Probabilités et Statistiques

Hugues GENVRIN

March 26, 2026

Table of contents

Analyse

Mauvaise interprétation de la courbe en cloche ou Gaussienne

Origine

Gauss définit la loi normale comme une loi des erreurs d'observations.

A.Fuchs : "En 1809, Carl Friedrich Gauss assimile des erreurs d'observation en astronomie à la courbe, dite des erreurs, de la densité d'une loi normale".

Dangers

Une distribution d'observations quelconque n'est pas forcément une courbe en cloche (faire un test de normalité).

Ce n'est pas parce que la moyenne arithmétique converge vers la moyenne par la Loi des Grands Nombres, que les écarts par rapport à la moyenne sont des erreurs.

Définitions et But

Statistique : Méthodes mathématiques pour découvrir des propriétés et relations entre des populations à partir d'échantillons.

Inférence : On appelle inférence une proposition issue d'un raisonnement logique.

Statistique inférentielle

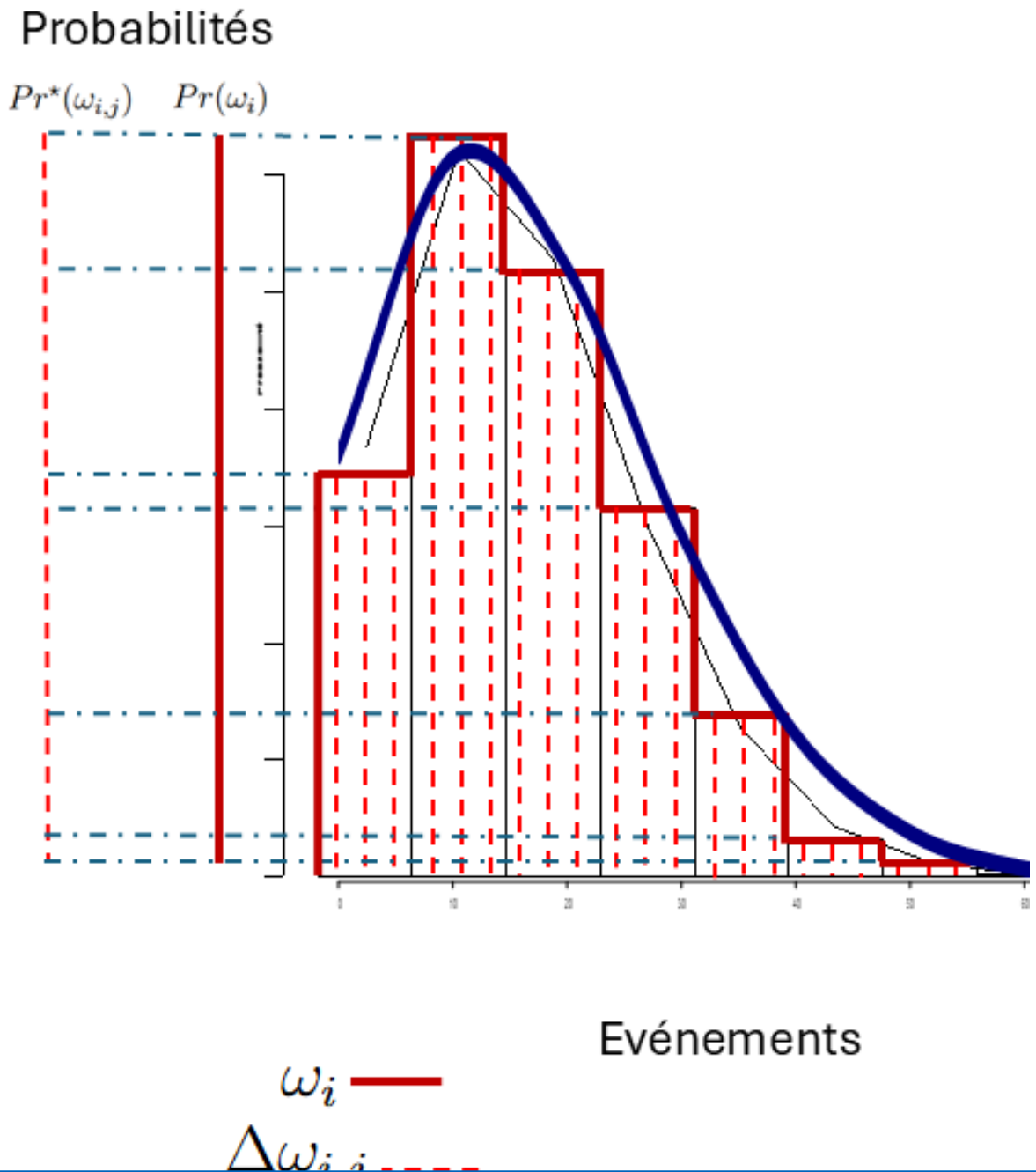
Saporta : "Son but est d'étendre les propriétés constatées sur l'échantillon à la population toute entière et de valider ou d'infirmer des hypothèses formulées après une phase exploratoire".

