# II - VARIABLES ALEATOIRES ET DONNEES

J-P. Croisille

Université de Lorraine

UEL - Année 2012/2013

# 1- PROBABILITES ET STATISTIQUES

## NOTION DE PROBABILITÉ

Situation: identification d'une quantité $x$ inconnue. L'accès à $x$ se fait non pas par une équation (dont $x$ serait l'unique solution), mais par des "mesures" qui présentent une **variabilité**. La situation est la suivante:

- On dispose d'un jeu de données représentant l'incertitude sur la valeur d'une quantité.
- On introduit une **modélisation** sous-jacente: Ces données sont un échantillon d'une quantité appelée **variable aléatoire**.
- La variable aléatoire est "en amont" des données et modélise de façon abstraite la non-connaissance (ou la connaissance partielle) de la valeur de $x$.

On note dans la suite $X$ une variable aléatoire type.

## EXEMPLE 1: SCIENCES DE LA VIE

On mesure sur un échantillon de 36 sujets masculins le taux de créatine phosphokinase. La variable aléatoire correspond au tirage abstrait d'un sujet masculin et de l'évaluation de son taux de créatine. On note $X =$"taux de créatine" la **variable aléatoire** correpondante sur la population sous-jacente (ici TOUS les sujets masculins). $X$ n'a pas une valeur fixée, mais au contraire une valeur variable en fonction du sujet tiré. On ne s'intéresse pas à la valeur de $X$ chez tel ou tel sujet précis, mais à la proportion des sujets qui ont un taux entre deux valeurs fixés.

## EXEMPLE 2 : TIRAGE D’UNE CARTE

On tire une carte à jouer au hasard. La variable aléatoire est $X$= la carte qui va être tirée. La population sous-jacente est le jeu de carte complet (52 cartes). On distingue

- **L’ensemble des sorties possibles:** Ici c’est la collection des 52 cartes.
- **Un évènement:** C’est un **sous-ensemble** de l’ensemble des sorties possibles.

## TIRAGE D'UNE CARTE (2)

On s'intéresse au fait que survienne un **évenement** à l'occasion du prochain tirage.

- Evènement="sortie du roi de cœur" (évènement élémentaire). *1 seule possibilité*.
- Evènement="sortie d'un cœur". *13 possibilités.*
- Evènement="sortie d'un roi". *4 possibilités.*
- Evenement="sortie d'un "pique". *13 possibilités.*

## EXEMPLE 3: JEU DE PILE OU FACE

On lance une pièce de monnaie non truquée. La variable aléatoire $X$ sous-jacente est la valeur du résultat de la "prochaine expérience". La population est l'ensemble de TOUS les lancers à venir (ou un très grand nombre), repérés par exemple par leur numéro. Sur un grand nombre de lancers, on a la même fréquence de sorties de l'évènement "face" que de l'évènement "pile". On attribue donc aux deux possibilités de sortie à venir un nombre appelé leur **probabilité**.

$$P(\text{face}) = 0.5 \text{ et } P(\text{pile}) = 0.5 \tag{1}$$

## PROBABILITÉ D'UN ÉVENEMENT

La survenue d'un "évènement" $E$ s'identifie à un caractère de la prochaine expérience aléatoire de $X$, (le prochain "tirage" de $X$). Par exemple "la prochaine carte est un pique" constitue un "évènement". On attribue à un évènement $E$ un nombre appelé sa "probabilité" noté $P(E)$. Ce nombre reflète la vraisemblance que l'évènement $E$ se produise effectivement lors du tirage. On a les propriétés suivantes:

- la probabilité de tout évènement est un nombre compris entre $0$ et $1$.

$$0 \leq P(E) \leq 1 \tag{2}$$

- Un évènement $E$ qui ne peut pas se produire a une probabilité $P(E) = 0$.
- Un évènement $E$ qui se produit de façon certaine a une probabilité $P(E) = 1$.

# 2- ALGÈBRE DES ÉVÈNEMENTS

La notion de probabilité est donc étroitement reliée à celle de sous-ensembles de l' ensemble de tous les possibles. Le point essentiel est que l'on peut combiner entre eux de façon algébrique les “évènements” puisqu'ils s'identifient à des sous-ensembles! Les trois opérations élémentaires sont:

- Somme: On note $A+B$ l'ensemble des évènements dans lesquels l'une au moins des deux marques $a$ ou $b$ apparaît. L'opération “+” correspond à l'union des ensembles $A$ et $B$.
- Produit: On note par $A.B$, l'ensemble des évènements dans lesquels les deux marques *a et b arrivent simutanément*. L'opération “.” correspond à l'intersection des ensembles $A$ et $B$.
- Opposé: On note par $A'$ l'évènement contraire de $A$, c'est-à-dire, l'évènement *a n'arrive pas*. Il correspond à l'opération “passage au complémentaire”.

