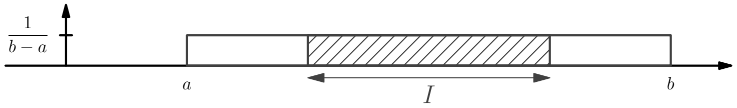
Kit de survie - Quatrième partie

11. Lois de probabilité

a) Loi uniforme

Définition

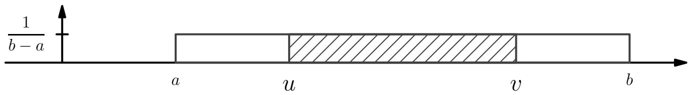
On dit qu'une variable aléatoire X suit la **loi uniforme** sur $[a\,;\,b]$ lorsque pour tout intervalle I, inclus dans $[a\,;\,b]$, la probabilité de l'événement « X appartient à I » est égale à l'aire du rectangle de base I et de hauteur $\dfrac{1}{b-a}$.



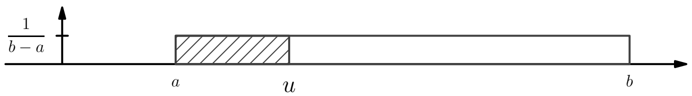
Propriété

Si une variable aléatoire X suit la **loi uniforme** sur $[a\,;\,b]$ alors pour tous réels u et v inclus dans $[a\,;\,b]$, on a :

- $p\,(u \leqslant X \leqslant v) = \dfrac{v-u}{b-a}$

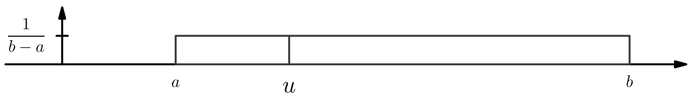


- $p\,(X \leqslant u) = p\,(a \leqslant X \leqslant u) = \dfrac{u-a}{b-a}$



- $p\,(X \geqslant v) = p\,(v \leqslant X \leqslant b) = \dfrac{b-v}{b-a}$

- $p\,(X = u) = 0$



(on a les mêmes résultats avec des inégalités strictes)

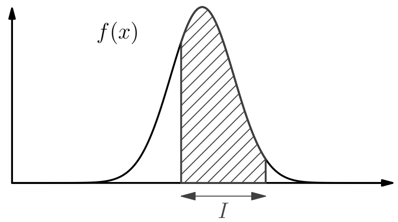
Propriété

Si une variable aléatoire X suit la **loi uniforme** sur $[a\,;\,b]$ alors **l'espérance** de X est égale à $\dfrac{a+b}{2}$.

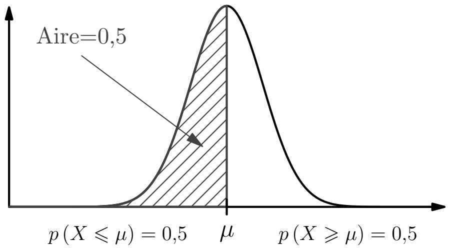
b) Loi normale

Définition

On dit qu'une variable aléatoire X suit la **loi normale d'espérance μ et d'écart-type σ** lorsque pour tout intervalle I la probabilité de l'événement « X appartient à I » est égale à l'aire sous la courbe sur I de la fonction f définie par $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-0,5\left(\frac{x-\mu}{\sigma}\right)^2}$



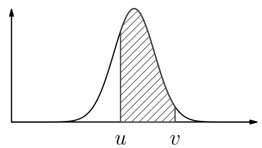
► **Remarque :** L'aire totale sous la courbe est égale à 1 (on dit que f est une densité de probabilité) et la courbe est symétrique par rapport à l'espérance μ. On a donc la situation suivante :



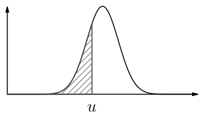
Propriété

- Si une variable aléatoire X suit la **loi normale d'espérance μ et d'écart-type σ** alors pour tous réels α et β, on a :

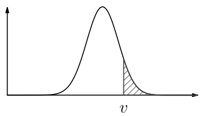
$p\,(u \leqslant X \leqslant v) =$
TI : DISTR (2nd+VARS) ; `normalcdf` (u,v,μ,σ)
CASIO : Menu STAT ; DIST ; NORM ; NCD avec
Lower : u ; Upper : v ; σ : σ ; μ : μ



$p\,(X \leqslant u) =$
TI : `normalcdf` $(-10^{99}, u, \mu, \sigma)$
CASIO : NCD avec
Lower : -10^{99} ; Upper : u ; σ : σ ; μ : μ



$p\,(X \geqslant v) =$
TI : `normalcdf` $(v, 10^{99}, \mu, \sigma)$
CASIO : NCD avec
Lower : v ; Upper : 10^{99} ; σ : σ ; μ : μ



