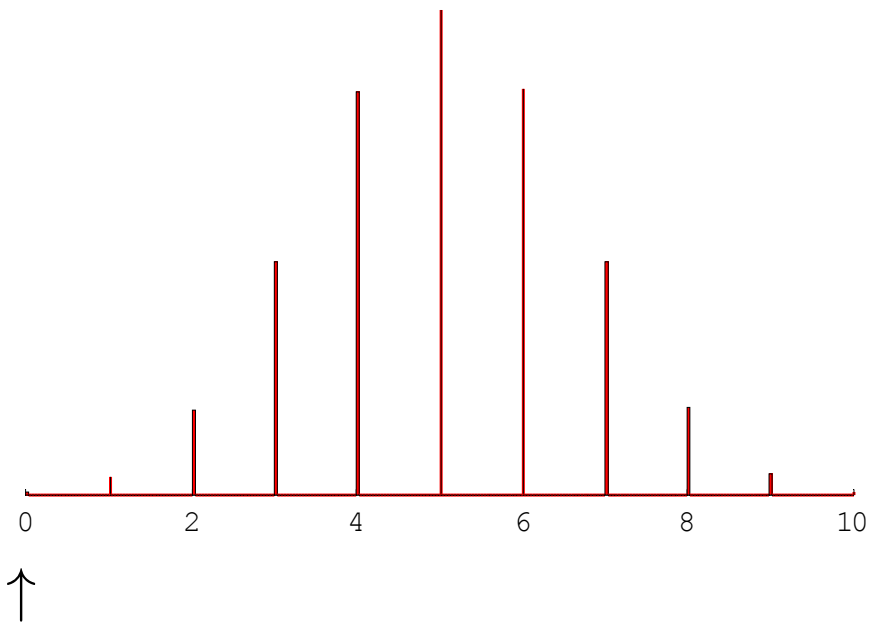


On doit fixer une règle de décision pour rejeter H_0 avec un risque maîtrisé. Par exemple avec 10 observations on regarde la distribution théorique de S sous H_0 :



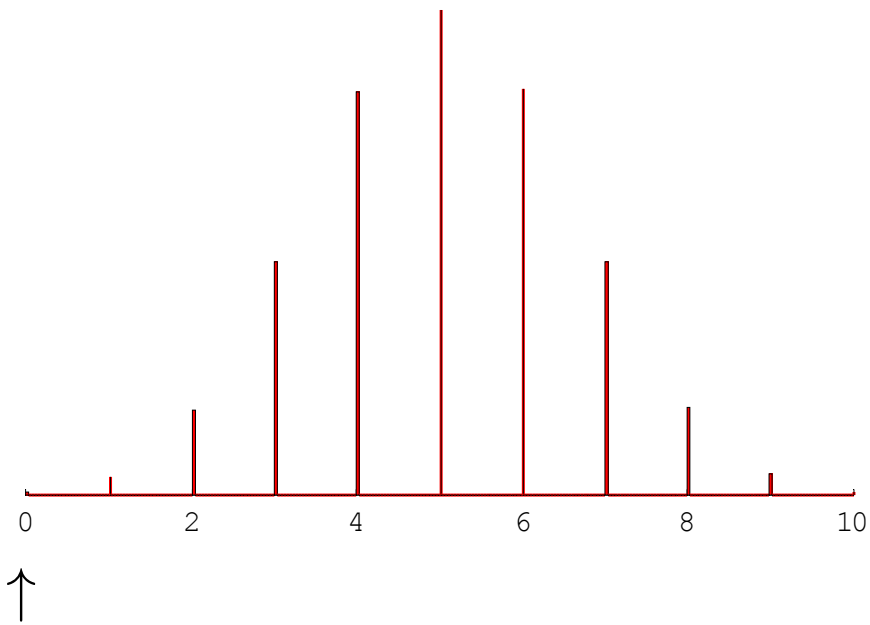
Probabilité faible sous H_0 ($\simeq 0.001$)

On doit fixer une règle de décision pour rejeter H_0 avec un risque maîtrisé. Par exemple avec 10 observations on regarde la distribution théorique de S sous H_0 :



