$\boxed{\text{Probabilités}}$

Martine Quinio

14 décembre 2012

Résumé

On suppose connus les premiers chapitres de base.
Les probabilités occupent une place importante au lycée. On verra les probabilités conditionnelles en exercice. On revoit la loi binomiale comme exemple en vue d'étudier la loi de Gauss, au centre de ce cours. Notions de statistique, échantillon. Source : Ouvrage « Probabilités et Statistiques aujourd'hui» Editions L Harmattan 2009 Martine Quinio

1 Variables aléatoires, généralités

Exemple 1.1 *Considérons un automobiliste qui, sur le trajet qui l'amène au travail, rencontre 8 feux tricolores ; les réglages des feux sont tels que la probabilité pour chaque feu d'être vert est de* $\frac{2}{3}$. *On suppose pour simplifier que ces feux sont indépendants les uns des autres.*

Définir un modèle probabiliste pour calculer les probabilités des événements suivants :
- lors d'un trajet, les 8 feux sont verts ;
- lors d'un trajet, exactement 5 feux sur 8 sont verts.

Nous pouvons, compte tenu des hypothèses, considérer la suite des 8 feux rencontrés un certain jour comme une succession de V et de non V :

Comme on préfère avoir un feu vert, on notera S (succès) : l'événement : «le feu est vert», et E (échec) sinon.

L'événement « Lors d'un trajet, les 8 feux sont verts » a pour probabilité : $(\frac{2}{3})^8$.

« Lors d'un trajet, exactement 5 feux sur 8 sont verts » a pour probabilité :

$$P(B) = \binom{8}{5} . (\frac{2}{3})^5 (\frac{1}{3})^3$$

Exemple 1.2 *Supposons que l'on joue à pile ou face : on lance une pièce truquée, 8 fois de suite. On donne la probabilité de "pile" :* $\frac{2}{3}$

Définir un modèle probabiliste pour calculer les probabilités des événements suivants :
- A : on obtient 8 fois pile lors des 8 lancers ;
- B : sur les 8 lancers, on obtient exactement 5 fois pile. Généraliser.

Nous donnerons le modèle suivant pour définir un cadre probabiliste à ce problème.

Les cas possibles sont toutes les suites (ordonnées, donc) de P ou F, avec 8 lettres en tout.

Ces 2^8 résultats possibles sont équiprobables : par exemple : (P, P, F, F, P, F, P, P) est un résultat possible.

On obtient 8 fois *pile* lors des huit lancers si on a seulement la lettre P dans la suite : il y a un seul cas : la probabilité cherchée de A est donc : $(\frac{2}{3})^8$.

Sur les 8 lancers, on obtient exactement 5 fois *pile* : il s'agit de compter de combien de façons possibles cet événement peut être réalisé :

Cela revient à choisir donc 5 places pour les 5 lettres P, les trois autres étant des F :

Il y a donc $\binom{8}{5}$ choix favorables ; il revient au même de placer les *face* : $\binom{8}{5}$.

La probabilité $P(B)$ cherchée est donc :

$$P(B) = \binom{8}{5}.(\frac{2}{3})^5.(\frac{1}{3})^3$$

Conclusion :

Modélisation.

Lors de chacune des expé riences, chaque é preuve,
deux issues sont possibles, que l'on symbolise par les noms
« succès », « échec », on connaît la probabilité p de succès, donc celle de l'échec : 1-*p*.

Alors, la probabilité d'obtenir exactement k succès au cours des n épreuves est :

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Exemple 1.3 *Une usine fabrique des boîtes rectangulaires aux dimensions commandées par un groupe industriel ; parmi ces boîtes, certaines, c'est inévitable, vont présenter un défaut ; le nombre X de ces pièces défectueuses, est en général inconnu mais estimé : la probabilité pour qu'une pièce soit défectueuse est p.*

Le nombre de ces pièces est une variable aléatoire ; chaque jour de fabrication, ce nombre varie, dans une fourchette que l'on aimerait bien évaluer ; l'impossibilité de dire le nombre exact de telles pièces nous amène à poser ce nombre comme dépendant du hasard.