• **Valeurs remarquables :**
$p\,(\mu - \sigma < X < \mu + \sigma) = 0{,}68$; $p\,(\mu - 2\sigma < X < \mu + 2\sigma) = 0{,}95$
$p\,(\mu - 3\sigma < X < \mu + 3\sigma) = 0{,}997$

► *Exemple 1:* (pour tester sa calculatrice)
Si X suit la loi normale d'espérance $\mu = 58$ et d'écart-type $\sigma = 6$, on doit avoir :
$p\,(52 \leqslant X \leqslant 64) \approx 0{,}682689$; $p\,(X \leqslant 55) \approx 0{,}308538$; $p\,(X \geqslant 62) \approx 0{,}252493$

► *Exemple 2:* Le diamètre X des barres métalliques sortant d'un atelier suit la loi normale d'espérance 12 mm (le diamètre attendu) et d'écart-type 0,08 mm. Un client refuse d'acheter des tubes dont le diamètre ne serait pas compris entre 11,9 mm et 12,2 mm. On cherche à déterminer le pourcentage de tubes acceptés par le client.
$p\,(11{,}9 \leqslant X \leqslant 12{,}2) \approx 0{,}888$, donc 88,8% des tubes sont acceptés par le client.

► *Exemple 3:* Une variable aléatoire suivant une loi normale est telle que $p\,(X < 2) = 0{,}067$ et $p\,(X < 3) = 0{,}159$. On peut en déduire que $p\,(X > 2) = 1 - p\,(X \leqslant 2) = 0{,}933$ et $p\,(2 < X < 3) = p\,(X < 3) - p\,(X < 2) = 0{,}092$.

12. Échantillonnage

a) Intervalle de fluctuation à 95%

Propriété

Étant donné une population dans laquelle la proportion connue d'un certain caractère est p. Si on prélève, avec remise, un échantillon de taille n dans cette population alors il y a 95% de chance (dans certaines conditions) que la proportion f du caractère au sein de cet échantillon appartienne à l'intervalle :

$$\left[p - 1{,}96\sqrt{\frac{p(1-p)}{n}}\ ;\ p + 1{,}96\sqrt{\frac{p(1-p)}{n}}\right]$$

Cet intervalle est appelé **intervalle de fluctuation à 95%** de l'échantillon associé à la proportion p.

b) Prise de décision à partir d'un intervalle de fluctuation

Propriété

Étant donné une population dans laquelle **on suppose que la proportion d'un certain caractère est** p. Si on prélève, avec remise, un échantillon de taille n dans cette population et si la **fréquence réelle observée** f du caractère dans cet échantillon **est comprise dans l'intervalle de fluctuation** alors on dit qu'**on accepte au seuil de 95% l'hypothèse que la proportion réelle du caractère dans la population est bien** p (dans le cas contraire, on dit qu'on rejette l'hypothèse).

► *Exemple :* Un candidat pense que 52% des électeurs lui sont favorables. On prélève avec remise un échantillon de 500 électeurs : 47% des électeurs interrogés de cet échantillon se déclarent favorable au candidat en question.
L'intervalle de fluctuation de l'échantillon associé à la proportion de 52% est $[0{,}476\,;\,0{,}564]$ car $0{,}52 - 1{,}96\sqrt{\dfrac{0{,}52 \times 0{,}48}{500}} \approx 0{,}476$ et $0{,}52 + 1{,}96\sqrt{\dfrac{0{,}52 \times 0{,}48}{500}} \approx 0{,}564$.
0,47 étant en dehors de l'intervalle de fluctuation, on peut rejeter au seuil de 95% l'hypothèse du candidat selon laquelle 52% des électeurs lui sont favorables.

c) Estimation par un intervalle de confiance

Propriété

On cherche à connaitre une estimation de la proportion p inconnue d'un certain caractère au sein d'une population. Pour cela, on prélève avec remise un échantillon de taille n au sein de la population et on note f la proportion observée du caractère au sein de l'échantillon. Il y alors 95% de chance (dans certaines conditions) que la proportion p du caractère au sein de la population totale soit comprise dans l'intervalle :

$$\left[f - \frac{1}{\sqrt{n}}\ ;\ f + \frac{1}{\sqrt{n}}\right]$$

Cet intervalle est appelé **intervalle de confiance à 95%** associé à la proportion f.

► *Exemple :* Un sondage réalisé sur un échantillon de 1000 personnes attribue à un candidat un score de 18%. L'intervalle de confiance à 95% associé à cette proportion observée de 18% dans l'échantillon est $[14{,}8\%\,;\,21{,}2\%]$ car $0{,}18 - \dfrac{1}{\sqrt{1000}} \approx 0{,}148$ et $0{,}18 + \dfrac{1}{\sqrt{1000}} \approx 0{,}212$.