On a en particulier la relation

$$\mathcal{P}(A+B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A.B) \tag{3}$$

## ALGÈBRE DE BOOLE

- L'opération d'addition "+" correspond au "OU" ensembliste. L'ensemble $A + B$ correspond à l'évènement *a OU b arrive*, c'est-à-dire la carte se trouve dans l'ensemble $\mathcal{A} \cup \mathcal{B}$.
- L'opération de multiplication "." correspond au "ET" ensembliste. L'ensemble $A.B$ correspond à l'évènement *a ET b arrivent*, c'est-à-dire la carte se trouve dans l'ensemble $\mathcal{A} \cap \mathcal{B}$.
- L'opération "$A \mapsto A'$" correspond à l'opération "passage au complémentaire". $A'$ correspond à l 'évènement *a n'arrive pas*, c'est-à-dire la carte se trouve dans l'ensemble complémentaire de $\mathcal{A}$.

NB: L'algèbre de Boole intervient également en électronique: analyse des "portes" logiques.

## VISION ENSEMBLISTE

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\cap$ et $\cup$ sont associatives, c'est-à-dire $(A \cap B) \cap C = A \cap (B \cap C)$ et $(A \cup B) \cup C = A \cup (B \cup C)$

## VISION ALGEBRIQUE

- $A.(B + C) = (A.B) + (A.C)$
- $A + (B.C) = (A + B).(A + C)$
- . et $+$ sont associatives, c'est-à-dire $(A.B).C = A.(B.C)$ et $(A + B) + C = A + (B + C)$

## QUELQUES RÉSULTATS

- $A + A = A$
- $(A + B)' = A'.B'$
- $(A.B)' = A' + B'$
- $A.(A + B) = A$
- $A + A'.B = A + B$
- $A.B + A.B' = A$

EXEMPLE

On considère les deux évènements $A =$ *un cœur est tiré* et $B =$ *un roi est tiré*. Alors on a

- $A + B =$*est l'évènement "tirage = roi ou cœur"*
- $A.B =$*est l'évènement "tirage = roi de cœur"*

On a donc

$$\begin{aligned} \mathcal{P}(\text{soit roi soit cœur}) &= \mathcal{P}(\text{roi}) + \mathcal{P}(\text{cœur}) - \mathcal{P}(\text{roi de cœur}) \\ &= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} \end{aligned}$$

## PROBABILITÉ CONDITIONNELLE

Supposons que

- l'évènement $A$ peut se produire de $n_a$ manières différentes
- l'évènement $B$ peut se produire de $n_b$ manières différentes
- l'évènement $AB$ peut se produire de $n_{ab}$ manières différentes

On appelle $N$ le nombre total de réalisations possibles. (Dans un jeu de cartes, c'est $52$).

$$\mathcal{P}(A) = \frac{n_a}{N} \quad , \quad \mathcal{P}(AB) = \frac{n_{ab}}{N} \quad , \tag{4}$$

On peut réécrire $\mathcal{P}(AB)$ comme

$$\mathcal{P}(AB) = \frac{n_a}{N}.\frac{n_{ab}}{n_a} = \mathcal{P}(A).\frac{n_{ab}}{n_a} \tag{5}$$

Le second terme est la probabilité que $B$ se produise, sachant que $A$ se produit. On l'appelle la *probabilité conditionnelle de $B$ sachant $A$*, notée $\mathcal{P}(B/A)$. On a donc

$$\mathcal{P}(AB) = \mathcal{P}(A).\mathcal{P}(B/A) \tag{6}$$

## EVÈNEMENTS INDÉPENDANTS (1)

Si l'évènement $A$ n'a pas d'influence sur l'évènement $B$, on dit que les deux évènements sont **indépendants**. On a dans ce cas

$$\mathcal{P}(AB) = \mathcal{P}(A).\mathcal{P}(B) \tag{7}$$

### EXEMPLE:

$A$=*tirage d'un roi*. On a

$$n_a = 4 \;,\; N = 52 \;,\; \mathcal{P}(A) = \frac{4}{52} \tag{8}$$

$B$=*tirage d'un cœur*. On a

$$n_b = 13 \;,\; N = 52 \;,\; \mathcal{P}(B) = \frac{13}{52} \tag{9}$$

$AB$=*tirage du roi de cœur*. On a

$$\mathcal{P}(AB) = \frac{1}{52} = \frac{4}{52}.\frac{13}{52} = \mathcal{P}(A).\mathcal{P}(B) \tag{10}$$

Les deux évènements $A$ et $B$ sont indépendants.