Si l'usine fabrique 1000 boîtes par jour, nous pourrons par exemple noter :

X « le nombre de pièces défectueuses parmi les 1000 » : alors les valeurs prises par X sont comprises entre 0 et 1000.

Peut-on calculer les probabilités des événements :

- « Le nombre de pièces défectueuses parmi les 1000 vaut exactement 10 » ?
- « Le nombre de pièces défectueuses parmi les 1000 est compris entre 2 et 10 » ?

Une solution pourra être donnée dès lors qu'un cadre, un modèle probabiliste sera défini.

On peut considérer l'expérience aléatoire suivante : imaginons tout le lot de boîtes devant nous, prélevons une boîte l'une après l'autre en examinant les défauts, et remettons-la dans le lot après chaque tirage, après avoir marqué B si la pièce est bonne, D si elle présente un défaut.

Après n tirages dans ces conditions, les résultats sont résumés par une séquence (-,-,-,-,-,-,-,-,-,-), où chaque tiret est un D ou un B : $P(D) = p$

Le cadre probabiliste est bien défini.

Par le même raisonnement, on a pour toute valeur de k compris entre 0 et n,

La probabilité de l'événement :"La séquence comporte k fois la lettre D" ou encore « X est égal à k » , notée $P[X = k]$, est :

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$$

1.1 Les principes de base

Cas général :

Définition 1.1 *Un univers et une probabilité étant associés à une expérience aléatoire, une variable aléatoire est une fonction qui permet d'associer à tout élément ω de l'univers Ω, une valeur numérique connue éventuellement seulement après l'expérience.*

On note les variables aléatoires par des grandes lettres X, Y, Z...

$$X : \Omega \to \mathbb{R}$$
$$\omega \to X(\omega)$$

C'est une définition qui permet d'englober, par exemple :

- la somme des points obtenus en lançant deux dés définit une variable aléatoire, dont l'ensemble des valeurs possibles, que l'on appelle le support de la variable aléatoire, numérique ici, sont les nombres compris entre 2 et 12 ;

- la durée de vie d'une ampoule est, en jours, un nombre compris entre 0 et 1000.

- le nombre nécessaire de lancers d'une pièce jusqu'à l'obtention de *pile* définit une variable aléatoire de support $[1, +\infty[$;

- la vitesse, la position d'une molécule d'un gaz ;

- le nombre de grippés sur une population d'une grande ville...

Ce sont là des exemples de variables aléatoires quantitatives.

Définition 1.2 *Une variable aléatoire est discrète si les valeurs prises par elle sont dénombrables, continue dans le cas contraire.*

1.2 Variables aléatoires discrètes

Une variable aléatoire discrète est définie dès que l'on connaît :

- toutes ses valeurs possibles ; on appelle leur ensemble le *support* de X
- les probabilités associées à ces valeurs :

C'est pour cela que l'on doit, en préalable à tout calcul, définir le cadre probabiliste.
Ces données constituent ce que l'on appelle la *loi de probabilité de X ; on dit aussi distribution de probabilité, qui veut bien dire de quelle façon la probabilité totale,* 1, *est « distribuée » entre les valeurs possibles !*

Définition 1.3 *On a modélisé une expérience aléatoire si on lui a associé une loi de probabilité.*

On note :

$$[X = k] = \{\omega \in \Omega, tel\ que\ X(\omega) = k\}$$

La loi de probabilité de X est la donnée de tous les nombres $P[X = k]$, pour toute valeur k susceptible d'être prise par X ; de même, on écrira, a et b étant des réels :
$[a \leq X \leq b] = \{\omega \in \Omega, tel\ que\ :\ a \leq X(\omega) \leq b\}$: c'est un événement associé à l'expérience aléatoire.

Remarque 1.1 *Il ne faut pas confondre "variable aléatoire" X et "loi de probabilité" de X : une loi donnée est unique, mais peut être celle d'une infinité de variables aléatoires ; par exemple, les deux variables aléatoires donnée en débute de chapitre ("nombre de pile au cours de 8 lancers" et "nombre de feux verts") obeissent à la même loi binomiale* $B(8; \frac{2}{3})$.

