Probabilité et statistique

Table des matières

Rappels

Nous avons défini :

Une expérience aléatoire Une expérience (souvent un tirage) dont seul le hasard détermine le résultat.
Les issues: tous les résultats possibles de l'expérience aléatoire
Issues équiprobables**:** dont aucune n'est favorisée ou défavorisée par le protocole de l'expérience aléatoire.
L'Univers Ω Ensemble des issues
Une tribu famille de sous-ensembles de Ω comprenant leur réunion, leur complémentaire, Ω, ∅
Les évènements Sous-ensembles appartenant à la tribu
Evènements incompatibles évènements dont l'intersection et nulle
Evènements composites: A ou B (A U B), A et B ($A \cap B$), non A (Ω – A ou $\bar{A}$)
Evènement mesurable: ensemble dont on peut évaluer la longueur, l'aire, le volume, le nombre d'éléments…
Probabilité d'un évènement rapport de la mesure de ce sous-ensemble à la mesure de Ω

On a vu quelques propriétés élémentaires de la probabilité

P(Ω)=1 P(∅) = 0
P(A ou B) = P(A U B) = P(A) + P(B) – P(A ∩ B) P(A ou B) = (P(A) + P(B) pour des évènements incompatibles
P(non A) = 1 – P(A)

On a étudié quelques techniques de dénombrement en vue d'étudier les expériences aléatoires complexes formées de plusieurs épreuves

■ Nous avons appris à compter le nombre d'éléments d'un ensemble produit qui est le nombre d'issues possibles quand l'univers est composé de n-Uplets dont chaque composante appartient à un ensemble distinct
Par exemple une issue X =(a,b,c) avec a∈E , b∈F, c∈G. On dit que X∈ ExFxG qui est un ensemble produit.
Ou quand on réitère une expérience avec remise dans le même ensemble X =(a,b,c) avec X∈E^3 (Ensemble ExExE)
Le nombre d'éléments d'un ensemble produit est égal au produit des nombres d'éléments des ensembles qui le composent.
■ Le nombre de sous-ensembles de p éléments qu'on peut faire avec un ensemble qui en contient n qu'on peut aussi définir comme le nombre de combinaisons possibles de n objets p à p est $\binom{n}{p}$.
■ Le nombre de permutations possibles qu'on peut faire avec p objets en les ordonnant différemment p! =p(p-1)..(2)1
Et le nombre d'arrangements possibles (ensembles ordonnés) que l'on peut faire avec n objets pris p à p A_n^p

On a vu comment calculer la probabilité d'une issue dans une expérience complexe

Si nécessaire en construisant un arbre pondéré recensant les issues possibles de chaque épreuve, chaque branche étant pondérée par la probabilité de l'issue correspondante, la probabilité du résultat d'une branche étant égale au produit des probabilités qui la composent.
Après définition d'une variable aléatoire …
On a plus particulièrement étudié le cas d'un schéma de Bernoulli où on prélève, dans une population un échantillon d'effectif n. Dans cet échantillon, on compte combien de fois on trouve un caractère donné, ce nombre k pouvant varier de 0 à n, et, sachant que la probabilité de rencontrer le caractère chez un individu est p, on s'intéresse à la probabilité pour que k prenne une valeur donnée.

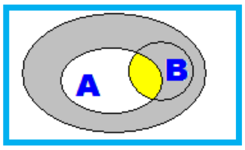
Cette expérience dont la loi de probabilités porte le nom de loi binomiale constitue un lien entre probabilités et statistiques.

Dans le domaine statistique on a étudié

Les notions de classe, d'effectifs, de fréquences, d'effectifs ou de fréquence cumulés.
Les caractéristiques de tendance centrale (moyenne, médiane) et de dispersion (quartiles, déciles, écart type, variance) d'une série statistique.
Dans le domaine de l'échantillonnage, on a défini l'intervalle de fluctuation d'une estimation au seuil de 95%, en utilisant notamment les chiffres fournis par la loi binomiale. Ce procédé consiste à déterminer dans quel intervalle doit se trouver une fréquence mesurée dans l'échantillon pour qu'on la juge acceptable au risque de 5%.
Si on connait la fréquence réelle dans la population où est prélevé l'échantillon cela peut servir à vérifier la validité de l'échantillonnage ou à évaluer une fréquence inconnue dans la population.

Probabilité de B sachant que A est réalisé

On note $P_A(B)$ La **probabilité de B sachant que A est réalisé**.



Quand A est réalisé l'univers des possibles se réduit à A.
Sachant que l'évènement A est réalisé, la probabilité que B se réalise est égale à la proportion d'issues favorables à B que contient A. Or ces issues sont celles de $A \cap B$.

Donc $P_A(B)$ n'est pas autre chose que la fréquence de $A \cap B$ dans A.

