Calcul d'erreur

F. Kany. ISEN-Brest & La Croix-Rouge

1 Problème de la marche au hasard. Loi de Gauss

1.1 Position du problème

Exemple :

On mesure 100 m avec 1 mètre que l'on reporte 100 fois avec, à chaque fois, une incertitude de 1 cm. Quelle est l'erreur totale commise ? 100 x 1 cm = 1 m ?

FAUX. Il est peu probable qu'à chaque fois, on fasse l'erreur de 1 cm dans le même sens.

Probabilité $\frac{1}{2}$ de déplacer le mètre de 1 cm vers la droite.
Probabilité $\frac{1}{2}$ de déplacer le mètre de 1 cm vers la gauche.
Probabilité de déplacer 100 fois de suite le mètre dans le même sens : $(\frac{1}{2})^{100} = 7,9.10^{-31}$.
La probabilité d'avoir une incertitude de 1 m est ridicule. Il serait exagéré d'écrire : il est probable que $X \in [99\text{ m}\,;\,101\text{ m}]$.

1.2 Marche au hasard

Généralisation du problème : la marche au hasard.
Soit p la probabilité de se déplacer vers la droite.
Soit $q = 1 - p$ la probabilité de se déplacer vers la gauche.

Au bout de N itérations, quelle est la probabilité P de faire au total : n_1 pas à droite et n_2 pas à gauche ?

Probabilité de faire n_1 pas à droite : p^{n_1} ; probabilité de faire n_2 pas à gauche : q^{n_2} ; nombre de façons de réaliser ce déplacement : $C_N^{n_1} = C_N^{n_2} = \frac{N!}{n_1!.n_2!}$

D'où : $P = p^{n_1}.q^{n_2}.\frac{N!}{n_1!.n_2!}$

Soit m : le déplacement total vers la droite : $m = n_1 - n_2$.
On a : $N = n_1 + n_2$. D'où : $n_1 = \frac{N+m}{2}$ et $n_2 = \frac{N-m}{2}$

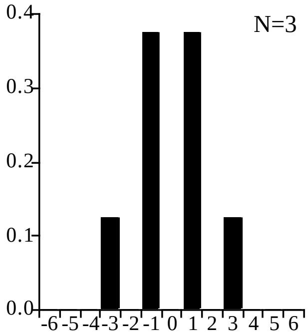
$$P_N(m) = p^{\frac{N+m}{2}}.q^{\frac{N-m}{2}}.\frac{N!}{(\frac{N+m}{2})!.(\frac{N-m}{2})!}$$

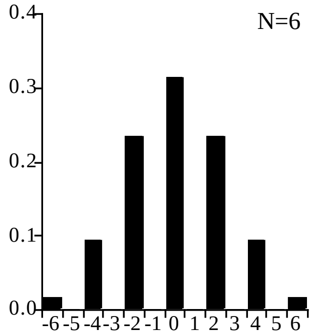
C'est la loi binomiale (cf. développement de $(p+q)^N = \sum_m P_N(m)$).

Calcul des C_j^i.

$$\begin{array}{ccccccc} 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 4 & 6 & 4 & 1 \\ 1 & 5 & 10 & 10 & 5 & 1 \\ 1 & 6 & 15 & 20 & 15 & 6 & 1 \end{array}$$

Cas particulier $p = q = 1/2$. $P_N(m) = (\frac{1}{2})^N.\frac{N!}{(\frac{N+m}{2})!.(\frac{N-m}{2})!}$

Exemple : 



1.3 Cas où N$\rightarrow +\infty$

Lorsque N est très grand, on montre[1] que la probabilité de se trouver à la position m est : $P_N(m) \approx \sqrt{\frac{2}{\pi.N}}.e^{\frac{-m^2}{2.N}}$

— Probabilité de se trouver entre m et $m+\Delta m$: $\Delta P = \frac{1}{2}P_N(m).\Delta m$
(car $P_N(m)$ est à peu près constante sur $[m, m+\Delta m]$; le facteur $\frac{1}{2}$ vient du fait que $P_N(m)$ n'est définit que pour une valeur de m sur deux : les m pairs (resp. impairs) si N est pair (resp. impair) ; il faut donc éliminer une valeur sur deux par rapport à une fonction $P_N(m)$ qui serait définie pour tout m.)