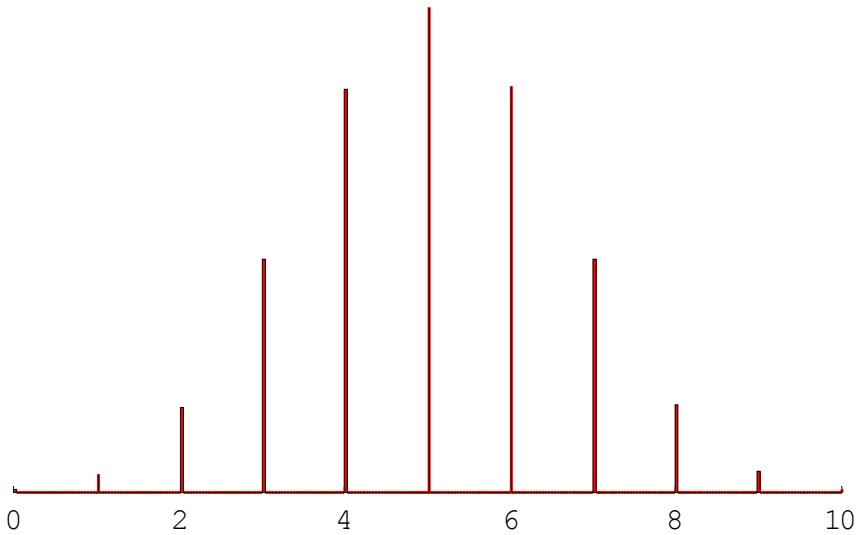
Probabilité faible sous H_0 ($\simeq 0.001$) $\hookrightarrow$ Règle : si $S = 0$ on rejette H_0

On doit fixer une règle de décision pour rejeter H_0 avec un risque maîtrisé. Par exemple avec 10 observations on regarde la distribution théorique de S sous H_0 :

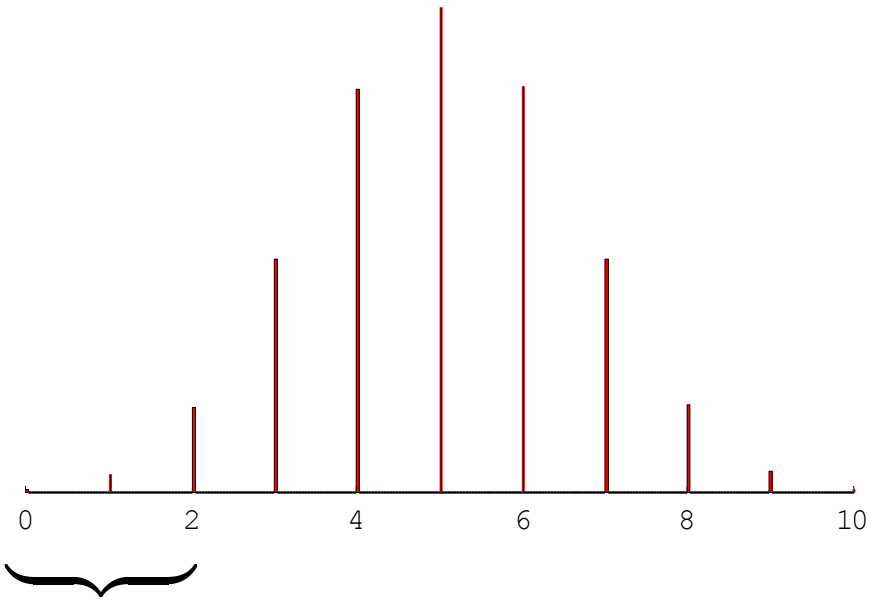


Probabilité faible sous H_0 ($\simeq 0.001$) $\hookrightarrow$ Règle : si $S = 0$ on rejette H_0 $\hookrightarrow$ Risque d'erreur en rejetant $H_0 \simeq 0.001$.

On peut aller plus loin (toujours avec 10 observations)