Si on note nb (E) le nombre d'issues ou la mesure d'un évènement E:
L'univers se réduisant à A on a

$$P_A(B) = \frac{nb\,(A \cap B)}{nb\,(A)} = \frac{nb(A \cap B)}{nb(\Omega)} : \frac{nb(A)}{nb(\Omega)} = \frac{P(A \text{ et } B)}{P(A)} \text{ (mesurées dans } \Omega)$$

et on déduit que

$$\boxed{P(A \text{ et } B) \text{ ou } P(A \cap B) = P(A).P_A(B)}$$

On voit aussi que $P_A(\overline{B}) = 1 - P_A(B)$

Car A constitue l'univers et dans A, B et $\overline{B}$ sont complémentaires.
Si maintenant on se place dans un contexte où c'est B qui est réalisé, on a

$$P_B(A) = \frac{P(A \text{ et } B)}{P(B)}\;) \quad \text{et} \quad P(A).P_A(B) = P(A \cap B) = P(B).P_B(A)$$

Derrière ces formules complexes se cache une réalité toute simple:
Supposons que la fréquence de A dans Ω soit 60% P(A) et que la fréquence de A∩B dans A soit 40% $P_A(B)$ quelle est la fréquence de A∩B dans Ω (soit P(A et B))?
Si A représente 60% de l'univers (P(A)=0,6) et A∩B représente 40% de A, ($P_A(B)$ = 0,4) alors
A∩B représente 40% de 60% de l'univers soit P(A∩B) = 40%x60% = 24% de l'univers =0,24.

Evènements indépendants.

$$\boxed{\text{A et B sont indépendants si et seulement si } P_A(B)=P(B). \text{ On a alors } P(A \text{ et } B) = P(A).P(B)}$$

En effet:

$P_A(B)=P(B)$ signifie que la fréquence de B est la même dans A que dans Ω.
Donc cette fréquence est aussi identique dans A et dans $\overline{A}$. $P_{\overline{A}}(B) = P(B)$.
Autrement dit la propriété A n'a aucune influence sur la propriété B. Elles sont indépendantes.
Et comme **$P(A \text{ et } B) = P(A).P_A(B)$** (cas général) **si $P_A(B)=P(B)$** (évènements indépendants) **$P(A \text{ et } B) = P(A).P(B)$**
Par exemple si la proportion de celles qui ont fait des études secondaires parmi les filles est la même que la proportion de ceux qui ont fait des études secondaires dans l'ensemble de la population (et donc la même que chez les garçons), cela signifie que l'évènement "avoir fait des études secondaires" est indépendant du sexe.
Si dans un groupe de 100 personnes il y a 60% de filles F et que la probabilité qu'un individu ait fait des études secondaires E est 80%. Quand on tire une personne au hasard dans ce groupe, la probabilité pour que ce soit une fille qui ait fait des études secondaires est P(F et E) = P(F).P(E) = 60% x 80% soit 48%. nb(F∩E) = 48% de l'effectif.

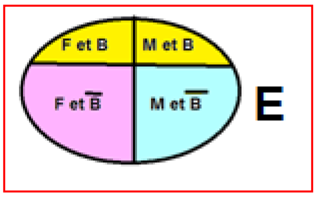
On a donc

Evènements compatibles **P(A ou B) = P(A U B)** **= P(A)+ P(B) – P(A et B)**	**Evènements incompatibles (disjoints)** **P(A ou B) = P(A U B)** **= P(A)+ P(B)**
Cas général **P(A et B) ou P(A ∩ B)** **= $P(A).P_A(B)$**	**Evènements indépendants** **P(A et B) ou P(A ∩ B)** **= P(A).P(B)**

Application : partition d'un ensemble selon des caractères croisés

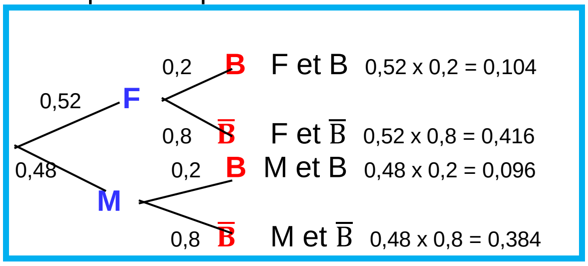
Caractères indépendants

Supposons qu'on s'intéresse au sexe et à la couleur des cheveux dans une population importante E.
Ces deux caractères sont indépendants.
Si la proportion de blonds dans la population est 20%, cette proportion doit être la même pour les individus de sexe masculin et pour les individus de sexe féminin.
Si on note F et M le sexe et la couleur de cheveux B pour blond et $\overline{B}$ pour non – blond, la population peut être scindée en 4 sous ensembles : F et B, F et $\overline{B}$, M et B , M et $\overline{B}$.



<u>Définition:</u> Ces 4 sous-ensembles forment ce qu'on appelle **une partition** de E, ce qui signifie que leur intersection 2 à 2 est nulle (ils ne se chevauchent pas) et leur réunion est égale à E.

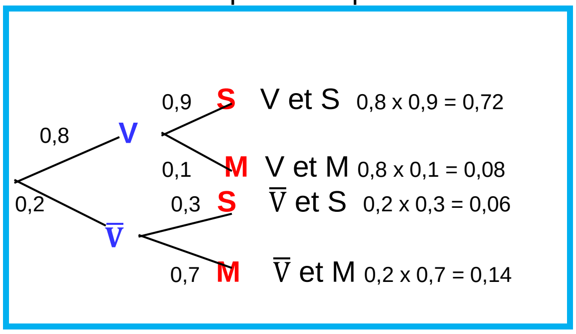
En supposant que la population comporte 52% de femmes et 48% d'hommes, on peut illustrer le mécanisme de la partition par l'arbre suivant:



Les 2 premières branches représentent la dissociation de E selon le sexe en 2 sous-ensembles de fréquence respective 0,52 et 0,48.
Les branches suivantes représentent la dissociation de chaque sexe en 2 sous ensemble B et $\overline{B}$ de fréquence respective 0, 2 et 0,8.
Au total on obtient une partition en 4 sous-ensembles dont la fréquence est donnée par le produit des fréquences qui pondèrent les branches menant au sous ensemble de la partition.
Par exemple la fréquence de {**F et B**} est donnée par le produit 0,52 par 0,2.