## EVÈNEMENTS INDÉPENDANTS (2)

Si un joker est ajouté au jeu, on $N = 53$. On a à présent

$$\mathcal{P}(A) = \frac{4}{53} \quad , \quad \mathcal{P}(B) = \frac{13}{53} \quad , \quad \mathcal{P}(AB) = \frac{1}{53} \tag{11}$$

Mais

$$\mathcal{P}(A).\mathcal{P}(B) = \frac{4}{53}.\frac{13}{53} = \frac{1}{53}.\frac{52}{53} < \frac{1}{53} = \mathcal{P}(AB) \tag{12}$$

donc dans un jeu de 53 cartes, les deux évènements ne sont plus indépendants. L'interprétation est la suivante: si nous savons qu'une carte tiré sur 53 est un roi, alors elle ne peut être le joker. Donc on a une information sur le second évènement, qui est la catégorie de couleur auquel la carte appartient.

# 3- INTERPRETATION FREQUENTISTE VERSUS INTERPRETATION BAYESIENNE

## INTERPRÉTATION “FRÉQUENTISTE”

Une façon naturelle d’attribuer le nombre $P(E)$ à l’évènement $E$ est (dans un premier temps) la suivante:
*La probabilité de l’évènement “E” est simplement la* **proportion** *des fois que “E” arrive lorsque l’on répète* **un grand nombre de fois** *le même tirage*. Cette interprétation s’appelle l’interprétation **fréquentiste**.

## INTERPRÉTATION “BAYESIENNE”

L’autre interprétation s’appelle l’interprétation **bayesienne** (Thomas Bayes - 1702-1761, mathématicien britannique) dans laquelle la notion de probabilité est reliée (mathématiquement !) à une certitude subjective.

## THEOREME DE BAYES

On a vu que

$$\mathcal{P}(A \cap B) = \mathcal{P}(A).\mathcal{P}(B/A) = \mathcal{P}(B \cap A) = \mathcal{P}(B).\mathcal{P}(A/B) \tag{13}$$

On en déduit immédiatement que

$$\mathcal{P}(A/B) = \frac{\mathcal{P}(B/A)\mathcal{P}(A)}{\mathcal{P}(B)} \tag{14}$$

Cette relation s'appelle le *théorème de Bayes*.
Identité de Bayes généralisée. Si $\Omega = \cup_i A_i$, $A_i$ disjoints, on a

$$\mathcal{P}(B) = \mathcal{P}(\cup_i (B \cap A_i)) = \sum_i \mathcal{P}(B \cap A_i) = \sum_i \mathcal{P}(B|A_i)P(A_i) \tag{15}$$

Donc on obtient la formule de Bayes généralisée:

$$\mathcal{P}(A/B) = \frac{\mathcal{P}(B/A)\mathcal{P}(A)}{\sum_i \mathcal{P}(B/A_i)\mathcal{P}(A_i)} \tag{16}$$

Exemple:

Probablité pour quelqu'un d'être malade avant tout test:

$$P(\text{malade}) = 0.001, \quad \mathrm{P}(\text{pas malade}) = 0.999$$

Probabilité d'identifier correctement un malade à l'aide du test:

$$P(+/\text{malade}) = 0.99, \quad \mathrm{P}(+/\text{pas malade}) = 0.01.$$

Probabilité d'identifier incorrectement un non malade à l'aide du test:

$$P(+/\text{pas malade}) = 0.02, \quad \mathrm{P}(-/\text{pas malade}) = 0.98.$$

On suppose qu'une personne a un résultat "+" au test. Doit-elle s'inquiéter ?

$$\mathcal{P}(\text{malade}/+) = \frac{\mathcal{P}(+/\text{malade})\mathcal{P}(\text{malade})}{\mathcal{P}(+/\text{malade})\mathcal{P}(\text{malade}) + \mathcal{P}(+/\text{pas malade})\mathcal{P}(\text{pas malade})}$$

On obtient $P(\mathsf{malade}/+) = 0.047$

- **Le point de vue de la personne:** la personne conclut qu'elle a 4.7 % de chances d'être malade.
- **Le point de vue du médecin:** Le médecin de la personne conclut que 4.7 % des patients comme la personne en question sont malades.