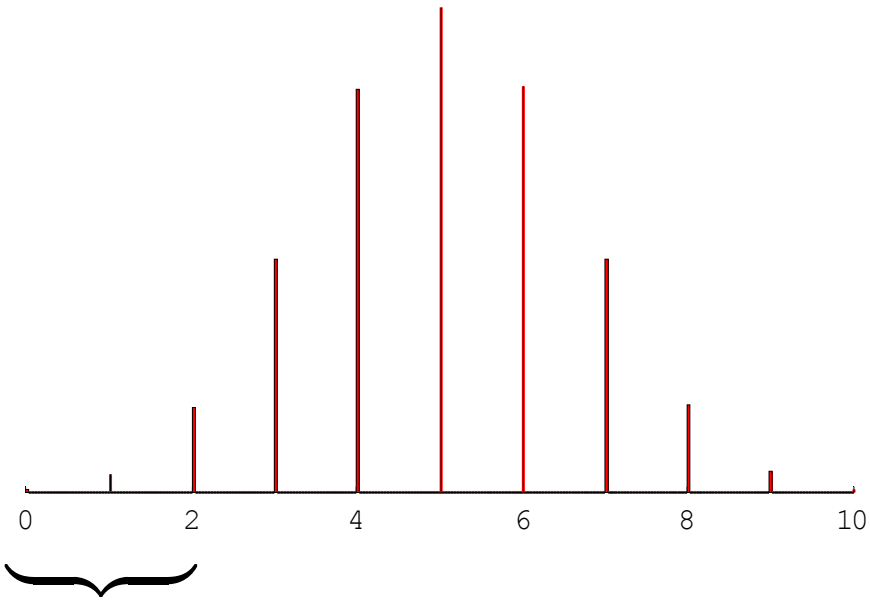
On peut aller plus loin (toujours avec 10 observations)



Probabilités assez faible sous H_0

$\mathbb{P}(S = 0, 1, 2) \simeq 0.055$

On peut aller plus loin (toujours avec 10 observations)



Probabilités assez faible sous H_0

$\mathbb{P}(S = 0, 1, 2) \simeq 0.055$

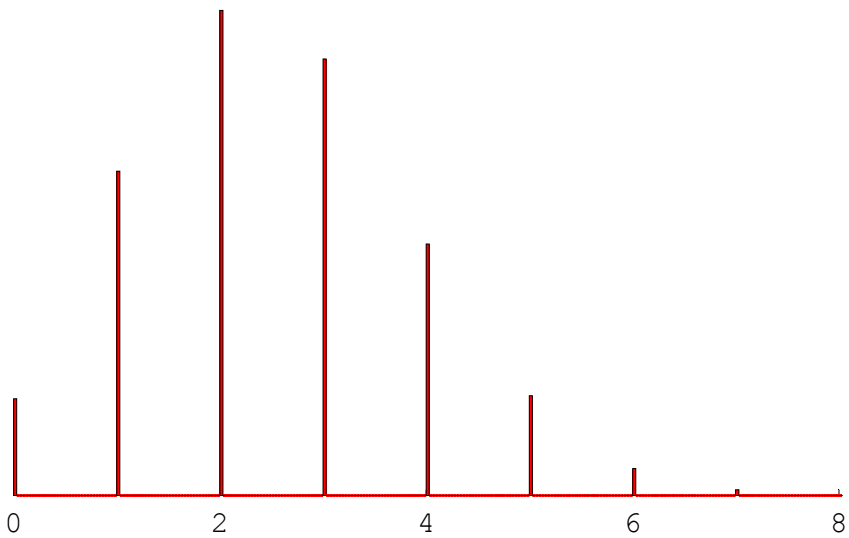
On rejette H_0 quand S vaut $0, 1$ ou 2.

Deux types d'erreurs :

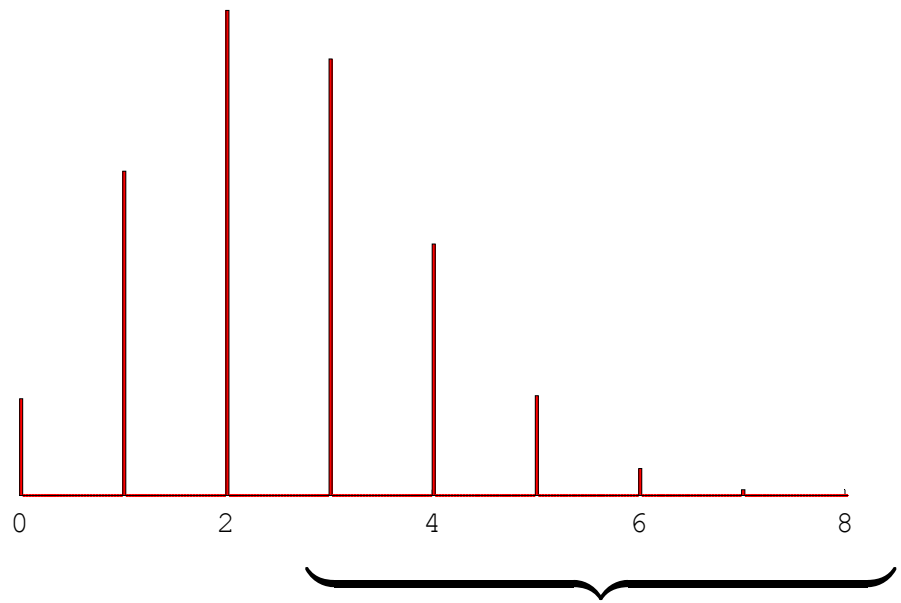
- En réalité H_0 est vraie (maladie) $\hookrightarrow$ erreur si je rejette H_0 $\hookrightarrow$ c'est le risque d'observer seulement $0, 1$ ou 2 pics de températures alors que l'individu est malade ($p = 1/2$). On le maîtrise : cette probabilité vaut environ 0.055.
- En réalité H_1 est vraie (individu sain) $\hookrightarrow$ erreur si j'accepte H_0 $\hookrightarrow$ c'est le risque d'observer $S = 3, 4, \cdots, 10$ pics de températures alors que l'individu est sain ($p = 1/4$).