1.3 Paramètres d'une variable aléatoire discrète

Définition 1.4 *L'espérance mathématique d'une variable aléatoire X discrète est le nombre noté $E(X)$ qui est la moyenne des valeurs prises par X, pondérées par leur probabilité respective.*

On a donc :

$$E(X) \ = \sum_{x_i \in SuppX} x_i P[\ X = x_i]$$

les réels x_i décrivant le support de X.
Supposons que la loi de probabilité d'une variable aléatoire Y soit donnée par :

k	1	2	3
$P[S = k]$	0.36	0.54	0.1

:

C'est bien une distribution de probabilité ; quelle est son espérance ?
C'est : $E(Y) = 1.74$:
Ainsi, l'espérance n'est pas 2, ce n'est pas une valeur entière.

L'espérance d'une variable aléatoire est une moyenne pondérée : il est naturel d'évaluer l'écart à la moyenne, comme nous l'avons fait dans la partie *Statistique descriptive* ; le vocabulaire, d'ailleurs, est le même.

Définition 1.5 *La variance d'une variable aléatoire est aussi une moyenne pondérée :*

$$Var\ X\ = \sum_{x_i} (x_i - E(\ X))^2 . P[\ X = x_i]$$

Par exemple pour la variable aléatoire Y ci-dessus, $VarY = 0.392\,4$.

Pour des raisons d'homogénéité, on définit *l'écart-type*, noté $\sigma(X)$: C'est la racine carrée de la variance.

$$\sigma(X) = \sqrt{VarX}$$

1.4 Règles de calcul

Pour deux variables aléatoires X et Y possédant espérance et variance, et tout a réel :
1) $E(X) + E(Y) = E(X + Y)$, ceci se généralise à une somme quelconque de variables aléatoires.
2) $E(X + a) = E(X) + a$.
3) $E(aX) = aE(X)$.
4) $Var(X + a) = Var(X)$.
5) $Var(aX) = a^2 Var(X)$, en particulier, $Var(-X) = Var(X)$.
6) En général, on a :
$Var(X + Y) \neq VarX + VarY$ et $E(X)E(Y) \neq E(XY)$.

7) $VarX = E(X^2) - (E(X))^2$

2 Indépendance

La notion d'indépendance d'événements est connue : rappelons que deux événements A et B définis pour une même expérience aléatoire sont indépendants si :

$$P(A \cap B) = P(A)P(B)$$

Alors on a la définition très simple :

Définition 2.1 *Deux variables aléatoires discrètes X et Y sont indépendantes si tout événement relatif à X est indépendant de tout événement relatif à Y, c'est à dire si, les supports de X et Y étant notés respectivement : $\{x_i, 1 \leq i \leq N\}$ et $\{y_j, 1 \leq j \leq M\}$, les événements $[X = x_i]$ et $[Y = y_j]$ sont indépendants, et ce pour tous les indices i et j possibles.*

proposition 2.1 *Si les variables aléatoires discrètes X et Y sont indépendantes, alors :*

$$Var(X + Y) = VarX + VarY$$

2.1 Loi de Bernoulli, fonction indicatrice

Un exemple simple de variable aléatoire est le suivant : à un événement A donné d'un espace probabilisé (Ω,P(Ω),p) on associe la fonction notée :

$1_A : \Omega \to \mathbb{R}, 1_A(\omega) = 1$ si $\omega \in \Omega, 1_A(\omega) = 0$ si $\omega \notin \Omega$:

1_A est une variable aléatoire et on a facilement : $E(1_A) = P(A), Var(1_A) = P(A)(1 - P(A))$

1_A est appelée fonction indicatrice d'un événement A

Exemple 2.1 *On lance un dé normal et l'on parie sur la sortie du numéro* 4 *; il y a donc un seul numéro gagnant, les autres sont perdants. Cette situation peut se modéliser par la variable aléatoire* X *suivante. Il y a deux issues possibles, représentables par* 0 *et* 1*, si l'on veut, avec la convention :*

$X = 1_{[\omega=4]}$

On a alors la loi de probabilité de X qui est définie par le tableau :

x_i	0	1
$P[X = x_i]$	$\frac{5}{6}$	$\frac{1}{6}$

Plus généralement, on appelle *variable de Bernoulli* une variable aléatoire X dont le support est constitué de deux valeurs, matérialisées par 0 et 1 ; il suffit alors de connaître la probabilité p de l'événement $[X = 1]$, celle de $[X = 0]$ étant alors $1 - p$. On dit que la variable aléatoire de Bernoulli est de paramètre p.