— Nombre d'événements ΔN tel que l'on se trouve entre m et $m+\Delta m$: $\Delta N = N.P_N(m)$

— Écart-type : $\sigma^2 = \frac{1}{N}.\sum((m-\overline{m})^2.\Delta N) = \sum m^2.\Delta P \approx \sqrt{\frac{2}{\pi.N}}.\sum m^2.e^{\frac{-m^2}{2.N}}.\frac{\Delta m}{2}$
Si le pas Δm est petit, alors, par passage à la limite, on a : $\sigma^2 = \sqrt{\frac{1}{2.\pi.N}}.\int_{-\infty}^{+\infty} x^2.e^{\frac{-x^2}{2.N}}.\mathrm{d}x.\frac{2.N}{2.N}.\frac{\sqrt{2.N}}{\sqrt{2.N}}$
$\sigma^2 = \frac{2.N}{\sqrt{\pi}}.\int_{-\infty}^{+\infty} u^2.e^{-u^2}.\mathrm{d}u$ avec $u = \frac{x}{\sqrt{2.N}}$.
Intégration par parties : $f = u/2$ et $g' = 2.u.e^{-u^2}$
$\Rightarrow f' = 1/2$ et $g = -e^{-u^2}$
$\int_{-\infty}^{+\infty} u^2.e^{-u^2}.\mathrm{d}u = \left[\frac{u}{2}.e^{-u^2}\right]_{-\infty}^{+\infty} + \frac{1}{2}.\int e^{-u^2}.\mathrm{d}u$
$= \frac{1}{2}.\int e^{-u^2}.\mathrm{d}u$

— Calcul de $I = \int_{-\infty}^{+\infty} e^{-x^2}.\mathrm{d}x$.
Astuce : poser $J = \int_{-\infty}^{+\infty} e^{-y^2}.dy$
et calculer $I.J = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)}.\mathrm{d}x.\mathrm{d}y$.
On passe en polaire : $I.J = \int_{r=0}^{r=+\infty}\int_{\theta=0}^{\theta=2.\pi} e^{-r^2}.r.\mathrm{d}r.\mathrm{d}\theta$.
On pose : $\alpha = r^2 \Rightarrow \mathrm{d}\alpha = 2.r.\mathrm{d}r$
$I.J = 2.\pi.\frac{1}{2}.\int_{\alpha=0}^{\alpha=+\infty} e^{-\alpha}.\mathrm{d}\alpha = \pi.[-e^{-\alpha}]_0^{+\infty} = \pi \Rightarrow I = \sqrt{\pi}$.
D'où : $\sigma^2 = \frac{2.N}{\sqrt{\pi}}.\frac{\sqrt{\pi}}{2} = N$

$$\boxed{\sigma = \sqrt{N}}$$

1.4 Autre démonstration

σ_N^2 représente la moyenne du carré de la distance (m^2) parcourue au bout de N itérations. σ_N est donc une mesure de la distance parcourue au bout de N itérations.

Si $m = 1$, alors il est évident que $\sigma_1 = 1$.

Si, au bout de N itérations, la distance parcourue est σ_N ; alors à la $(N+1)^{ème}$ itération, on a :

$$\sigma_{N+1} = \sigma_N + 1 \quad \textbf{ou} \quad \sigma_{N+1} = \sigma_N - 1$$

D'où : $\begin{cases} \sigma_{N+1}^2 = \sigma_N^2 + 2.\sigma_N + 1 \\ \quad\quad \textbf{ou} \\ \sigma_{N+1}^2 = \sigma_N^2 - 2.\sigma_N + 1 \end{cases}$

1. Ce résultat se démontre en utilisant la formule de Stirling $\ln(n!) \simeq \frac{1}{2}\ln(2\pi) + (n+\frac{1}{2}).\ln n - n$ et en faisant : $\frac{m}{N} \ll 1$

et donc en moyenne : $\sigma_{N+1}^2 = \sigma_N^2 + 1$.