INTERPRETATION FREQUENTISTE La probabilité d'un évènement est vu comme une limite : si l'évènement $E$ est observé $n$ fois sur un total de $N_e$ tirages. Alors la probabilité de $E$ est

$$\mathcal{P}(E) = \lim_{N_e \to +\infty} \frac{n}{N_e} \tag{17}$$

On a

$$0 \leq \mathcal{P}(a) \leq 1 \tag{18}$$

INTERPRETATION SUBJECTIVE OU BAYESIENNE
$E$ est une hypothèse (affirmation vraie ou fausse). Alors $\mathcal{P}(E)$ est un degré de certitude que $E$ est vraie.

# 4- DENSITE D'UNE VARIABLE ALEATOIRE

VARIABLE ALEATOIRE

Une variable aléatoire est une variable dont la valeur dépend du résultat d'une expérience aléatoire, (expérience dont on ne connaît pas le resultat).

Exemple:

On effectue des mesures de la largeur d'une pièce. Le résultat de la mesure est une variable aléatoire $X$.

## HISTOGRAMMES DES FREQUENCES RELATIVES

On réalise 10 mesures, qui donnent

| valeur | 323 | 324 | 325 |
|---|---|---|---|
| nombre de fois | 1 | 5 | 4 |

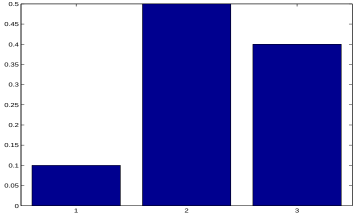



Figure: Histogramme 1

## HISTOGRAMMES DES FREQUENCES RELATIVES (2)

On effectue à présent les mesures avec un appareil plus précis. On effectue une trentaine de mesures. On obtient

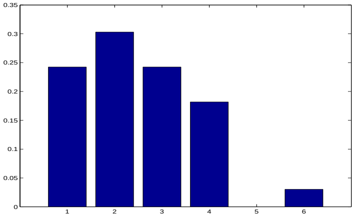
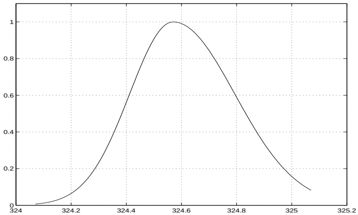



Figure: Histogramme 2

## FONCTIONS DE DISTRIBUTION LIMITE

On continue le processus: mesures de plus en plus précises et mesures de plus en plus nombreuses. On obtient à la limite une courbe en cloche: la **fonction de distribution de la variable aléatoire** $X$. C'est cette démarche qui donne sa réalité à la notion de variable aléatoire.



## FONCTION DE DISTRIBUTION (1)

La courbe limite ainsi construite, a la propriété suivante:
La probabilité $\mathcal{P}$ que la valeur mesurée soit dans l'intervalle $(x_1, x_2)$ est **l'intégrale de** $f(x)$ **sur l'intervalle** $(x_1, x_2)$.

$$\mathcal{P} = \int_{x_1}^{x_2} f(x)dx \tag{19}$$

C'est la **surface sous la courbe** $f(x)$ **entre les abscisses** $x_1$ **et** $x_2$.
La propriété générale de $f(x)$ est que

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \tag{20}$$

Le passage d'un histogramme à une fonction de distribution continue est analogue au passage d'une somme de Riemann à une intégrale.

## FONCTION DE DISTRIBUTION (2)

Comment caractériser une fonction de distribution $f(x)$ en l'absence d'une connaissance de $f(x)$ en tout $x$? On introduit deux paramètres fondamentaux, la *moyenne* et la *variance*.

### Définition ( MOYENNE)

*La moyenne de la fonction de distribution* $f(x)$ *est*

$$\mu := \int_{-\infty}^{+\infty} x f(x) dx \tag{21}$$

Pour une variable $x$ discrète, on a

$$\bar{x} := \sum_{i=1}^{+\infty} x_i \mathcal{P}(x_i) \tag{22}$$

## Définition (VARIANCE)

*La variance de la fonction de distribution $f(x)$ est la moyenne du carré de l'écart à la moyenne.*

$$\sigma^2 := \overline{(x-\bar{x})^2} = \int_{-\infty}^{+\infty} (x-\bar{x})^2 f(x)dx \tag{23}$$

Pour une variable $x$ discrète, on a

$$\sigma^2 := \sum_{i=1}^{+\infty} (x_i-\bar{x})^2 \mathcal{P}(x_i) \tag{24}$$

La variance mesure l'étalement de la densité autour de la moyenne $\mu$.