Précisons *l'espérance et la variance d'une telle variable aléatoire de Bernoulli :*

- L'espérance est : $1 \times p + 0 \times (1 - p) = p = E(X)$;
- La variance est : $E(X^2) - p^2 = p(1 - p)$.

2.2 Loi binomiale

- *répétition de n épreuves aléatoires indépendantes ;*
- *chaque épreuve associée à deux issues du type S « succès » ou nonS « échec »*
- *la probabilité p est donnée* : $P(S) = p$

X est la variable aléatoire « nombre de succès au cours des n épreuves » : on reconnait le schéma de Bernoulli.

Pour toute valeur de k compris entre 0 et n, on a le résultat suivant :

La probabilité de l'événement : « X est égal à k » , est :

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$$

Lorsqu'une variable aléatoire suit une loi de probabilité qui répond à ce modèle, on dit que la loi suivie par X est une loi binomiale de paramètres : n, le nombre total de répétitions d'épreuves, et p, la probabilité de succès pour chaque épreuve.

proposition 2.2 *Toute variable aléatoire qui suit une une loi binomiale de paramètres n (nombre total d'épreuves indépendantes) et p, probabilité de succès, est la somme de n variables aléatoires de Bernoulli, indépendantes, de paramètre p. Réciproquement, la somme de n variables aléatoires de Bernoulli indépendantes est une variable aléatoire dont la loi est une loi binomiale de paramètres n et p.*

Ce résultat a déjà une conséquence intéressante : il nous permet de calculer l'espérance et la variance d'une variable aléatoire qui suit une loi binomiale de paramètres n et p :

Puisqu'une telle variable aléatoire s'écrit : $Y = X_1 + X_2 + X_3 + ... + X_n$ et que les X_i suivent des lois de Bernoulli et sont indépendantes :

L'espérance est la somme des espérances, toutes égales à p ; la variance est la somme des variances car les variables sont indépendantes :

$$\boxed{\text{E(X)=np}}$$

$$\boxed{\text{Var(X)=np(1-p)}}$$

3 La loi normale

3.1 Courbe de Gauss

3.1.1 Historique de la Courbe de Gauss

Laplace, en France : Il étudie les naissances en posant la probabilité d'avoir une fille p=0,49 ; le nombre k de filles parmi n enfants est une variable aléatoire qui suit une loi binomiale de paramètres n et p :

Il trace un diagramme, avec, en abscisse, les diffrentes valeurs de k, et, en ordonnée, les probabilités associées ; Laplace cherche la courbe qui lisse les observations. Laplace obtient les seules fonctions du type :

$$e^{-\frac{1}{2}(x-m)^2}$$

Ensuite, Laplace s'intéresse au problème des erreurs, erreurs de mesure faites sur une série d'observations de pression données par un baromètre divers moments de la journée. Problème :

Comment obtenir l'estimation d'une mesure compte tenu des erreurs dues aux instruments, des erreurs de lecture et de multiples petites causes perturbatrices ; Laplace a l'idée de faire un grand nombre de mesures, ce qui constitue une approche statistique du problème.

Supposons que l'on effectue une série d'observations et dsignons par Ei l'erreur associ

Gauss, en Allemagne : Gauss se demande comment estimer la valeur d'un paramètre partir de rsultats observés ; il développe alors la théorie de l'estimation par la méthode des moindres carrés, pose une fonction dite fonction de vraisemblance, qui est une probabilité ; posant quelques hypothèses de régularité, Gauss démontre qu'alors, les seules fonctions possibles sont les fonctions du type

$$e^{-\frac{1}{2}x^2}$$

Et retrouve les fonctions de Laplace dont la courbe représentative a l'allure de la courbe en cloche...

4 La probabilité est une aire...

Pour les variables aléatoires continues, la probabilité est une aire qui se calcule par une intégrale ;

On définit la :

La fonction de répartition F, telle que :

$$F(a) = P[X \leq a]$$

Et aussi :

$$P[a \leq \ X \leq b] = F(b) - F(a)$$

Pour toute valeur a, F(a) est l'aire de la surface délimitée par **:**

- la courbe de la fonction f densité
- la droite d'équation « $x = a$ » ;
- l'axe des abscisses.

