TRAITEMENT STATISTIQUE DES DONNÉES

S1 - Master PHEAPC

Chapitre 1 : Rappels Générеaux

Pr. M. EL KACIMI[1]

[1]Université Cadi Ayyad
Faculté des Sciences Semlalia
Département de Physique

Année universitaire 2019/2020

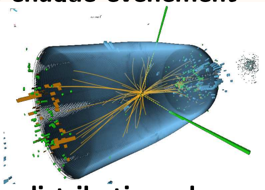
Chapitre 1 : Rappels Génèreaux

1. **Introduction**
2. **Analyse combinatoire**
3. **Probabilités : Définition et propriétés**
4. **Probabilités conditionnelles**
5. **Théorème de Bayes**
6. **Variables aléatoires et loi de probabilité**
7. **Fonction de densité de probabilité : Propriétés**
   - **Fonction de partition**
   - **Fonctions de densité marginale**
   - **Fonction de densité de probabilité conditionnelle**
8. **Changement de variable**
   - **Cas d'une seule v.a.**
   - **Cas de plusieures v.a.**
9. **Propriétés des distributions**
   - **Fonction génératrice des moments**
10. **Propagation des erreurs**

Chapitre 1 : Rappels Généreaux

Introduction

On observe des évènements d'un certain type et on mesure les propriétés de chaque évènement



- Moment de la particule ;
- Nombre de muons dans l'évènement ;
- Energie des jets ;
- Energie déposée dans le CM $\cdots$;

Les distributions de ces mêmes propriétés sont prédites par les modèles théoriques, tel que le modèle standard (MS), et ce en fonction des paramètres libres, comme pour l'exemple α, G_F, M_Z, α_Z, M_H, $\cdots$

L'analyse des données permet, entre autres, de :

- Estimer les paramètres libres du modèle ;
- Quantifier les erreurs sur les paramètres estimés avec un niveau de confiance donné ;
- Tester jusqu'à quelle limite une théorie décrit bien les données.

Introduction

Composer avec les incertitudes

En physique des particules, nombre de causes peuvent être sources des incertitudes :

- **la théorie n'est pas déterministe : mécanique quantique ;**
- **erreurs aléatoires (dites statistiques) des mesures : présentes même sans l'effet quantique**
- **méconaissance de certaines choses : limitations des prix, temps, $\cdots$**

Récapitulatif

Nous pouvons quantifier les incertitudes en utilisant les probabilités

Chapitre 1 : Rappels Génèreaux

Analyse combinatoire

Rappels de quelques cas

Dénombrement des dispositions possibles des éléments d'un ensemble fini :
Soient les ensembles A, B,$\cdots$,S avec card$(A) = n_a$, card$(B) = n_b$, $\cdots$, card$(S) = n_s$.
On peut distinguer

	Sans remise	Avec remise
Order important	Arrangement sans répétition	Arrangement avec répétition
Ordre sans importance	Combinaisons sans répétition	Combinaisons avec répétition

Paires(x_a, x_b)	:	$n_a \times n_b$
Multiplets$(x_a, x_b, \cdots, x_s)$	:	$n_a \times n_b \times \cdots \times n_s$
Arrangements de p parmi n avec répétition	:	n^p
Arrangements de p parmi n sans répétition	:	$n!/(n-p)!$
Permutations de n sans répétition	:	$n!$
Permutations de n avec répétition	:	$n!/(n_1! \times \cdots \times n_s!$
Combinaisons de p parmi n sans répétition	:	$n!/[p!(n-p)!]$
Combinaisons de p parmi n avec répétition	:	$(n+p-1)!/[p!(n-1)$

Analyse combinatoire

Rappels de quelques cas

Partitions généralisées : on introduit cette notion par l'exemple suivant

	A	B	C	D	Total
♣	3	5	4	1	13
♠	3	3	4	3	13
♡	5	1	2	5	13
♢	2	4	3	4	13
Total	13	13	13	13	52

$\Longrightarrow A$: $n_A = 13!/(3! \times 3! \times 5! \times 2!)$

$\Longrightarrow B$: $n_B = 13!/(5! \times 3! \times 1! \times 4!)$

$\Longrightarrow C$: $n_C = 13!/(4! \times 4! \times 2! \times 3!)$

$\Longrightarrow D$: $n_D = 13!/(1! \times 3! \times 5! \times 4!)$

Le nombre total de possibilités est donné par $N = n_A \times n_B \times n_c \times n_D$..