## FONCTION DE DISTRIBUTION (4)

On introduit également les “moments d'ordre supérieur” $\mu_k$, $k = 1, 2, ...,$

$$\mu_k = \int (x - \mu)^k f(x) dx \tag{25}$$

- “skewness”:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \tag{26}$$

  La “skewness” est un coefficient d'asymétrie.

- “kurtosis”

$$\gamma_2 = \frac{\mu_4}{\sigma^4} \tag{27}$$

  La “kurtosis” est un coefficient d'aplatissement.

## FONCTION DE DISTRIBUTION (5)

Cas de plusieurs variables: on mesure à présent deux variables aléatoires $X_1$ et $X_2$, par exemple la longueur et la largeur d'une pièce. La fonction de distribution dépend du couple $(X_1, X_2)$. la probabilité qu'une mesure $(x_1, x_2)$ de $(X_1, X_2)$ soit telle que $x_1 \in (a_1, a_2)$ et $x_2 \in (b_1, b_2)$ est

$$\mathcal{P}(a_1 < X_1 < a_2, b_1 < X_2 < b_2) = \int_{(x_1,x_2)\in(a_1,a_2)\times(b_1,b_2)} f(x_1, x_2) dx_1 dx_2 \tag{28}$$

La caractérisation du fait que $X_1$ et $X_2$ sont **indépendantes** est que

$$f(x_1, x_2) = f_1(x_1).f_2(x_2) \tag{29}$$

où $f_1$ est la densité de $X_1$ et $f_2$ est la densité de $X_2$. On a dans ce cas

$$\mu_{X_1+X_2} = \mu_{X_1} + \mu_{X_2} \tag{30}$$

et

$$\sigma^2_{X_1+X_2} = \sigma^2_{X_1} + \sigma^2_{X_2} \tag{31}$$

## CORRELATIONS (1)

Deux évènements $A$ et $B$ sont corrélés lorsqu'ils ne sont pas indépendants. On a dans ce cas

$$\mathcal{P}(A).\mathcal{P}(B) < \mathcal{P}(AB) \tag{32}$$

## Définition

*La covariance de $X_i$ et $X_j$ est*

$$\text{cov}(X_i, X_j) = \int (x - \mu_X)(y - \mu_Y) f(x, y) dx dy \tag{33}$$

On introduit également le coefficient de corrélation $q_{ij}$, qui caractérise la corrélation des variables $X_i$ , $X_j$, défini par

$$q_{ij} = \frac{\text{cov}(X_i, X_j)}{\sigma_i \sigma_j} \tag{34}$$

## CORRELATIONS (2)

Si les variables $X_i$, $X_j$ sont indépendantes, leur coefficients de corrélation est nul: $q_{ij} = 0$. A l'opposé, si $X_i$ est proportionnelle à $X_j$, $X_i = aX_j + b$, alors, on a $q_{ij} = 1$ si $a > 0$, et $q_{ij} = -1$ si $a < 0$,

## CORRELATIONS (3)

Exemple de corrélation:

Soit deux variables $X_1$, $X_2$ indépendantes de moyenne $\mu$ et de variance $\sigma^2$. On considère les deux variables $Y_1$, $Y_2$ définies par

$$\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 \\ Y_2 = a_{21}X_1 + a_{22}X_2 \end{cases} \tag{35}$$

Alors on a

$$\text{cov}(Y_1, Y_2) = (a_{11}a_{21} + a_{12}a_{22}) \tag{36}$$

Donc les deux variables $Y_1$ et $Y_2$ sont corrélées en général.

# 5- EXEMPLES DE LOI

## QUELQUES LOIS USUELLES

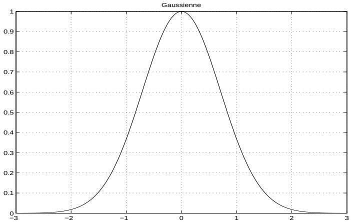
- **Lois discrètes**
  Loi binomiale, loi de Poisson, loi binomiale négative, loi hypergéométrique,...
- **Lois continues**
  Loi de Gauss, loi du $\chi^2$, loi de Student, loi de Weibull, loi gamma, loi F,...