On va calculer ce deuxième risque.

Calcul de la probabilité d'erreur en acceptant H_0 (alors que l'individu est sain)



Calcul de la probabilité d'erreur en acceptant H_0 (alors que l'individu est sain)



Sous H_1, $\mathbb{P}(S = 3, 4, \cdots, 10) \simeq 0.71$
Donc si l'individu est sain on peut se tromper avec une forte probabilité (dans 71% des cas).

On peut calculer l'erreur globale de se tromper (quelle que soit la décision). On suppose (au début de l'expérience) que l'individu a une probablité q d'être malade.

$$\begin{aligned}
\mathbb{P}(erreur) &= \mathbb{P}(erreur \cap H_0) + \mathbb{P}(erreur \cap H_1) \\
&= \mathbb{P}(erreur|H_0) * \mathbb{P}(H_0) + \mathbb{P}(erreur|H_1) * \mathbb{P}(H_1) \\
&= 0.055 * q + 0.71 * (1 - q) \\
&= \begin{cases} 0.64 & q = 0.1 \\ 0.38 & q = 0.5 \\ 0.07 & q = 0.9 \end{cases}
\end{aligned}$$

On voit bien ici qu'il vaut mieux s'intéresser aux taux de faux négatifs...

Cas intéressant : on ne fait qu'un seul relevé de température ($n = 1$).

Règle de décision :

- On observe de la température $\hookrightarrow$ on décide H_0 (malade).
- On n'observe pas de température $\hookrightarrow$ on décide H_1 (sain).

Les erreurs possibles :

* Erreur si l'individu est malade : $P(\bar{T}|H_0) = 1/2$ (faux négatif)
* Erreur si l'individu est sain : $P(T|H_1) = 1/4$ (faux positif)
* Erreur globale :
  $P(pile \cap H_1 \cup face \cap H_0) = (1/4 * q + 1/2 * (1-q)) = 1/2 - q/4$

Si on propose une autre règle :

Règle de décision 2 :

- On observe de la température $\hookrightarrow$ on décide H_0 (malade).
- On n'observe pas de température $\hookrightarrow$ on décide H_1 (sain).