Un théorème de densité

Theorem

La densité d'une variable aléatoire vaut un.

C'est l'aire délimitée par la courbe et l'axe des abscisses.


Probabilités
$Pr^*(\omega_{i,j})$
$Pr(\omega_i)$
Evénements
ω_i
$\Delta\omega_{i,j}$

La courbe normale

Theorem

Une fonction de Laplace-Gauss décrit une variable aléatoire.

Soit une loi normale centrée et réduite, alors elle est d'équation $f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}}$. Montrons que sa densité est unitaire.

Autrement dit : $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1.$

Proof.

On remarque que f est symétrique par rapport à $(y'Oy)$, de telle sorte que $f(x) = f(-x)$. Posons $I = \int_{-\infty}^{+\infty} e^{\frac{x^2}{2}} dx$. Par le théorème de Fubini,
$J^2 = \int_{-\infty}^{+\infty} e^{\frac{x^2}{2}} dx \int_{-\infty}^{+\infty} e^{\frac{y^2}{2}} dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{(x^2+y^2)}{2}} dxdy$. On procède à un changement de variables, on passe en coordonnées polaires $(x, y) \mapsto (\phi, r)$. Alors ϕ est définie sur $[0; 2\pi[$, et r sur $[0; +\infty[$. On a $r^2 = x^2 + y^2$. Et $d(\frac{-1}{2}r^2) = rdr = rd(-r)$.
Comme f est symétrique, I et I^2 le sont aussi. D'où $d(-r) = dr$.
Alors $I = \int_0^{2\pi} d\phi \int_0^{+\infty} re^{\frac{-1}{2}r^2} dr$.

Posons $J = \int_0^{+\infty} re^{\frac{-1}{2}r^2} dr = -\int_0^{+\infty} -\frac{1}{2} \times 2re^{\frac{-1}{2}r^2} dr =$
$-\int_0^{+\infty} d(\exp \circ \frac{-1}{2}r^2) = -[e^{-\frac{1}{2}r^2}]_0^{+\infty} = 1$. D'où
$I^2 = 2\pi \Leftrightarrow I = \sqrt{2\pi}$. En conséquences :
$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1$.

Statistisques

Tests d'hypothèses

Soient deux estimateurs d'une expérience aléatoire :

- L'hypothèse nulle H_0 suppose qu'il n'y a pas de différence entre ces deux estimateurs.
- Le rejet de l'hypothèse nulle $\overline{H_0}$ montre une différence entre ces deux estimateurs.

Il y a deux risques d'erreur :

- L'erreur de type I (ou alpha) qui consiste à dire qu'il y a un rejet de l'hypothèse nulle, sans avoir une différence significative supportée par une p-value.
- L'erreur de type II (ou beta), qui consiste à dire qu'il n'y a pas de différence significative, par manque de puissance du test.