# 6- LOI NORMALE

## DISTRIBUTION DE GAUSS (1)

La distribution de Gauss, ou Gaussienne, joue un rôle fondamental en pratique. Elle est basée sur la fonction

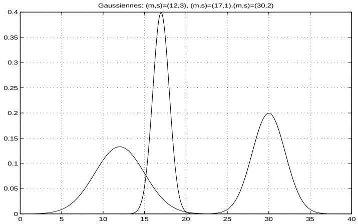
$$x \in \mathbb{R} \mapsto \exp(-x^2) \tag{37}$$



# DISTRIBUTION DE GAUSS (2)

La distribution de Gauss a deux paramètres: sa moyenne $m$ et sa variance $\sigma^2$. Elle s'écrit

$$x \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \qquad (38)$$




XX

## DISTRIBUTION DE GAUSS (3)

THEOREME “LIMITE CENTRALE”

Soit $X$ une grandeur physique ayant une moyenne $\mu$ et une variance $\sigma^2$. Si la variance $\sigma^2$ est finie, alors la ditribution de la valeur moyenne sur un grand nombre de mesures ($n \to +\infty$), c-a-d,

$$X = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{39}$$

tend vers une distribution de Gauss avec moyenne $\mu$ et variance $\sigma$.

## SENS INTUITIF DE LA DISTRIBUTION DE GAUSS

La distribution de Gauss est entièrement déterminée par deux paramètres:

- Sa moyenne $\mu$
- Son écart-type $\sigma$

On la note $N(\mu, \sigma)$. La plupart des grandeurs physiques mesurées peuvent être décrites par cette distribution. Donc la plupart des résultats expérimentaux peuvent être caractérisés par ces deux valeurs seulement. On présente habituellement les résultats sous la forme

$$x_{\text{exp}} = \bar{x} \pm \Delta x \equiv \mu + \sigma \tag{40}$$

## SENS INTUITIF DE LA DISTRIBUTION DE GAUSS (2)

En présentant les résultats de cette manière, on fait de façon sous-jacente les hypothèses suivantes:

- On suppose que la grandeur mesurée est effectivement gaussienne.
- On prend la moyenne de la distribution pour la valeur “réelle” de la grandeur $x$ et l’écart-type $\sigma$ pour l’”’erreur” sur la variable $x$..

Attention, cela ne signifie pas que la valeur “réelle” varie de $x_{\text{min}} = \mu - \sigma$ à $x_{\text{max}} = \mu + \sigma$.
L’interprétation est de type probabiliste.

## PROPRIÉTÉS DE LA GAUSSIENNE CENTRÉE RÉDUITE

La gaussienne de moyenne $\mu$ et d'écart-type $\sigma$ a pour loi

$$x \in \mathbb{R} \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{41}$$

Cela signifie que la probabilité que la valeur de $x$ soit dans l'intervalle $[x_1, x_2]$ est

On a les probabilités suivantes

- $\mathcal{P}[\mu - \sigma, \mu + \sigma] \simeq 68.27$
- $\mathcal{P}[\mu - 2\sigma, \mu + 2\sigma] \simeq 95.45$
- $\mathcal{P}[\mu - 3\sigma, \mu + 3\sigma] \simeq 99.73$

Autrement dit, une variable $x$ qui suit une loi gaussienne a environ $2/3$ de chances de se trouver dans l'intervalle $[\mu - \sigma, \mu + \sigma]$.

## SURFACE SOUS LA GAUSSIENNE

- L'aire totale sous la Gaussienne centrée réduite est $1$.
- La courbe normale s'étend indéfiniment dans les deux directions. Elle s'approche mais ne touche jamais l'axe des $x$
- La Gaussienne centrée réduite est symétrique par rapport à $0$.
- Presque toute la surface de la courbe est comprise entre $-3$ et $3$.

## NIVEAU DE CONFIANCE, INTERVALLE DE CONFIANCE

Au lieu de décrire des données expérimentales par leur fonctions de distributions de tel ou tel type (Gaussienne, Weibull,.,.), on peut utiliser la notion de probabilité comme langage unificateur. Considérons par exemple la distribution gaussienne de paramètres $(\mu, \sigma)$. Avec $x_1 = \mu - r\sigma$ et $x_2 = \mu + r\sigma$, la probabilité que $X$ soit dans l'intervalle $[x_1, x_2]$ est

$$\mathcal{P}_r = \frac{1}{\sqrt{2\pi\sigma}} \int_{x_1}^{x_2} \exp(-\frac{(x-\mu)^2}{2\sigma^2})dx = \frac{1}{\sqrt{2\pi}} \int_{-r}^{r} \exp(-\frac{z^2}{2})dz \qquad (42)$$

Au lieu de caractériser la variable $X$ par $\mu$ et $\sigma$, on la décrit par l'intervalle $[x_1, x_2]$ et par la probabilité que $X$ se trouve dans cet intervalle.