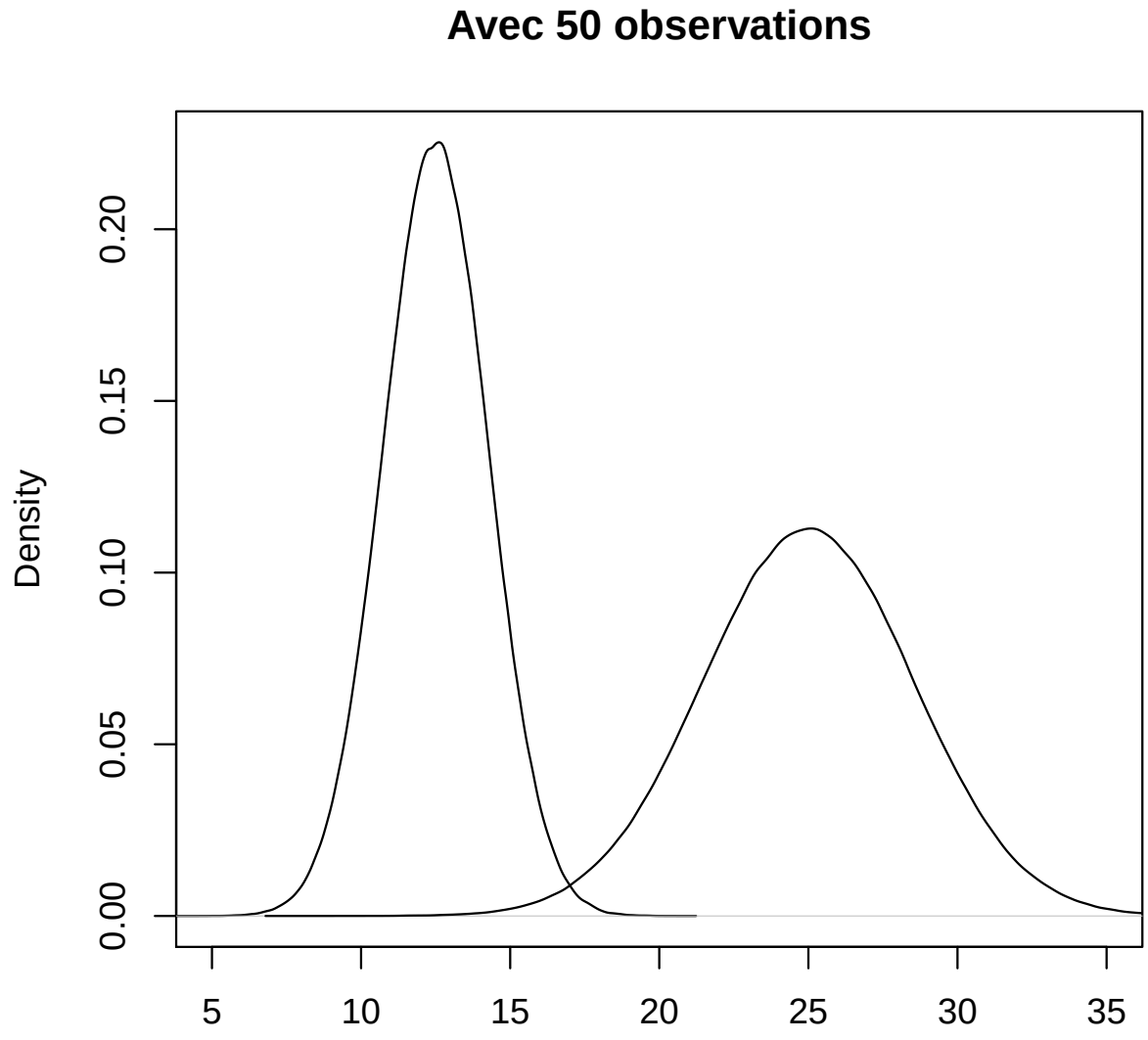
Les erreurs possibles :

* Erreur si individu sain : $P(face|H_1) = 3/4$
* Erreur si individu malade : $P(pile|H_0) = 1/2$
* Erreur globale : $P(T \cap H_0 \cup \bar{T} \cap H_1) = 1/2 + q/4$