Interprétation de la p-value

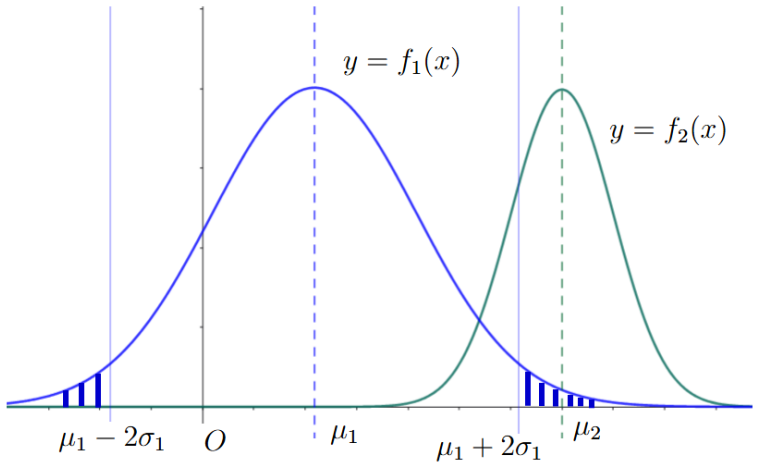


Figure: Origine de la p-value.

Pour une p-value de 0.05

$$\Pr\left(|\mu_2 - \mu_1| > \mu_1 + 2\sigma_1\right) < 0.05$$

Limite des erreurs de type II

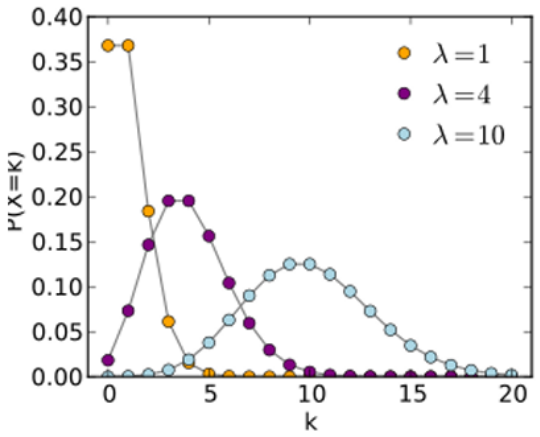
Les erreurs de type II s'appuient sur un écart-type qui pourrait tendre vers zéro, ce qui est faux dans sa généralité. Nous allons montrer que l'écart-type d'une série converge, sans forcément tendre vers zéro.
Ceci est particulièrement vrai si la variable aléatoire suit une distribution de Poisson.
Nécessité de trouver un indicateur d'arrêt d'une inclusion de cas dans une étude.

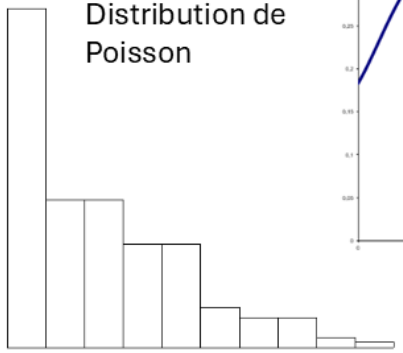
Distribution de Poisson

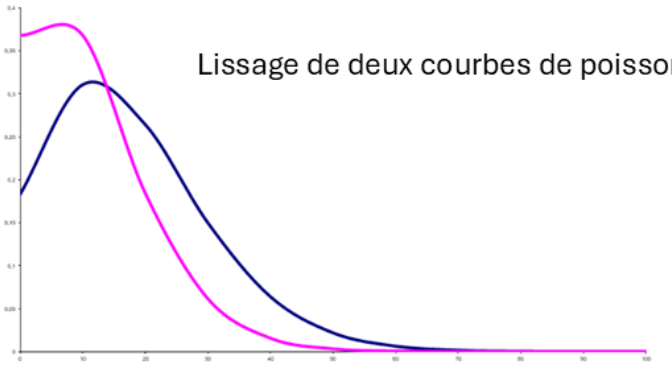
D'après D.Raupp "De l'extinction des espèces", la plupart des distributions de la nature sont des lois des Poisson.