Caractères liés

Supposons maintenant que l'on distribue dans cette population un vaccin contre une maladie. Au bout d'un certain temps la population va être scindée selon 2 caractères:
L'état de santé: (malade (M) / sain (S)) et le statut par rapport au vaccin (vacciné (V) / non vacciné ($\overline{V}$)).
Les caractères composites terminaux de cette partition seront V et M, V et S, $\overline{V}$ et M, $\overline{V}$ et S.
Mais ces caractères V et M n'étant pas indépendants, la proportion de malades devrait être plus importante chez les non vaccinés que chez les non vaccinés.
L'arbre donnera par exemple:

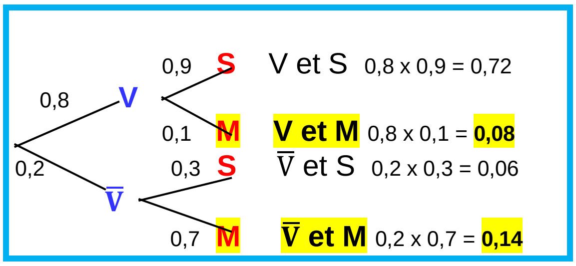


Dans tous les cas
1) la somme des fréquences de chaque fourche doit être 1 (ou 100%)
→ 0,8 + 0,2 = 1 → 0,9 + 0,1 = 1 → 0,3+ +0,7 = 1
2) la somme des fréquences des caractères terminaux portant le caractère V (ou $\overline{V}$) doit être égale à la fréquence de V (ou $\overline{V}$) → 0,72 + 0,08 = 0,8.
3) la somme des fréquences des caractères terminaux portant le caractère S (ou M) est égale à la fréquence de S (ou M) dans la population
→ fréquence de S = 0,72 + 0,06 = 0,78 → 78% de la population est saine.

Probabilités conditionnelles

On peut nous demander par exemple sachant qu'une personne tirée au hasard dans E est malade quelle est la probabilité pour qu'elle soit vaccinée. C'est ce qu'on appelle la probabilité des causes découverte par un mathématicien appelé Bayes.

Dans ce cas on fait un zoom sur le sous ensemble malade (M) .



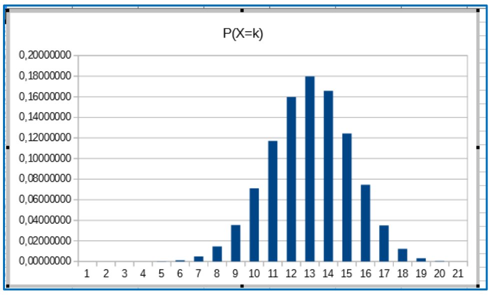
Sa fréquence dans la population est 0,08 + 0,14 = 0,22 soit 22% .
Tandis que la fréquence des vaccinés et malade est 0,08 soit 8%.
Donc la fréquence des vaccinés parmi les malades est

$$\frac{0,08}{0,08+0,14} = \frac{8}{22} = 0,36$$

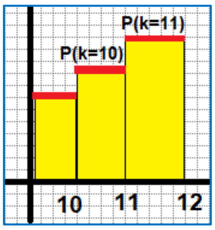
La probabilité de vacciné sachant malade est 0,36 ou 36%.

Ce type de calcul de probabilité est pertinent que les caractères soient dépendants ou indépendants.
Ici les caractères sont déclinés en 2 modalités. On peut imaginer des caractères déclinés en 3 modalités ou plus.
Le nombre de fourches de l'arbre varie mais le principe est le même.

Du discret au continu

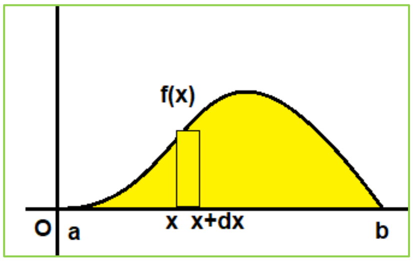


Ce diagramme illustre la répartition de la probabilité dans un schéma de Bernoulli respectant la loi B(20;0,6).
Autrement dit on tire au hasard un échantillon d'effectif 20, et dans cet échantillon k est le nombre d'individus possédant la caractéristique recherchée dont la probabilité, dans l'univers est 0,6.
La hauteur de chaque barre du diagramme est la probabilité pour k = x, (x variant de 0 à 20).
Ici k prend des valeurs entières (0,1,2,….,20) . C'est une variable discrète.
Si on fait la somme des hauteurs des 21 barres on trouve la probabilité totale $\sum P(K = x)$ = 1.

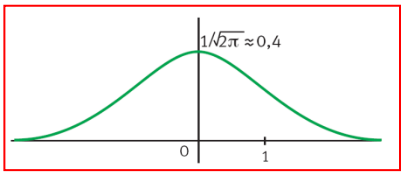


En supposant que P(K=10) = 0,07 et que P(K=11) = 0,18 par exemple, rien ne nous empêche de dire que P(K ∈ [10,11[) = 0,07 , P(K∈ [11; 12[= 0,18 et de procéder ainsi pour tout k ∈[0;21[.
Pour montrer que la probabilité est constante sur un intervalle [n ;n+1[on va utiliser un histogramme.
La largeur de chaque rectangle sera 1, sa hauteur sera P(K=Partie entière de x) ou P(K=PE(x)).
La surface des rectangles sera donc P(K=0) , P(K=1), P(K=2) …. jusqu'à P(K=20) et la surface totale des rectangles sera égale à 1 qui est la probabilité totale.
La fonction en escalier qui coiffe les rectangles (en rouge sur le dessin) est f(x)=P(k=PE(x)).
Et on peut dire que $\int_0^{21} f(x)dx = 1$ (en unité d'aires)