### Définition

*L'**intervalle de confiance** est $[x_1, x_2]$, il est paramétré par $r$. Le **niveau de confiance** est la probabilité $\mathcal{P}_r$ correspondante.*

## Cas de la Gaussienne $N(0,1)$

Pour $r$ fixé, le tableau suivant donne le niveau de confiance

| r | 0.0 | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}_r$ | 0.00 | 38.29 | 68.27 | 86.64 | 95.45 | 98.76 | 99.73 | 99.95 |

Soit $X \sim N(0,1)$. On a

- $\mathcal{P}_{0.5} = 0.3829$ = probabilité que $X$ soit dans $(-0.5, 0.5)$.
- $\mathcal{P}_2 = 0.9545$ = probabilité que $X$ soit dans $(-2, 2)$.
- $\mathcal{P}_{3.5} = 0.9995$ = probabilité que $X$ soit dans $(-3.5, 3.5)$.

# 7- DISTRIBUTION BINOMIALE

Une série d'expériences aléatoires est un **tirage de Bernoulli** si

- L'expérience possède deux résultats possibles notés "1" (succès) et "0" (échec).
- Les titages sont indépendants.
- la probabilité de succès $p, 0 \leq p \leq 1$ est constante au cours des tirages.

La variable aléatoire $X_N$= nombre de succès après $N$ tirages de Bernoulli s'appelle la **loi binomiale**.

$$P_N(X = x) = \binom{N}{x} p^x (1-p)^{N-x} \tag{43}$$

$$\text{Moyenne :}\bar{n} = Np, \quad \text{Variance :}\sigma^2 = Np(1-p) \tag{44}$$

# DISTRIBUTION BINOMIALE (2)

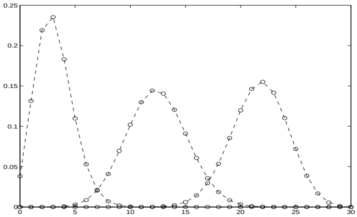



Figure: Loi binomiale pour $p = 0.1$, $p = 0.4$, $p = 0.7$, $N = 30$

## DISTRIBUTION BINOMIALE (3)

La moyenne de la variable $n$ (nombre d'apparitions de l'évènement $A$ sur $N$ tirages) est donc proportionnel à $N$ ( le nombre de tirages). De plus, l'écart-type $\sigma$ de $n$ est proportionnel à $\sqrt{N}$. On effectue à présent le quotient

$$\delta = \frac{\sigma}{\bar{n}} \tag{45}$$

c'est-à-dire le rapport entre l'écart-type sur $N$ tirages par la valeur moyenne sur ces mêmes $N$ tirages. La quantité $\delta(N)$ mesure l'erreur relative sur l'évènement mesuré (nombre d'apparitions de l'évènement $A$ sur $N$ tirages). On a

$$\delta = \frac{\sqrt{Np(1-p)}}{Np} = \frac{K}{\sqrt{N}} \tag{46}$$

où $K$ est une constante.

- Ceci confirme l'intuition que plus on fait de mesures (plus $N$ est grand) , plus la précision est grande.
- La précision dépend de $N$ par un facteur $1/\sqrt{N}$. Cela signifie que pour diminuer l'incertitude de 10 il faut faire 100 fois plus d'expériences.

## DISTRIBUTION BINOMIALE (4)

Identification d'une situation concrète de type “loi binomiale”.

- N tirages sont effectués
- Il n'y a que deux résultats possibles s'identifiant à “succès” et “échec”.
- Les tirages sont indépendants.
- La probabilité de succès reste la même de tirage en tirage.

## DISTRIBUTION BINOMIALE (5)

Procédure “distribution binomiale”:

1. Identifier la notion de succès.
2. Déterminer $p =$ probabilité de succès.
3. Déterminer $N =$ nombre de tirages.
4. En déduire la loi binomiale:

$$\text{Pour } x \text{ entier }, P_N(X = x) = \left( \begin{array}{c} N \\ x \end{array} \right) p^x (1-p)^{N-x} \tag{47}$$

# 8 - DISTRIBUTION DE POISSON

Sert à modéliser la fréquence avec laquelle un évènement particulier intervient dans une certaine période de temps.