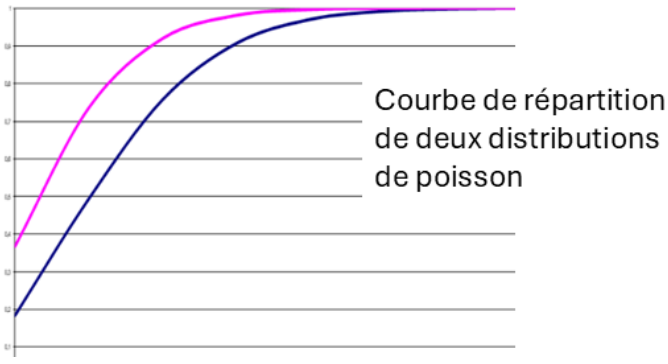
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda}$$









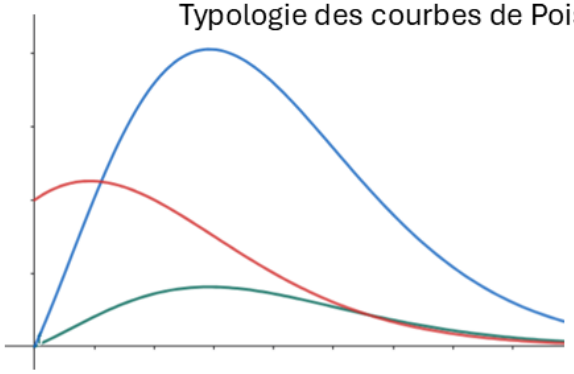


Figure: Les distributions de Poisson.

Expression de la variance

Soit une série de données $(x_i)_{1\leq i\leq n}$. On note $\overline{x}$ la moyenne arithmétique, $\overline{x}_q = \sum_{i=1}^{n} \frac{x_i^2}{n}$ la moyenne quadratique, et on appelle $V_n = \sum_{i=1}^{n} \frac{(x_i-\overline{x})^2}{n}$ la variance.
Montrons que la variance ne converge pas systématiquement vers zéro.

Montrons tout d'abord que (V_n) converge.
Soit $V_{n+1} = \sum_{i=1}^{n+1} \frac{(x_i-\overline{x})^2}{n} = \sum_{i=1}^{n} \frac{(x_i-\overline{x})^2}{n} \times \frac{n}{n+1} + \frac{(x_{n+1}-\overline{x})^2}{n+1} = V_n \times \frac{1}{1+\frac{1}{n}} + \frac{(x_{n+1}-\overline{x})^2}{n+1}$.
Comme $(x_i)_{1\leq i\leq n}$ est bornée, lorsque n tend vers l'infini, l'expression de la variance $V_{n+1} \to V_n \Leftrightarrow V_{n+1} - V_n \to 0$.
Donc, $(V_i)_{1\leq i\leq n}$ est convergente vers une valeur réelle positive.
En conséquence l'écart-type $(\sigma_i)_{1\leq i\leq n} = \sqrt{V_i}$ est convergent.

Condition nécessaire

Montrons que l'expression de $(V_i)_{1\leq i\leq n}$ ne converge pas systématiquement vers zéro. Soit
$V_n = \sum_{i=1}^{n} \frac{(x_i-\overline{x})^2}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2) =$
$\sum_{i=1}^{n} \frac{x_i^2}{n} - 2\overline{x}\sum_{i=1}^{n} \frac{x_i}{n} + \overline{x}^2 = \overline{x}_q - 2\overline{x}\ \overline{x} + \overline{x}^2 = \overline{x}_q - \overline{x}^2$. Donc, une condition nécessaire pour que la variance converge vers zéro est que la moyenne quadratique vaut la moyenne arithmétique au carré.

D'un contre-exemple vers un cas de base

Considérons un nombre pair de données. On suppose que $x_i = \text{Sup}$ pour la moitié des données et $x_i = \text{Inf}$ pour l'autre moitié (pour maximiser la variance). Alors $\overline{x} = \frac{1}{2}(\text{Sup} + \text{Inf})$, tandis que $\sqrt{\overline{x}_q} = \sqrt{\frac{\frac{n(\text{Sup}^2+\text{Inf}^2)}{2}}{n}}$.
En élevant au carré de part et d'autre, et en simplifiant les expressions. Il vient $\frac{1}{4}\text{Sup}^2 + \frac{1}{4}\text{Inf}^2 + \frac{1}{2} \times \text{Sup} \times \text{Inf} = \frac{1}{2}\text{Sup}^2 + \frac{1}{2}\text{Inf}^2 \Leftrightarrow \text{Sup} \times \text{Inf} = \frac{1}{2}\text{Sup}^2 + \frac{1}{2}\text{Inf}^2$, ce qui est manifestement faux pour une multitude de cas.S

On conclut que la variance et l'écart-type ne convergent pas systématiquement vers zéro.