La densité est le rapport entre une quantité d'une chose et un volume, une surface, une longueur, ... mais contrairement aux quantités volumiques, surfaciques ou linéiques, la densité est une donnée locale. Elle indique la valeur du rapport en cet endroit précis mais elle ne garantit pas que ce rapport sera stable quand l'endroit variera. De ce point de vue la fonction f(x) de cet exemple peut être considérée comme **une densité de probabilité** sur un intervalle puisque c'est le rapport entre la probabilité sur un intervalle et la longueur de cet intervalle (qui ici est 1).

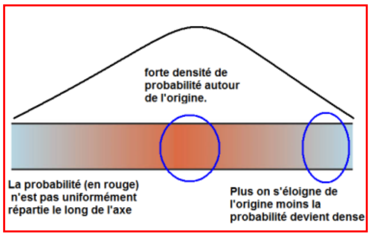


Supposons maintenant qu'on passe d'une variable discrète k à une variable continue X pouvant prendre toutes les valeurs d'un intervalle [a,b].
On sait que la probabilité pour que X prenne une valeur particulière est nulle P(X=x) = 0 mais par contre P(X∈[x_1, x_2]) a un sens. Notamment P(a≤ X≤ b) = 1.
Sur un intervalle infinitésimal I= [x,x+dx] on peut représenter la probabilité P(X∈I) comme l'aire d'un élément d'aire de largeur dx et de hauteur f(x).
f(x) = P(x ≤ X ≤ x+dx) / dx est donc le rapport entre la probabilité sur un intervalle et la longueur de cet intervalle. f(x) est une **densité de probabilité.** C'est une fonction continue et positive de x. Et on a bien
P(a ≤ X≤ b)=$\int_a^b f(x)dx = 1$ et P(x_1 ≤ X ≤ x_2)=$\int_{x_1}^{x_2} f(x)dx$ (aires en unité d'aires).



Voici un exemple ou X peut prendre toutes les valeurs de –∞ à + ∞ et où la densité de probabilité est la fonction **f(x)=** $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$.
C'est une fonction paire et on démontre que

$$\lim_{a\to+\infty} \int_{-a}^{+a} f(x)dx = 1$$



Comme toujours quand f(x) est une densité de probabilité on a
P(x_1 ≤ X ≤ x_2)=$\int_{x_1}^{x_2} f(x)dx$
L'allure de la courbe nous renseigne sur la répartition de la probabilité presque toute concentrée autour de X = 0.
C'est tellement vrai que $\int_{-1,96}^{+1,96} \mathrm{f(x)dx} = 0{,}95$. Autrement dit X a 95% de chances de se trouver entre -1,96 et +1,96.

f(x) définie sur un intervalle I est une **densité de probabilité** si
f(x) est continue sur I
f(x) est positive sur I
L'aire sous la courbe $\int_I f(x)dx = 1\ U.A$

Exemples de densités de probabilité:
f(t) = –2t + 2 sur [0,1] En effet f(t) est positive et continue sur [0;1] et $\int_0^1 f(t)dt = [-t^2 + 2t]_0^1$ = – 1 + 2 = 1
f(t)= $1/t^2$ sur [1; +∞[En effet f(t) est positive et continue sur [0;1] et $lim_{x\to+\infty} \int_1^x \frac{1}{t^2}dt$ = $lim_{x\to+\infty}$ $[-\frac{1}{t}]_1^x$ = 1

Lois de probabilités à densité

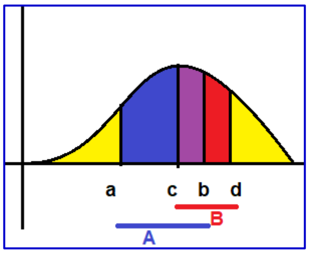
Quand le résultat d'une expérience aléatoire est une variable aléatoire continue X,
Ω est un intervalle de ℝ (ou ℝ lui-même)
Et les évènements sont des intervalles de Ω.
Quand la variable X suit une loi à densité f(x) sur Ω,
la probabilité d'un évènement [a ; b] est la probabilité P(a ≤ X ≤ b) donnée par $\int_a^b f(x)dx$ en unités d'aires

Il est équivalent d'écrire P(X∈ [a;b]) ou P(a≤ X≤ b) ou P([a;b]) puisque les évènements sont des intervalles de Ω .

On retrouve les propriétés caractéristiques de la probabilité

On a P(Ω) = 1 (aire totale située entre le graphe de f(x) et l'axe des x sur Ω)
La probabilité d'un évènement [a;b] est bien le rapport entre sa mesure $\int_a^b f(x)dx$ et la mesure de Ω (= 1)
P(X =a) = 0 Car la mesure d'un point est nulle ($\int_a^a f(x)dx = 0$)
Donc exclure ou inclure les bornes de l'évènement est sans influence P([a;b]) = P(]a;b]) = P([a;b[) = P(]a;b[)
Si [a;b] et [c,d] sont des intervalles disjoints on a P(X ∈ [a;b] U [c;d]) = $\int_a^b f(x)dx + \int_c^d f(x)dx$
Si $\overline{[a\,;b]}$ est le complémentaire de [a,b] on a P(X ∈$\overline{[a\,;b]}$) = 1 – $\int_a^b f(x)dx$

Probabilité de B sachant A



En supposant que l'intersection entre A et B existe.