- Nombre de patients arrivant aux urgences entre 6h et 7h
- Nombre d'appels téléphoniques reçus à un standard.
- Nombre de particules $\alpha$ émises par une substance radioactive par minutes.
- Nombre de colonies bactériennes apparaissant dans une boîte de Petri pendant un laps de remps $\Delta t$.

# DISTRIBUTION DE POISSON (2)

C'est une **loi discrète**: prend des valeurs entières.

$$P(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2, 3... \tag{48}$$

- Paramètre unique $\lambda$. Moyenne et écart-type:

$$\mu = \lambda, \quad \sigma = \sqrt{\sigma} \tag{49}$$

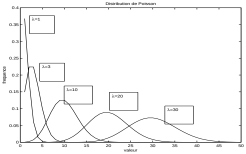
- 


Figure: Distribution de Poisson

## DISTRIBUTION DE POISSON (3)

On suppose que le nombre de patients qui arrivent aux urgences entre 6h et 7h du soir suit une loi de Poisson de paramètres $\lambda = 6.9$.

- Probabilité que le nombre de patients qui arrivent soit exactement 4:

$$P(X = 4) = e^{-6.9}\frac{6.9^4}{4!} = 0.095 \tag{50}$$

Il y a 9.5% de chances qu'il y ait 4 patients qui arrivent entre 6h et 7h.

- Probabilité qu'au plus 2 patients qui arrivent:

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= e^{-6.9}\left(\frac{6.9^0}{0!} + \frac{6.9^1}{1!} + \frac{6.9^2}{2!}\right) = 0.032 \end{aligned}$$

Il y a 3.2% de chances qu'il y ait moins de 2 patients qui arrivent entre 6h et 7h.

- $\mu = 6.9$, $\sigma = 2.6$. En moyenne 6.9 patients arrivent aux urgences, avec écart-type de 2.6.

# 9- RELATION ENTRE LOI BINOMIALE ET LOI DE POISSON

Lorsque $N$=nombre de tirages dans la loi binomiale est grand, les valeurs des probabilités données par la formule

$$P_N(X = x) = \left( \begin{array}{c} N \\ x \end{array} \right) p^x (1-p)^{N-x} \tag{51}$$

deviennent difficiles à calculer. On a recours à une approximation. Lorsque $N \to +\infty$ et $p$ petit avec $Np = \lambda$ fixé, on a

$$\left( \begin{array}{c} N \\ x \end{array} \right) p^x (1-p)^{N-x} \simeq e^{-\lambda} \frac{(np)^x}{x!} \quad \text{loi de Poisson} \tag{52}$$

## RELATION ENTRE LOI BINOMIALE ET LOI DE POISSON (2)

Procédure pratique

1. Trouver $N$= nombre de tirages et $p$= probabilité de succès.
2. Conditions pratiques pour effectuer l'approximation:

$$N \geq 100, \quad \text{et } Np \leq 10 \tag{53}$$

3. Effectuer l'approximation

$$P(X = x) \simeq e^{-np} \frac{(Np)^x}{x!} \tag{54}$$

Exemple: Mortalité infantile en Suède.

# 10 - RELATION ENTRE LOI BINOMIALE ET LOI NORMALE

Comme précédemment, considérons une loi binomiale avec grand nombre de tirages ($N$ grand).
Exemple de calcul de table de mortalité (assurances). On sait que 80% des personnes âgées de 20 ans atteignent l'âge de 65 ans.
Trouver la probablité pour que entre 375 et 425 personnes parmi 500 soient en vie à 65 ans.
Le calcul direct par la loi binomiale donne

$$P(375 \leq X \leq 425) = \sum_{x=375}^{425} = \left( \begin{array}{c} 500 \\ x \end{array} \right) 0.8^x 0.2^{500-x} \tag{55}$$

n'est pas pratiquable.

## RELATION ENTRE LOI BINOMIALE ET LOI NORMALE (2)

Procédure pratique

1. Trouver $N$= nombre de tirages et $p$= probabilité de succès.
2. Conditions pratiques pour effectuer l'approximation:

$$Np \geq 5, \quad \text{et} N(1-p) \leq 5 \tag{56}$$

3. Calculer moyenne $\mu$ et écart-type $\sigma$

$$\mu = Np, \quad \sigma = \sqrt{Np(1-p)} \tag{57}$$

4. Trouver l'aire sous la courbe de la loi normale $N(\mu, \sigma)$ correspondant aux probabilités demandées.