Définition 4.1 *On dit qu'une variable aléatoire X suit une loi de Gauss (ou loi normale) si sa fonction de répartition a comme densité de probabilité sur $\mathbb{R}$, la fonction définie par :*

$$f : x \to \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-m)^2}.$$

La courbe de f est symétrique par rapport à la droite d'équation « $x = m$ ».

Le paramètre σ, mesure le degré d'étalement de la courbe

Un calcul mathématique d'intégration montre que cette fonction est bien une densité de probabilité :

L'aire totale sous la courbe vaut 1. On dit X que suit une loi de Gauss N(m,σ)

Mais on ne peut pas calculer de primitive de cette fonction f, alors :

Définition 4.2 *On dit qu'une variable aléatoire Z suit une loi normale centrée réduite si sa fonction de répartition*

$$F(x) = P[Z \leq x]$$

a comme densité de probabilité sur $\mathbb{R}$, la fonction définie par :

$$f : x \to \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Par simple changement de variable on a :

proposition 4.1 *Si X suit une loi de Gauss N(m,σ), $alorsZ$= $\dfrac{X-m}{\sigma}$ suit une loi normale centrée réduite*

proposition 4.2 *L'espérance d'une variable aléatoire qui suit une loi de Gauss est : $E(X) = m$ et sa variance est : $VarX = \sigma^2$.*

Ainsi :

L'espérance d'une variable aléatoire qui suit une loi normale centrée réduite est : $E(X) = 0$ et sa variance est : $VarX = 1$.

Remarque 4.1 *Les valeurs de F(a), pour une loi centrée réduite sont tabulées.*

Ainsi, dans une table de Gauss, on peut lire, pour toute valeur de a comprise entre 0 et 2.99, les valeurs de $F(a)$:

Par exemple, on lit pour $a = 1.56$: $F(a) = 0.9406$;

Inversement, pour 0.9750, on lit la valeur 1.96.

On peut aussi, si on a beaucoup de lecture et de calculs à faire, utiliser des logiciels (EXCEL, R par exemple), qui donnent les valeurs de la loi de Gauss.

Il est traditionnel de noter Π la fonction de répartition de la loi centrée réduite ; les tables donnant les valeurs de $\Pi(a)$, pour a compris entre 0 et 2.99, il faut utiliser la symétrie de la courbe pour obtenir les valeurs de $\Pi(a)$ pour a négatif : On a facilement :

proposition 4.3

$$\Pi(-a) \;=\; 1-\Pi(a)$$

On en déduit :

$$\Pi(a)-\Pi(-a)=2\Pi(a)-1=P[|X|\leq a]$$

Par exemple, cherchons pour quelle valeur de a l'aire mesurée par $\Pi(a)-\Pi(-a)$, c'est à dire, l'aire comprise entre les droites d'équation « $x=a$ » et « $x=-a$ », est égale à 0.95 :

Il suffit de résoudre la petite équation :

$2\Pi(a)-1=0.95$, soit, $\Pi(a)=\dfrac{1.95}{2}=0.975$: on l'a vu, la valeur correspondante de a est : $a=1.96$.

proposition 4.4 *Pour une variable aléatoire X qui suit loi normale centrée réduite, la probabilité*

$$P[-1.96\leq\; X\leq 1.96]=0.95$$

Or, l'écart-type d'une telle variable est $\sigma=1$: donc la probabilité pour que X soit compris entre -2 et 2 est environ 95%.. C'est la règle du double de l'écart-type.

Définition 4.3 *Pour une variable aléatoire X qui suit loi normale centrée réduite, on dira que l'intervalle* $[-1.96;1.96]$ *est un* intervalle de confiance *pour X au niveau de confiance de 95%;*

En fait, la loi de Gauss est intéressante surtout par ses nombreuses applications...