P(X∈[c;d]) sachant que X∈ [a;b] = $\frac{P(X\in[a;b]\cap[c;d])}{P(X\in[a;b])} = \frac{P([c;b])}{P([a;b])}$

Sur notre figure $P_{X\in[a;b]}$(X ∈[c;d]) ou $P_{[AB]}$([c;d]) est le rapport de l'aire en violet à la somme des aires en violet et en rouge.

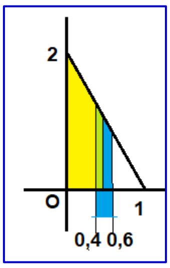
Espérance mathématique de X

Pour une variable aléatoire discrète on a E(X) = $\sum XiP(X = Xi)$.
Par analogie, dans le cas où X est continue et suit une loi à densité f(x)

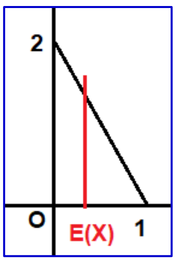
L'espérance mathématique de X sur [a;b] est E(X)=$\int_a^b xf(x)dx$

Méthodes

Soit X de densité f(x) = –2x + 2 sur [0 ; 1]



Déterminer P([0,4 ; 0;6]) quand [0 ; 0,5] est réalisé
1) je cherche l'intersection de [0,4; 0,6] avec [0;0,5] je trouve [0,4 ; 0,5]
2) je cherche la probabilité de l'intersection $\int_{0,4}^{0,5}(-2x+2)dx$ = $[-x^2+2x]_{0,4}^{0,5}$ = 0,75 – 0,64 = 0,11
3) je cherche la probabilité de [0 ; 0,5] = $[-x^2+2x]_0^{0,5}$ = 0,75.
4) la probabilité cherchée [P([0,4 ; 0;6]) quand [0 ; 0,5] est réalisé] est égale au quotient 0,11 / 0,75 = 0,147



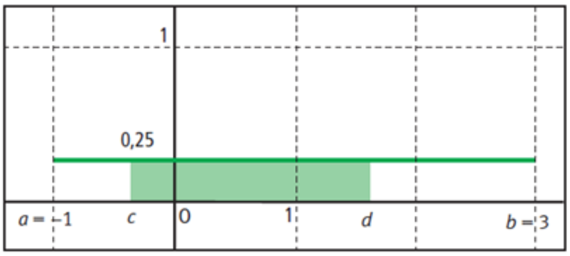
Déterminer E(X) sur [0;1]
E(X) = $\int_0^1 x(-2x+2)dx = \int_0^1(-2x^2+2x)dx$ = $[-\frac{2}{3}x^3+x^2]_0^1$ = $\frac{1}{3}$
Ce qui traduit une densité de probabilité plus importante près du 0 que près du 1.

Loi uniforme

On parle de loi uniforme quand la densité de probabilité est une constante K sur l'intervalle [a;b] caractérisant Ω .
Si $\int_a^b K\,dt = 1$ c'est que Kb – Ka = 1 et donc que K = $\frac{1}{b-a}$.

La densité de probabilité d'une loi uniforme sur [a;b] est donc **f(x)=** $\boldsymbol{\frac{1}{b-a}}$.

Et il en découle que P[c,d] = $\frac{d-c}{b-a}$ quotient de l'amplitude de l'évènement par l'amplitude de l'univers.



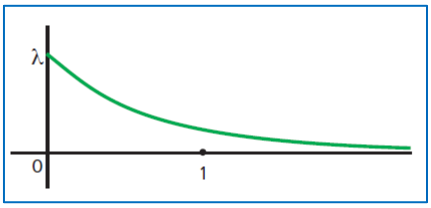
Exemple de loi uniforme:
On pique un point au hasard sur l'intervalle [–1 ; +3]. Quelle est la probabilité de piquer un point de l'intervalle [– 0,2 ; 2,5].
Il est évident que la densité de probabilité doit être la même partout. Il n'est pas plus probable de piquer un point à un endroit qu'à un autre.

En faisant le rapport mesure de l'intervalle évènement sur mesure de l'univers P[–0,2 ; 2,5] = 2,7 / 4 = 0,675.
En utilisant la loi à densité uniforme P[–0,2 ; 2,5] = $\int_{-0,2}^{2,5} 0{,}25dt$ = $[0{,}25t]_{-0,2}^{2,5}$ = 0,625 +0,5 = 0,675.

E(X) = $\int_a^b \frac{x}{b-a}\,dx = \left[\frac{x^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}$

L'espérance mathématique de la loi uniforme sur [a,b] est la moyenne des bornes de l'intervalle $\frac{a+b}{2}$

Loi exponentielle



Une variable aléatoire suit une loi exponentielle de paramètre λ sur [0 ; +∞[si sa densité de probabilité est f(x)=$\lambda e^{-\lambda x}$.

Probabilité de X

P([a;b])=$e^{-\lambda a} - e^{-\lambda b}$
P(X≤ a)=P([0;a]) = 1 – $e^{-\lambda a}$
P(X≥ a)= $e^{-\lambda a}$

Espérance mathématique de X

E(X) =lim $_{x\to+\infty}\int_0^x \lambda t e^{-\lambda t}dt = \frac{1}{\lambda}$

Probabilité de X ≥ t+ h quand X ≥ t est réalisé .