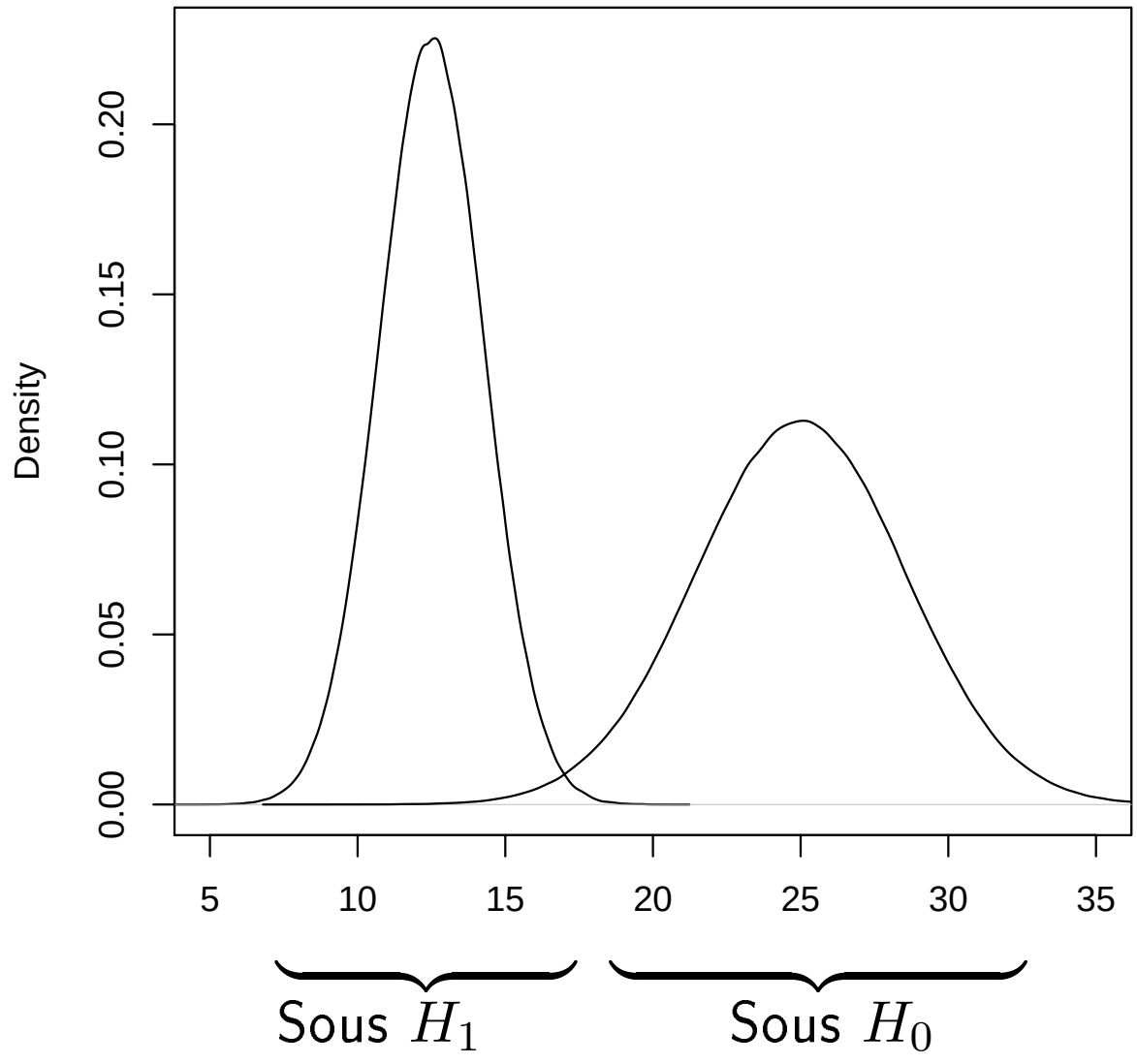
Remarque : dans ce cas il vaut mieux décider au hasard (une chance sur deux de se tromper).

Approximation par une loi normale

Et avec $n = 50$ observations, si nous faisons l'approximation de la binomiale par une loi normale alors :


Avec 50 observations
Density
0.00
0.05
0.10
0.15
0.20
5
10
15
20
25
30
35

Avec 50 observations



Démonstration du TLC

Pour simplifier supposons que les variables de la suite $X_1, \cdots, X_n$ sont centrées ($m = \mathbb{E}(X) = 0$) et réduites ($\sigma^2 = \mathbb{V}(X) = 1$) (quitte à retrancher m et à diviser par σ).
On sait que la fonction caractéristique de la loi normale centrée réduite $\mathcal{N}(0,1)$ est

$$\varphi_Z(t) = \mathbb{E}\big(\exp(itZ)\big) \quad = \quad \exp(-t^2/2),$$

où $Z \sim \mathcal{N}(0,1)$.
Si la fonction caractéristique de S/n tend (quand $n \to \infty$) vers $\exp(-t^2/2)$ alors la loi de S/n tend vers la loi normale centrée réduite (Théorème de Lévy).