4.1 En pratique...

Exemple 4.1 *La répartition du taux de cholestérol suit une loi normale N(163; 13,3) Donc la moyenne est de 163 cg, l'écart-type est 13,3;*
Quelle est la proportion de personnes dont le taux est plus grand que 180 cg? Solution :
$P[X\leq 180]=P[\;\dfrac{X-163}{13,3}\leq\;\dfrac{180-163}{13,3}]$
On lit dans la table $P[Z\leq 1,28]=0,9$ *Donc la proportion de personnes dont le taux est plus grand que 180 cg est 0,10. On peut dire aussi que* 95% *des personnes ont un taux de cholestérol*
compris entre 150*cg et* 190*cg...*
d' après la règle du double de l'écart-type.

Lecture de table :
X est une variable aléatoire qui obéit à une loi normale de paramètres m et σ : - Si $m=0$, $\sigma=\;1$, calculez :

$a)$ $P[X\leq 1,45]$, $b)$ $P[X\geq 1,82]$ $c)$ $P[X\leq -1,65]$ $d)$ $P[|X|\leq 1,45]$.

$e)$ $P[X\leq 1,435]$ $f)$ $P[X\leq 1,452]$.

- Si $m=25$, et $\sigma=5$: calculez :

$g)$ $P[X\leq 32]$ $h)$ $P[X>-28]$ $i)$ $P[|X|\leq 30]$.

5 Somme de gaussiennes

Th'eorème 5.1 *Si X et Y sont deux variables aléatoires indépendantes qui suivent, respectivement, les lois normales de paramètres m et σ, et paramètres m' et σ', alors la somme $Z = X + Y$ suit une loi normale paramètres $m" = m + m\prime$ et $\sigma" = \sqrt{\sigma^2 + \sigma'^2}$:*

Ce théorème affirme surtout que la somme de deux variables indépendantes gaussiennes est gaussienne ; de plus, l'indépendance sert à calculer la variance de Z qui est la somme des variances de X et de Y.

Preuve 5.1 *(Partielle)*
Soient X et Y deux variables aléatoires :

$$E(X+Y) = E(X) + E(Y)$$

Comme elles sont indépendantes :

$$Var(X+Y) = VarX + VarY$$

$$Var(X+Y) = \sigma^2 + \sigma'^2$$

Notation abrégée :
$X \sim N(m, \sigma)$ et $Y \sim N(m\prime, \sigma')$ et X et Y indépendantes, alors :

$$X + Y \sim N(m + m\prime, \sqrt{\sigma^2 + \sigma'^2}).$$

Exemple 5.1 *Les responsables de compagnies aériennes admettent que le poids X des passagers suit une une loi normale de moyenne 70 kg, d'écart-type : 8 kg ; le poids des bagages, Y, lui, suit une loi normale de moyenne 15 kg, d'écart-type : 5 kg. En admettant que ces variables aléatoires sont indépendantes, quelle est la loi suivie par le poids total d'un passager et de ses bagages ? L'avion comporte 300 places : quelle est la loi de probabilité de la variable aléatoire : « poids total » ?*

Solution :
Le poids total d'un passager et de ses bagages suit une loi normale, puisque les variables « poids passager » et « poids bagages » sont indépendantes, de moyenne 85 kg et d'écart-type :$\sqrt{8^2 + 5^2} = \sqrt{89}$ kg.
La loi de la variable aléatoire « poids total » est celle de la somme : $Z = \sum_{i=1}^{300} Z_i$, où chaque Z_i suit une loi normale de moyenne 85 kg et d'écart-type : $\sqrt{8^2 + 5^2}$kg ; Z suit exactement une loi normale de moyenne $E(Z) = 300.E(Z_i)$, et d'écart-type :
$\sigma(Z) = \sqrt{300.VarZ_i} = \sqrt{300}\sigma(Z_i)$.

Exemple 5.2 *On suppose que la température moyenne du mois de mai à Marseille est une variable aléatoire qui suit une loi normale de moyenne : 19° et d'écart-type : 2° ; on suppose que ces températures moyennes sont indépendantes d'une année sur l'autre. Quelle est la loi suivie par la moyenne des températures moyennes en mai, à Marseille, sur trois années consécutives ? Quelle est la probabilité pour que, durant trois années consécutives, la température moyenne dépasse 20° ?*

Solution :

Notons X_i , $i \leq 3$, la variable aléatoire : « température moyenne du mois de mai, année i, à Marseille »

La moyenne sur 3 ans est la variable aléatoire : $X = \dfrac{X_1 + X_2 + X_3}{3}$: l'espérance de X est 19°, la variance de X est : $VarX = \dfrac{1}{9}(Var\ X_1 + Var\ X_2 + Var\ X_3) = \dfrac{3}{9}.\ 2^2 = \dfrac{4}{3}$.