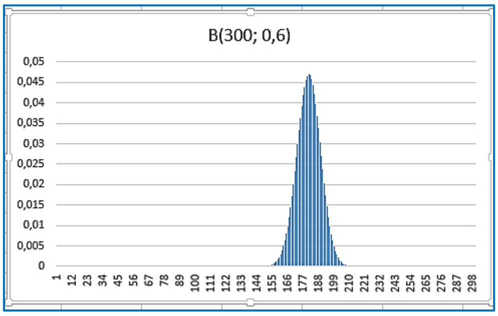
L'intersection des 2 intervalles est X ≥ t + h de probabilité $e^{-\lambda(t+h)}$.
Tandis que P(X≥ t) = $e^{-\lambda t}$.
La probabilité cherchée est le quotient P(X≥ t+h) / P(X ≥ t) = $e^{-\lambda h}$

Quand X suit une loi exponentielle de paramètre λ , **P(X ≥ t+h) quand X ≥ t = P(X ≥ h)**

La loi exponentielle servant souvent à mesurer la probabilité de dépasser une durée de vie, cette formule exprime que quand la durée de vie a atteint le seuil t, la probabilité qu'elle dépasse ce seuil de h est égale à la probabilité de dépasser h au début de sa vie. Théorème de la durée de vie sans vieillissement.

Loi normale centrée réduite N(0 , 1)

De la loi binomiale à la loi normale centrée réduite



Soit une variable X suivant une loi binomiale B(n,p).
X comptabilise les succès d'une expérience de probabilité p renouvelée n fois.
X peut prendre les valeurs 0,1,….,n. La loi P(X=x) est connue.
L'espérance ou moyenne de X est m = E(X) = np
Et son écart type est $\sigma = \sqrt{np(1-p)}$
Le Diagramme à barre représentant la probabilité P(X=x) en fonction de x pour B(300,0,6) a l'allure ci-contre: maximum très prononcé pour X=np=180 et décroissance assez rapide de la probabilité autour de ce maximum.

Faisons maintenant **le changement de variable Z = $\frac{X-m}{\sigma} = \frac{X-np}{\sqrt{np(1-p)}}$** . La probabilité de Z = $\frac{k-m}{\sigma}$ est P(X=k).

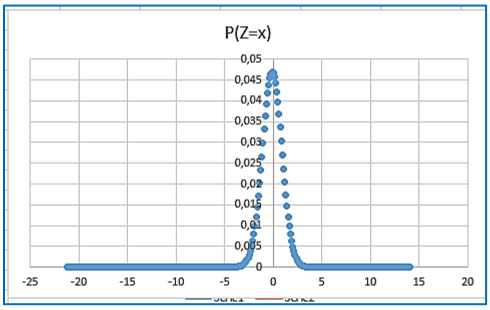
Comme **le maximum de probabilité** se produisait pour X =m, il va maintenant se produire pour Z =$\frac{m-m}{\sigma}$ = 0.

La moyenne de Z est $\frac{np-np}{\sqrt{np(1-p)}} = 0$ et **son écart type** est $\frac{\sqrt{np(1-p)}}{\sqrt{np(1-p)}} = 1$ (obtenus à partir des transformations subies par X)

Z = $\frac{X-m}{\sigma}$ est ce qu'on appelle une variable binomiale centrée réduite

Les valeurs extrêmes de Z sont $\frac{-np}{\sqrt{np(1-p)}} = -\frac{\sqrt{np}}{\sqrt{1-p}}$ et $\frac{n-np}{\sqrt{np(1-p)}} = -\frac{\sqrt{n(1-p)}}{\sqrt{p}}$.
Elles vont tendre vers -∞ et +∞ quand n→+∞

L'écart entre 2 valeurs consécutives est $\frac{X+1-np}{\sqrt{np(1-p)}} - \frac{X-np}{\sqrt{np(1-p)}} = \frac{1}{\sqrt{np(1-p)}}$ Il va tendre vers 0 quand n → +∞ .

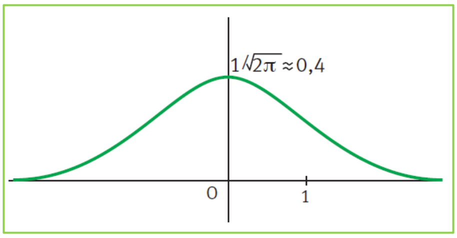


Pour n=300 et p=0,6 le graphe de P(Z=x) va avoir l'allure ci-contre.
Il est centré sur l'origine (m=0) et la réduction de l'écart type de $\sqrt{72}$ à 1 a pour effet de rapprocher les valeurs autour de la moyenne qui est Z =0.
D'après ce qu'on vient de voir, quand n va tendre vers +∞ l'amplitude du graphe va s'étendre sur ℝ tout entier et l'écart entre 2 valeurs consécutives va diminuer et tendre vers 0. En somme les discontinuités du graphe dues à ce que Z prend des valeurs discrètes vont tendre à disparaître et les probabilités du type P(Z=k) vont tendre vers 0 du fait que la probabilité totale (1) va être répartie entre un nombre de valeurs de Z qui va tendre vers l'infini.

Quand n → ∞ seule la valeur de P(a ≤ Z ≤ b) pour a et b réels a une signification.