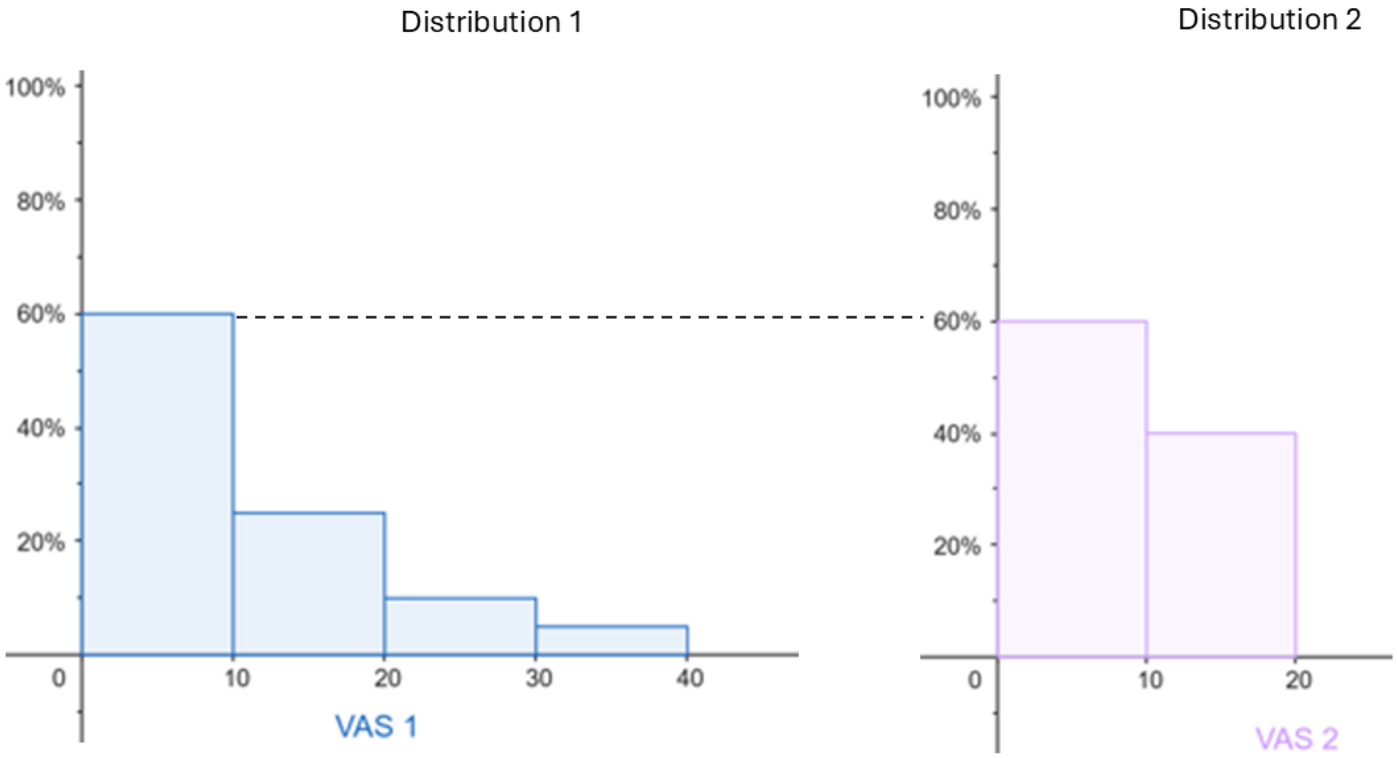


Figure: Exemple d'une distribution du cas de base.

Quand stopper une inclusion de cas ?

Heuristique

On ne sait pas calculer la puissance d'un test, nécessaire quelle que soit la courbe. Il faut donc utiliser une heuristique pour arrêter l'inclusion de cas.