La loi suivie par la moyenne des températures moyennes en mai, à Marseille, sur trois années consécutives est une loi normale de moyenne : 19° et d'écart-type : $\sqrt{\dfrac{4}{3}}$

La probabilité pour que, durant trois années consécutives, la température moyenne dépasse 20° peut se calculer comme la probabilité de trois événements indépendants :

C'est donc $(P[X_i \geq 20])^3$: Or, $P[X_i \geq 20] = P[\dfrac{X_i - 19}{2} \geq \dfrac{20 - 19}{2}] = 1 - \Pi(\dfrac{1}{2})$.

Donc : $P[X_i \geq 20])^3 \simeq 2.936\ \times\ 10^{-2}$: environ 3 pour 100 est la probabilité cherchée.

6 Convergence vers la loi normale

Le point suivant justifie l'appellation « loi normale »

Th'eorème 6.1 *(dit : « central limite » ou de la limite centrale) :*

Si (X_i) $i \leq n$, sont des variables aléatoires indépendantes et de même loi, quelle que soit cette loi, m leur espérance commune et σ^2 la variance commune, alors la suite des variables aléatoires $(\dfrac{\overline{X} - m}{\frac{\sigma}{\sqrt{n}}})$ converge en loi vers la loi normale centrée réduite $N(0,1)$, en notant $\overline{X}$ la variable aléatoire :

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Définition 6.1 *Une suite de variables aléatoires indépendantes et de même loi est un n-échantillon de cette loi.*

Cela signifie que la loi de probabilité de la moyenne ("moyenne d'échantillon") de variables aléatoires indépendantes et de même loi, ***quelle que soit cette loi****, se comporte, pour n assez grand, comme une loi normale* $N(m, \frac{\sigma}{\sqrt{n}})$:

Remarque 6.1 *Montrons à présent la puissance du théorème :* Son intérêt est que l'on peut évaluer des probabilités du type :

$P[\frac{\sum_{i=1}^{n} X_i}{n} \leq a]$, ou encore, $P[\sum_{i=1}^{n} X_i \leq na]$, même si l'on ne connaît pas la loi des X_i, ou même si cette loi est compliquée, pourvu que ces variables aléatoires soient indépendantes, de même loi et que n soit assez grand : c'est un résultat très puissant !

Exemple 6.1 *On suppose que la température moyenne du mois de mai à Marseille est une variable aléatoire qui suit une loi inconnue de moyenne :* $19°$ *et d'écart-type :* $2°$; *on suppose que ces températures moyennes sont indépendantes d'une année sur l'autre. Quelle est la loi suivie par la moyenne des températures moyennes en mai, à Marseille, sur 50 années consécutives ?*

La loi suivie par la moyenne $X = \dfrac{\sum_{i=1}^{50} X_i}{50}$ des températures moyennes en mai, à Marseille, sur cinquante années consécutives, est une loi normale de moyenne : $E(X) = 19° = E(X_i)$, et d'écart-type : $\sigma(X) = \dfrac{1}{50}\sqrt{50.VarX_i} = \dfrac{\sigma(X_i)}{\sqrt{50}} = \dfrac{2}{\sqrt{50}}$.

6.1 Loi des erreurs

Supposons qu'une usine fabrique des pièces dont l'une des dimensions est une variable aléatoire qui suit une loi normale ; l'erreur est une variable aléatoire qui suit une loi normale N(0 ; σ) :

Donc pour chaque pièce, l'erreur Ei est telle que :

$$P[-1.96\sigma < Ei \leq 1.96\sigma] = 0.95$$

Alors si l'erreur moyenne est : $\overline{X} = \frac{\sum_{i=1}^{n} Ei}{n}$:

$\sigma(\overline{X}) = \dfrac{\sigma}{\sqrt{n}}$:

Donc sur n pièces la précision est $\sqrt{n}$ fois meilleure puisque :

$$P[-1.96\frac{\sigma}{\sqrt{n}} < \frac{\sum_{i=1}^{n} Ei}{n} \leq 1.96\frac{\sigma}{\sqrt{n}}] = 0.95$$

Application :

Soit une série de n mesures xi d'une distance (exemple : Terre Soleil à t donné) :

Xi : la variable aléatoire "mesure i"

Posons : Ei=d-Xi :

L'espérance de Xi est d.