Chapitre 1 : Rappels Généreaux

Probabilités

Définition

$\mathcal{P}$ **est une loi de probabilité sur** Ω **si :**

Axiome de positivité

$$\mathcal{P}(X_i) \geq 0 \quad \forall X_i \in \Omega;$$

Axiome d'additivité

$$\mathcal{P}(X_i \ \textbf{ou} \ X_j) = \mathcal{P}(X_i) + \mathcal{P}(X_j);$$

Axiome de certitude

$$\sum_{\Omega} \mathcal{P}(X_i) = 1;$$

Axiomes de Kolmogorov 1933

Propriétés

$$\begin{array}{rcl}
\mathcal{P}(\overline{A}) & = & 1 - \mathcal{P}(A) \\
\mathcal{P}(A \cup \overline{A}) & = & 1 \\
\mathcal{P}(\emptyset) & = & 0 \\
\textbf{if} \quad A \subset B & \Longrightarrow & \mathcal{P}(A) \leq \mathcal{P}(B) \\
\mathcal{P}(A \cup B) & = & \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)
\end{array}$$

Probabilités

Propriétés : Loi d'addition

Soit $A \subset \Omega$, $A = \{\cdots, X_i, \cdots\}$.

card(A)=1$\Longrightarrow$ A évènement élémentaire.

On généralise la loi d'addition $\mathcal{P}(A_1 \cup A_2) = \mathcal{P}(A_1) + \mathcal{P}(A_2) - \mathcal{P}(A_1 \cap A_2)$ au cas où $A_i, i = 1, \cdots, n$ et $A_i \cap A_j \neq \varnothing$, et on utilise les notations suivantes

$\mathcal{P}_i = \mathcal{P}(A_i)$, $\mathcal{P}_{ijk} = \mathcal{P}(A_i \text{ et } A_j \text{ et} A_k)$ avec $i < j < k$, $S_1 = \sum_i \mathcal{P}_i$,
$S_2 = \sum_{i<j} P_{ij}$, $S_3 = \sum_{ijk} \mathcal{P}_{ijk}$, $\cdots$,

la probabilité pour qu'au moins un évènement se produise, appelée aussi la formule du crible de Poincaré, est

$$\mathcal{P}(A_1 \text{ ou } A_2 \cdots \text{ ou } A_n) = S_1 - S_2 + S_3 + \cdots - (-1)^n S_n.$$

Chapitre 1 : Rappels Généreaux

Probabilités conditionnelles

Définition

La probabilité conditionnelle d'avoir A sachant que B est réalisé est définie par

$$\boxed{\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}}$$

que l'on peut écrire sous la forme

$$\mathcal{P}(A \cap B) = \mathcal{P}(A|B)\mathcal{P}(B) = \mathcal{P}(B|A)\mathcal{P}(A)$$

Deux évènements sont indépendants si $\mathcal{P}(A \cap B) = \mathcal{P}(A) \times \mathcal{P}(B)$.

A ne pas confondre avec $A \cap B = \varnothing$.

La probabilité conditionnelle de réalisation de A sachant B dans ce cas est

$$\mathcal{P}(A|B) = \mathcal{P}(A) \quad \text{et} \quad \mathcal{P}(B|A) = \mathcal{P}(B)$$

Différentes écoles de probabilités

Fréquentiste ou Bayesienne ?

Fréquentistes :

Les évènements peuvent être répétés :

$$\mathcal{P}(A) = \lim_{N \to +\infty} \frac{N_A}{N}$$

Mécanique quantique, diffusion, radioactivité, $\cdots$

Bayesienne :

La probabilité est une notion subjective. Les évènements sont des hypothèses et la probabilité d'un évènement A est définie par le degré de conviction que A soit vrai.

Les deux interprétations sont consistantes avec les axiomes de Kolmogorov.

Chapitre 1 : Rappels Généreaux

Théorème de Bayes

Enoncé

Théorème de Bayes :

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A) \times \mathcal{P}(A)}{\mathcal{P}(B)}.$$

Révérend T. Bayes (1702/1761)

Soient $(H_1, \cdots, H_i, \cdots)$ **Hypothèses exclusives et exhaustives** $H_i \cap H_j = \emptyset$ **et** $\cup_i H_i = \Omega$**. Alors**

$$B = B \cap (\cup_i H_i) = \cup_i (B \cap H_i)$$

et le théorème devient

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A) \times \mathcal{P}(A)}{\sum_i \mathcal{P}(B \cap H_i)} = \frac{\mathcal{P}(B|A) \times \mathcal{P}(A)}{\sum_i \mathcal{P}(B|H_i) \times \mathcal{P}(H_i)}$$

Théorème de Bayes

Cas d'étude d'hypothèses

Dans le cas de l'étude d'hypothèses au lieu d'évènements, le théorème de Bayes est bien adapté.