Comme $\sigma_1 = 1$, il est évident que $\sigma_{N+1}^2 = N + 1$;

d'où : $\boxed{\sigma_N = \sqrt{N}}$

1.5 Retour sur l'exemple initial

Application :

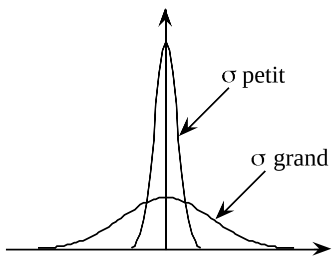
1 mètre que l'on reporte 100 fois $\Rightarrow N = 100 \Rightarrow \sigma = 10$.
L'incertitude est donc de 10x1 cm=10 cm (et non 100 cm !)
Donc, il est probable que $\ell \in$[99,9 m ; 100,1 m]

Nouvelle question :

Puisque l'on a considérablement réduit l'intervalle, quelle est la probabilité que ℓ soit réellement dans cet intervalle ?

1.6 Loi de Gauss

— Passage en variable continue :
$\Delta P = \frac{1}{2} P_N(m).\Delta m = \frac{1}{\sqrt{2.\pi}.\sigma}.e^{-\frac{x^2}{2.\sigma^2}}.\Delta m$
devient : $\mathrm{d}P = p(x).\mathrm{d}x$
avec : $\boxed{p(x) = \frac{1}{\sqrt{2.\pi}.\sigma}.e^{-\frac{x^2}{2.\sigma^2}}}$

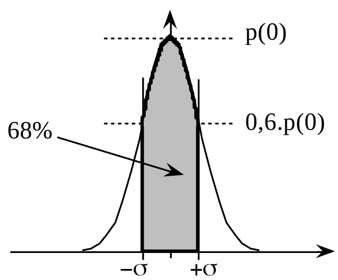


On a : $p(0) = \frac{1}{\sqrt{2.\pi}.\sigma}$ et $p(\sigma) = p(0).e^{-1/2} \approx 0,6.p(0)$

— **Probabilité que** $x \in [-\sigma, +\sigma]$
$P = \int_{-\sigma}^{+\sigma} p(x).\mathrm{d}x = \frac{1}{\sqrt{2.\pi}.\sigma}.\int_{-\sigma}^{+\sigma} e^{-\frac{x^2}{2.\sigma^2}}.\mathrm{d}x$
On pose : $t = \frac{x}{\sqrt{2}.\sigma}$
$\Rightarrow P = \frac{1}{\sqrt{\pi}}.\int_{-1/\sqrt{2}}^{+1/\sqrt{2}} e^{-t^2}.\mathrm{d}t = \frac{2}{\sqrt{\pi}}.\int_0^{+1/\sqrt{2}} e^{-t^2}.\mathrm{d}t$

On appelle $\theta(u) = \frac{2}{\sqrt{\pi}}.\int_0^u e^{-t^2}.\mathrm{d}t$ la fonction d'erreur (fonction tabulée dans la plupart des ouvrages et qui existe sur certaines calculatrices : UTPN sur "Hewlett-Packard").
D'où : $P = \theta(u = \frac{\sqrt{2}}{2}) = \boxed{68\%}$

1.7 Conclusion



On adopte σ comme valeur de l'incertitude absolue :
il y a $\boxed{68\%}$ de chances que $X \in [\overline{x} - \sigma; \overline{x} + \sigma]$
il y a $\boxed{95.4\%}$ de chances que $X \in [\overline{x} - 2.\sigma; \overline{x} + 2.\sigma]$

et $\boxed{99.7\%}$ de chances que $X \in [\overline{x} - 3.\sigma; \overline{x} + 3.\sigma]$

Remarque : si $\overline{x} \neq 0$, alors : $p(x) = \dfrac{1}{\sqrt{2.\pi}.\sigma}.e^{-\frac{(x-\overline{x})^2}{2.\sigma^2}}$

2 Question

1. Simuler 100 000 marches au hasard de $N = 2500$ pas de ± 1.
2. Calculer la proportion de mesures dans l'intervalle $[\overline{x} - \sigma; \overline{x} + \sigma]$
3. Calculer la proportion de mesures dans l'intervalle $[\overline{x} - 2.\sigma; \overline{x} + 2.\sigma]$
4. Calculer la proportion de mesures dans l'intervalle $[\overline{x} - 3.\sigma; \overline{x} + 3.\sigma]$