Le résultat précédent donne :

$P[d - 1.96\dfrac{\sigma}{\sqrt{n}} < \frac{\sum_{i=1}^{n} Xi}{n} \leq d + 1.96\dfrac{\sigma}{\sqrt{n}}] = 0.95$:

Donc la moyenne des mesures, m=$\frac{\sum_{i=1}^{n} xi}{n}$ étant une réalisation de la variable aléatoire $\frac{\sum_{i=1}^{n} Xi}{n}$:

La moyenne m donne la distance d avec une précision $\sqrt{n}$ fois meilleure que une seule mesure !

Exemple :

400 mesures donne une précision 20 fois meilleure qu'une seule mesure :

Remarque 6.2 *La multilication des observations augmente la précision, puisque l'écart-type diminue !*

7 Théorème de Moivre Laplace

Th'eorème 7.1 *Une variable aléatoire X qui suit une loi binomiale de paramètres n et p, converge en loi vers une variable aléatoire qui suit une loi normale $N(m, \sigma)$, ce qui signifie que la fonction de répartition de X, qui dépend de n, peut être approchée par la fonction de répartition d'une loi normale de même espérance et de même variance.*

C'est une consèquence du Théorème central limite :
Si (X_i) $i \leq n$, sont des variables aléatoires indépendantes et de même loi de Bernoulli, m leur espérance commune et σ^2 la variance commune, alors la suite des variables aléatoires $(\dfrac{\overline{X} - m}{\frac{\sigma}{\sqrt{n}}})$ converge en loi vers la loi normale centrée réduite $N(0, 1)$, en notant $\overline{X}$ la variable aléatoire :

$$\overline{X} = \frac{X}{n}$$

Si $X \sim B(n,p)$ on peut approcher la loi de X par une loi normale :

$$P[-b \leq \frac{\overline{X}-m}{\frac{\sigma}{\sqrt{n}}} \leq b] \simeq \int_{-b}^{+b} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

Pour $b = 1.96$:
$P[-b \leq \frac{\overline{X}-m}{\frac{\sigma}{\sqrt{n}}} \leq b] = 0.95$, soit,

$$P[-1.96\frac{\sigma}{\sqrt{n}} \leq \overline{X} - m \leq 1.96\frac{\sigma}{\sqrt{n}}] = 0.95$$

Ainsi l'écart, en valeur absolue, entre la moyenne d'échantillon $\overline{X}$ et son espérance m est "contrôlé" :

$$P[\left|\overline{X} - m\right| \geq 1.96\frac{\sigma}{\sqrt{n}}] = 0.05$$

7.1 Vitesse de convergence et approximation

Quelles sont au juste, les conditions d'approximation ? Que signifie l'expression « n assez grand » ?

On retiendra les principes suivants d'approximation d'une binomiale $B(n,p)$ par une gaussienne $N(np, \sqrt{np(1-p)})$

Si p est proche de 0.5, l'approximation est très bonne dès que n ≥ 20;
Si p est proche de 0, il faut que $np \geq 10$;
Si p est proche de 1, il faut que $n(1-p) \geq 10$;

conclusion 7.1 La loi normale intervient comme approximation à deux niveaux :

- Au niveau d'un seul échantillon*, série de* 100 *tirages, pour approcher une loi exacte, la loi binomiale de paramètres* 100 *et* 0, 5*, par une loi normale de paramètres* 50 *et* 5.

*- Au niveau de l'*échantillonnage*, c'est à dire des séries de 1000 échantillons, pour prévoir le nombre de tirages associés aux intervalles du tableau, et ainsi, comparer des effectifs théoriques à ceux observés : c'est de la statistique..*