Soient $(\theta_1, \cdots, \theta_n)$ n hypothèses exclusives 2 à 2 et exhaustives. L'on cherche à tester une hypothèse à partir d'un échantillon de données observées X^0, alors :

$$\mathcal{P}(\theta_i|X^0) = \frac{\mathcal{P}(X^0|\theta_i)\mathcal{P}(\theta_i)}{\mathcal{P}(X^0)}$$

- $\mathcal{P}(\theta_i|X^0)$: la probabilité à postériori de l'hypothèse θ_i, sachant X^0.
- $\mathcal{P}(X^0|\theta_i)$: la probabilité d'obtenir les données X^0 dans le cadre de l'hypothèse θ_i.
- $\mathcal{P}(\theta_i)$: la probabilité à priori de θ_i, qui représente le degré de conviction dans cette hypothèse avant l'expérience.
- $\mathcal{P}(X^0)$ est la probabilité d'obtenir les données X^0 quelque soit θ_i. Elle peut être considérée comme un facteur de normalisation où $\mathcal{P}(X^0) = \sum \mathcal{P}(X^0|\theta_j)\mathcal{P}(\theta_j)$.

$$\mathcal{P}(\theta_i|X^0) \propto \mathcal{P}(X^0|\theta_i)\mathcal{P}(\theta_i)$$

Si $\mathcal{P}(X^0)$ est inconnue, on ne peut pas calculer la probabilité à postériori de θ_i mais on peut calculer le rapport suivant

$$\frac{\mathcal{P}(\theta_i|X^0)}{\mathcal{P}(\theta_j|X^0)}$$

Chapitre 1 : Rappels Généreaux

Variables aléatoires et loi de probabilité

Variables aléatoires

Définition

Soit Ω l'ensemble des éventualités, dit aussi univers. La v.a. X est une application définie comme suit

$$\begin{array}{llll} X & : \Omega & \rightarrow & \mathcal{R} \\ \forall e_i \in \Omega & \quad e_i & \rightarrow & X(e_i) \in \mathcal{R} \end{array}$$

X peut être continue ou discrète et prend ses valeurs dans $D_X = \{x_1, \cdots, x_n\}$ où $n = Card(\Omega)$.

Variables aléatoires et loi de probabilité

Loi de probabilité

Définition

X **est une v.a. prenant ses valeurs dans** $D_X = \{x_1, \cdots, x_n\}$**. La loi de probabilité associée à** X **est définie comme suit**

$$\begin{array}{cccl} \mathcal{P} & : & D_X & \rightarrow & \mathcal{R} \\ & & x_i & \rightarrow & P(X = x_i) = P(e_i) = p_i \end{array}$$

Propriété

$$\sum_i \mathcal{P}(X = x_i) \quad = \quad \sum_i p_i = 1$$

En effet, $\sum_i \mathcal{P}(e_i) = \mathcal{P}(\cup_i e_i)$ **car** e_i **sont disjoints et comme** $\cup_i e_i = \Omega \Longrightarrow \mathcal{P}(\cup_i e_i) = \mathcal{P}(\Omega) = 1.$

Chapitre 1 : Rappels Génréreaux

Fonction de densité de probabilité : fdp

Définition, Histogramme

Définition

La probabilité par unité de longueur de x est définie comme

$$f(X=x) \quad = \quad \lim_{\Delta x \to 0} \frac{P(X=x)}{\Delta x} \quad \text{et} \quad d\mathcal{P}(x \in [x, x+dx]) = f(x)dx$$

Propriété : $\int_{D_X} f(x)dx = 1.$

Représentation par un histogramme :

On répartit les données en N_t classes (bin) entre x_{min} et x_{max} : choix du pas à optimiser(!) $\Delta x = (x_{max} - x_{min})/N_t$.

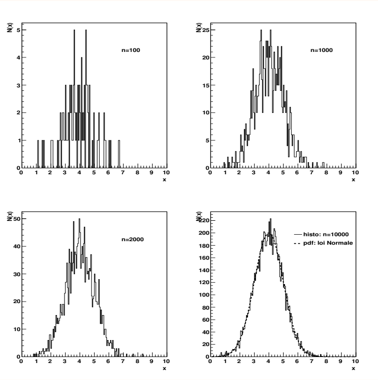
Le bin i d'abscisse $x = x_{min} + (i-1) \times \Delta x$, comprend $N(x)$ évènements : alors

$$f(x) \quad = \quad \frac{N(x)}{N_{tot}\Delta x}$$

Fonction de densité de probabilité : fdp

Exemple Histogramme

Exemple de données présentées sous forme d'histogramme :

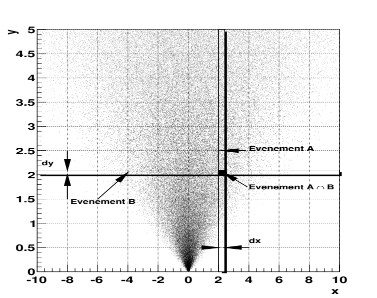


- Données suivant une loi normale : Autant $N_t = n$ est grand autant les données sont bien lissées

$\Longrightarrow$ Choix du pas (binning) : faire apparaître l'allure de la pdf.