Lorsque le rapport de la variance $\frac{V_n}{V_{2n}} \geq 0.95$. Où n renseigne le nombre de cas.

Probabilité baysésienne

Généralisation de la loi de Bayes

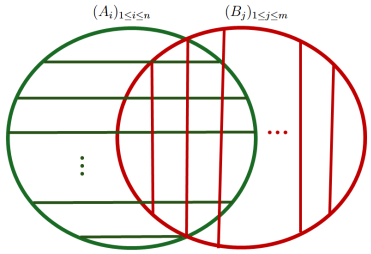


Figure: Preuve de la loi de Bayes.

D'où :

$$\Pr\big((\cup \mathrm{A_i}) \cap (\cup \mathrm{B_j})\big) = \Pr(\cup_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{A_i} | \cup_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{B_j}) \times \Pr(\cup_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{B_j}) \quad (1)$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{m}} \Pr(\mathrm{A_i}|\mathrm{B_j}) \times \Pr(\mathrm{B_j}) \quad (2)$$

$$= \Pr(\cup_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{B_j} | \cup_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{A_i}) \times \Pr(\cup_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{A_i}) \quad (3)$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{m}} \Pr(\mathrm{B_j}|\mathrm{A_i}) \times \Pr(\mathrm{A_i}) \quad (4)$$

Causes et Effet

Soit $(C_j)_{1\leq j\leq n}$ un ensemble de causes, et C_i une cause parmi les n causes potentielles. On a $\overline{C_i}$ qui signifie l'ensemble des causes complémentaires. On appelle E l'effet généré par les causes (C_j).

$$\begin{aligned}
Pr(C_i|E) &= \frac{Pr(E|C_i)\times Pr(C_i)}{Pr(E)}\text{; un cas pratique : } Pr(E)=1\\
&= \frac{Pr(E|C_i)\times Pr(C_i)}{Pr(E|C_i)\times Pr(C_i)+Pr(E|\overline{C_i})\times Pr(\overline{C_i})}\\
&= \frac{Pr(E|C_i)\times Pr(C_i)}{Pr(E|C_i)\times Pr(C_i)+Pr(E|\overline{C_i})\times (1-Pr(C_i))}
\end{aligned}$$

Par argument de symétrie :

$$\boxed{Pr(E|C_i) = \frac{Pr(C_i|E)\times Pr(E)}{Pr(C_i|E)\times Pr(E)+Pr(C_i|\overline{E})\times (1-Pr(E))}}$$

Expression de la probabilité de Laplace

Principe de raison insuffisante

Nous allons alors appliquer une probabilité inverse pour $\widehat{\mathcal{H}_{\mathrm{d}}}$, qui sera une probabilité conditionnelle. En conséquence, la composition sera aussi définie par une probabilité bayésienne. Voici les termes engagés :

- $\Delta\mathbb{E}_{\mathrm{d}} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \Pr_{\mathrm{i}}[\text{signe i}|\text{fait}] \times \mathcal{I}\big(\text{fait}\big)$
- $\Delta\mathbb{E}_{\mathrm{c}} = \sum_{\mathrm{j}=1}^{\mathrm{m}} \Pr_{\mathrm{j}}[\text{fait}|\text{signe j}] \times \mathcal{I}\big(\text{signe j}\big)$
- $\Delta\mathbb{E}_{\mathrm{d}}^{|\mathrm{i}} = \Pr_{\mathrm{i}}[\text{signe i}|\text{fait}] \times \mathcal{I}\big(\text{fait}\big)$
- $\Delta\mathbb{E}_{\mathrm{c}}^{|\mathrm{j}} = \Pr_{\mathrm{j}}[\text{fait}|\text{signe j}] \times \mathcal{I}\big(\text{signe j}\big)$

$\mathcal{I}$ est le sens relatif à l'argument. On peut développer les probabilités bayésiennes et considérer les restrictions des espérances $\mathcal{E}^{|}$. Pour $\mathcal{E}_{\mathrm{d}}^{|\mathrm{k_i}}$, on définit la restriction de l'espérance de l'unité $\mathrm{k_i}$ pour la décomposition.

Table of contents