On va utiliser trois propriétés importantes de la fonction caractéristique :

$$\varphi_X\left(\frac{t}{\sqrt{n}}\right) \quad = \quad \varphi_{\frac{X}{\sqrt{n}}}(t),$$

et si X_1 et X_2 sont indépendantes alors

$$\varphi_{X_1+X_2}(t) \quad = \quad \varphi_{X_1}(t)\varphi_{X_2}(t),$$

et donc si X_1 et X_2 ont même fonction caractéristique (i.e. même loi) :

$$\varphi_{X_1+X_2}(t) \quad = \quad \varphi_X(t)^2$$

Et une dernière propriété : si la variance de X existe alors

$$\begin{aligned} \varphi'_X(0) &= i\mathbb{E}(X) \\ \varphi''_X(0) &= -\mathbb{E}(X^2). \end{aligned}$$

Posons (après avoir centré et réduit les X_i)

$$U = \sqrt{n}\,\frac{S/n - m}{\sigma} = S/\sqrt{n}.$$

Les X_i étant indépendantes et de même fonction caractéristique on a

$$\begin{aligned}
\varphi_U(t) &= \varphi_{\frac{S}{\sqrt{n}}}(t) \\
&= \varphi_S\Big(\frac{t}{\sqrt{n}}\Big) \\
&= \varphi_{X_1+\cdots+X_n}\Big(\frac{t}{\sqrt{n}}\Big) \\
&= \varphi_{X_1}\Big(\frac{t}{\sqrt{n}}\Big)\cdots\varphi_{X_n}\Big(\frac{t}{\sqrt{n}}\Big) \\
&= \varphi_X\Big(\frac{t}{\sqrt{n}}\Big)^n
\end{aligned}$$

On fait alors un d.l. à l'ordre 2 :

$$\begin{aligned}
\left(\varphi_X\left(\frac{t}{\sqrt{n}}\right)\right)^n &= \left(\varphi_X(0) + i\frac{t}{\sqrt{n}}\varphi'_X(0) + (i^2)\frac{t^2}{2n}\varphi''_X(0) + o(1/n)\right)^n \\
&= \left(1 - \frac{t^2}{2n} + o(1/n)\right)^n \\
&\longrightarrow \exp(-t^2/2) \ \ (\text{quand } n \to \infty)
\end{aligned}$$

Test sur deux valeurs

On observe deux échantillons de Bernoulli indépendantes : $X_1, \cdots, X_n$ et $Y_1, \cdots, Y_k$. On teste :

$$H_0 : p_X = p_Y \quad VS \quad H_1 : p_X \neq p_Y$$

On utilise l'approximation par la loi normale :

$$\begin{aligned} S_X/n &\approx N(p_X, p_X(1-p_X)/n) \\ S_Y/k &\approx N(p_Y, p_Y(1-p_Y)/k). \end{aligned}$$

L'indépendance nous permet d'écrire

$$S_X/n - S_Y/k \approx N(p_X - p_Y, V)$$

avec $V = \dfrac{S_X/n(1-S_X/n)}{n} + \dfrac{S_Y/k(1-S_Y/k)}{k}$.

Donc, si H_0 est vraie on s'attend à avoir une valeur issue de la loi normale centrée réduite de T

$$T = \frac{S_X/n - S_Y/k}{\sqrt{V}},$$

on décide alors de rejeter ou non H_0.

Généralisation au cas d'une moyenne

Le TCL s'applique pour toute suite de variables i.i.d. de moyenne m ayant une variance σ^2 finie. On a

$$U = \sqrt{n}\,\frac{\bar{X} - m}{\sqrt{\sigma}} \quad \rightarrow \quad \mathcal{N}(0,1),$$

ce qui donne l'approximation pour n grand :

$$\bar{X} \quad \approx \quad \mathcal{N}(m, \sigma^2/n).$$

Intervalle de confiance

On observe $X_1, \cdots, X_n$ i.i.d. de moyenne m et de variance σ^2 inconnues. Par exemple des durèes de vie.
On utilise l'approximation précédente combinée à la convergence (en probabilité) de la variance empirique $\mathbf{S^2} = \sum(X_i - \bar{X})^2/n$:

$$\begin{aligned} U = \sqrt{n}\frac{\bar{X} - m}{\sigma} &\rightarrow \mathcal{N}(0,1) \text{ (en loi)}, \\ \mathbf{S} &\rightarrow \sigma \text{ (en probabilité)}, \end{aligned}$$

pour conclure

$$T = \sqrt{n}\frac{\bar{X} - m}{\mathbf{S}} \approx \mathcal{N}(0,1).$$

Prenons ensuite u fractile de la loi normale tel que :

$$\mathbb{P}(-u \leq T \leq u) \approx 0.95,$$

on en déduit l'intervalle de confiance de niveau 0.95 asymptotique pour m

$$IC(p, 0.95) = [\bar{X} \pm 1.96\frac{\mathbf{S}}{\sqrt{n}}]$$