Moivre et Laplace ont démontré que quand n → ∞, si Z est une variable binomiale centrée réduite
$\lim_{n \to \infty}$ P(a < Z < b) = $\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Loi normale centrée réduite



Une variable aléatoire X suit la loi normale centrée réduite $\mathcal{N}(0\,;1)$ si, pour tous réels a et b tels que $a < b$, on a :

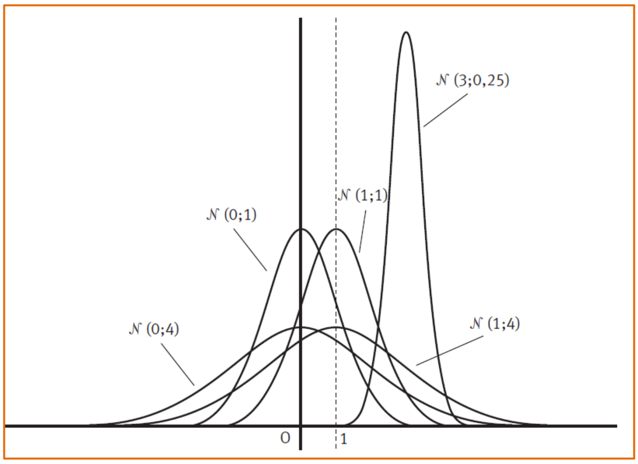
$$P(a \le X \le b) = \int_a^b f(x)\,dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx.$$

On voit ci-contre le graphe de cette densité de probabilité(fonction paire).

A savoir:

Si X suit une loi normale centrée réduite : **E(X) = 0** , **σ(X) = 1**
Il existe un unique nombre Up tel que P(– Up < X < Up) = 1 – p (p étant une probabilité)
$U_{0,05}$ = 1,96 signifie que P(–1,96 < X < +1,96) = 0,95 **$U_{0,01}$=2,58** signifie que P(–2,58 < X < +2,58) = 0.99
C'est donc que respectivement 95% et 99% de la probabilité sont concentrés ente ces valeurs de X.

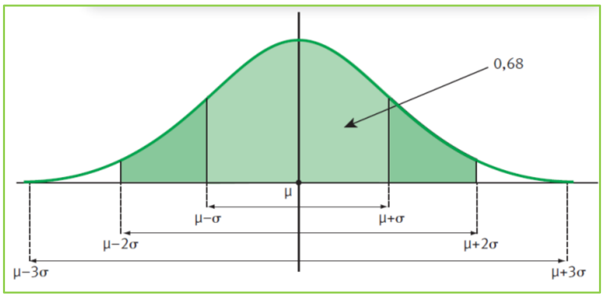
Autres lois normales N (m, σ^2)



Une variable aléatoire X suit une loi normale $\mathcal{N}(\mu\ ;\ \sigma^2)$ si la variable aléatoire $Z = \dfrac{X-\mu}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0\ ;\ 1)$.

Si une variable aléatoire X suit la loi N(m,σ^2)
E(X) = m et σ(X) = σ.

Ci-contre, les graphes de quelques densités de probabilité de ce type montrent que m est l'abscisse d'une densité de probabilité maximale et que σ détermine la dispersion de cette densité autour de m.



A savoir :

$P(\mu-\sigma \le X \le \mu+\sigma) \approx 0{,}68$ (à 10^{-2} près) ;

$P(\mu-2\sigma \le X \le \mu+2\sigma) \approx 0{,}95$ (à 10^{-2} près) ;

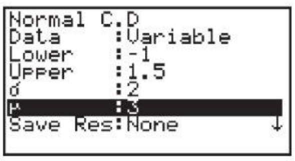
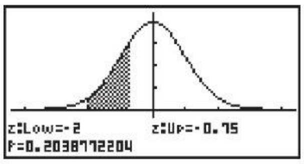
$P(\mu-3\sigma \le X \le \mu+3\sigma) \approx 0{,}997$ (à 10^{-3} près).

Résolution des problèmes

Si X suit la loi N(m, σ^2) , il faut savoir évaluer les probabilités
P(a< X <b) , P(X >a) , P(X<a) a et b étant des nombres réels donnés.
Inversement, il faut savoir trouver le réel a tel que P(X <a) = p , p étant une probabilité donnée
Les calculatrices scientifiques ou des tables permettent de trouver de telles valeurs.

Déterminer P(a ≤ X ≤ b) si X a pour moyenne μ et pour écart type σ selon la loi N(μ , σ^2)

Par exemple $P(-1 \le X \le 1{,}5) \approx 0{,}203877$ où X suit la loi $\mathcal{N}(3\ ;\ 4)$.

Texas	Casio Graph 35…
On choisit Distr (par 2nd Var) puis normalcdf (ou, en français, normalFRep). On indique les données dans l'ordre a, b, μ et σ (attention à ne pas confondre σ et σ^2). Voici un exemple avec $a=-1$ $b=1{,}5$ $\mu=3$ et $\sigma=2$: 	Dans le menu Stat, on choisit Distr, puis NormCD. On indique les données dans l'ordre a, b, σ et μ (attention à ne pas confondre σ et σ^2). Voici un exemple avec $a=-1$ $b=1{,}5$ $\mu=3$ et $\sigma=2$:  

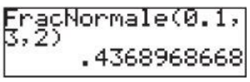
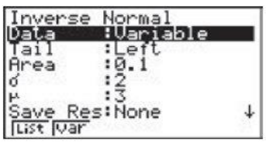
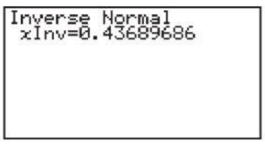
(Selon le site Académie en ligne)

Les modules de calcul adéquats sont
Distr puis normalFrep (TI) ou
Stat puis Distr puis Norm CD (Casio)
On passe à la calculatrice les paramètres a, b, moyenne écart type et elle nous donne le résultat.
Problème si on veut connaître
P(X ≤ 5)?
il faut donner à a la valeur –∞ et on ne sait pas faire.
Mais il suffit de prendre a = –100000 par exemple et on aura une valeur très voisine de la probabilité cherchée.

Déterminer x tel que P(X ≤ x) = p , p étant une probabilité donnée selon la loi N(μ , σ^2).

- Déterminer x tel que $P(X \le x) = p$, p étant une probabilité donnée.

La plupart des calculatrices permettent de trouver directement le résultat.

Texas	Casio graph 35 et plus
On choisit Distr (par 2nd Distr) puis invNorm (ou, en français, FracNorm), puis on donne p, μ et σ. Voici un exemple avec $p=0{,}1$ $\mu=3$ et $\sigma=2$: 	Dans le menu Stat, on choisit Distr, puis Inverse Normal. On indique les données dans l'ordre p, σ et μ . Voici un exemple avec $p=0{,}1$ $\mu=3$ et $\sigma=2$:  

On choisit le même module de calcul que précédemment
Distr (TI) ou Stat puis Distr (Casio) puis on indique à la calculatrice l'opération inverse de l'opération normale
soit par invNorm (TI)
 soit par Inverse Normal (Casio).
Enfin on donne les paramètres p, μ , σ.
Si on veut a et b tels que
P(m - a ≤ X ≤ m + b) ≤ 95% par exemple on cherche
P(X ≤x) ≤ 0,975 avec x = m+b
P(X ≤ y) ≤ 0,025 avec y = m-a
Dans la loi N(0,1) centrée a = b.

Intervalle de fluctuation

On s'intéresse maintenant non plus au nombre de succès X de la loi binomiale mais à leur fréquence $F=\frac{X}{n}$.

$F=\frac{k}{n}$ avec k entier de 0 à n	$P(F=\frac{k}{n})=P(X=k)$	$m=E(F)=p$	$\sigma(F)=\sqrt{\frac{P(1-p)}{n}}$	$Z=\frac{F-p}{\sigma}$ est une variable centrée réduite.

Quand $n \to \infty$ Z suit donc une loi N(0,1) et $F= Z\sigma + p$
On en déduit que les propositions suivantes sont équivalentes

$-U_\propto < Z < U_\propto$	$-U_\propto < \frac{F-p}{\sigma} < U_\propto$	$-U_\propto\sigma < F - p < U_\propto\sigma$	$-U_\propto\sigma + p < F < U_\propto\sigma + p$	$P(-U_\propto < Z < U_\propto) = P(p - U_\propto\sigma < F < p - U_\propto \sigma)$

Autrement dit:

Si F est la fréquence d'une modalité constatée dans un échantillon d'effectif n prélevé selon le protocole B(n,p)
Quand $n \to \infty$, la variable centrée réduite $Z=\frac{F-p}{\sigma}$ suit une loi normale N(0,1) et …
Si $P(-U_\propto < Z < U_\propto) = 1-\propto$ c'est aussi la probabilité pour que F se trouve dans l'intervalle $[p - U_\propto\frac{\sqrt{p(p-1)}}{\sqrt{n}} ; p + U_\propto\frac{\sqrt{p(p-1)}}{\sqrt{n}}]$

Par exemple

$U_{0,05}=1,96$ donc la probabilité pour que F se trouve dans l'intervalle $[p -1,96\frac{\sqrt{p(p-1)}}{\sqrt{n}} ; p +1,96\frac{\sqrt{p(p-1)}}{\sqrt{n}}]$ est $1-0,05 = 0,95$

$[p -\mathbf{1,96}\frac{\sqrt{p(p-1)}}{\sqrt{n}} ; p +\mathbf{1,96}\frac{\sqrt{p(p-1)}}{\sqrt{n}}]$ est **l'intervalle de fluctuation asymptotique de F au risque de 95%**

Il suffit de remplacer 1,96 par $U_\propto$ pour définir l'intervalle de fluctuation asymptotique (IFA) au risque $1 - \propto$.
L'exigence minimale de fiabilité du test est $n \geq 30$ $np \geq 5$ et $n(1-p) \geq 5$.

Utilisation au niveau de la décision

Si F mesurée dans un échantillon appartient à l' IFA au risque de 95% on considère que l'échantillon est compatible avec le modèle. Sinon on le rejette.

Intervalle de confiance

Alors que l'IFA est utilisé **dans la prise de décision**, l'intervalle de confiance est utilisé lors de **l'estimation.**

Autrement dit on ne connaît pas p et on l'estime à une fréquence F mesurée dans un échantillon de taille n. Comme on ne peut affirmer avec certitude que la valeur de F mesurée est égale à la fréquence p réelle, on détermine un intervalle dans lequel p devrait se situer au risque de 5%. Cet intervalle dépend évidemment de la taille n de l'échantillon. Plus n sera grand plus notre estimation sera fiable et plus l'intervalle de confiance sera réduit autour de la fréquence mesurée.

L'intervalle de confiance de l'estimation de p au risque de 5% est $\mathbf{[F - \frac{1}{\sqrt{n}} ; F + \frac{1}{\sqrt{n}}]}$

L'amplitude de cet intervalle est $\frac{2}{\sqrt{n}}$. Elle décroit avec n.

L'exigence minimale de fiabilité du test est $n \geq 30$ $np \geq 5$ et $n(1-p) \geq 5$.