Fonction de densité de probabilité : fdp

Cas de Plusieures variables



Si l'on prend $A \cap B$ infinitésimal, on peut définir la probabilité élémentaire
$\Longrightarrow d\mathcal{P}(A \cap B) = f(X = x, Y = y)dxdy$
ce qui permer d'en déduire la densité jointe de probabilité dans le cas de deux variables par

$$f(X = x, Y = y) = \lim_{\Delta x \to 0, \Delta y \to 0} \frac{\mathcal{P}(A \cap B)}{\Delta X \Delta Y}$$

Et l'on peut déduire que la définition de la densité jointe de probabilité dans le cas de plusieurs variables est donnée par :

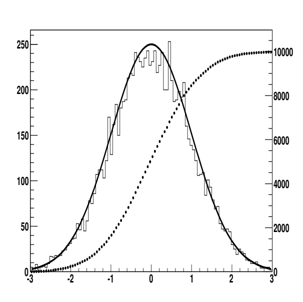
$$f(X_1 = x_1, ..., X_n = x_n) \quad = \quad \lim_{\Delta x_1 \to 0, \ldots, \Delta x_n \to 0} \frac{\mathcal{P}(A_1 \cap ... \cap A_n)}{\Delta x_1 ... \Delta x_n}$$

Fonction de répartition

Une variable aléatoire $X \in [X_{min}, X_{max}]$ est bien décrite soit par sa fonction de densité de probabilité $f(X)$ soit par sa fonction de répartition $F(X)$ définie par

$$F(X) = \int_{Xmin}^{X} f(X')dX'.$$

Dans le cas d'une v.a. discrète, $F_k = \mathcal{P}(i \leq k) = \sum_{i \leq k} \mathcal{P}(k)$.



Quelques propriétés :

- $F(X_{min}) = 0$ et $F(X_{max}) = 1$.
- $F(X)$ est monotone dans $[X_{min}, X_{max}]$;
- $F(X) = P(X' \leq X)$.
- $F(x \in [X_1, X_2]) = F(X_2) - F(X_1)$.

En général, la probabilité que les variables aléatoires X et Y prennent leurs valeurs dans $D(X, Y)$ est

$$P(X, Y \in D) = \iint_{D(x,y)} dF(X, Y).$$

Fonctions de densité marginale

Cas de deux variables

Considérons le cas de deux variables : $X \in [X_{min}, X_{max}]$ et $Y \in [Y_{min}, Y_{max}]$; Soit $f(X, Y)$ la densité jointe de probabilité.

Densité marginale $f_x(X)$ de X : Projection de $f(X, Y)$ sur l'axe OX :

$$f_x(X) = \int_{Y_{min}}^{Y_{max}} f(X, Y) dY.$$

Densité marginale $f_Y(Y)$ de Y : Projection de $f(X, Y)$ sur l'axe OY :

$$f_Y(Y) = \int_{X_{min}}^{X_{max}} f(X, Y) dX.$$

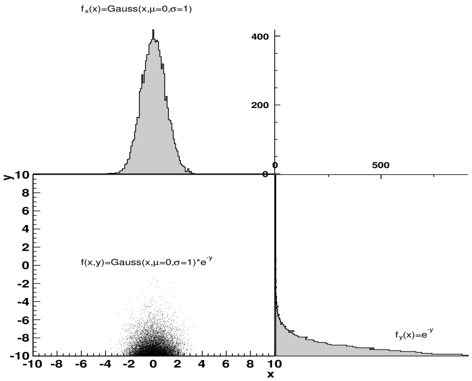
Dans le cas des variables discrètes, on obtient

$$\mathcal{P}_k(k) = \sum_j \mathcal{P}(k, j).$$

Fonctions de densité marginale

Exemple : $f(X, Y) = 1/\sqrt{2\pi}e^{-x^2/2} \times e^{-y}$

La figure ci-dessous illustre le cas d'une pdf jointe $f(x, y) = 1/\sqrt{2\pi}e^{-x^2/2}e^{-y}$ ainsi que les distributions marginales qui lui sont associées.



Fonctions de densité de probabilité conditionnelle

Définition

Evènement A **:** $x \in [x, x+dx]\ \forall y \Longrightarrow d\mathcal{P}(A) = f_x(x)dx$
Evènement B **:** $y \in [y, y+dy]\ \forall x \Longrightarrow d\mathcal{P}(B) = f_y(y)dy$
Evènement $A \cap B$ **:** $(x, y) \in [x, x+dx] \times [y, y+dy] \Longrightarrow d\mathcal{P}(A \cap B) = f(x, y)dxdy$

La probabilité conditionnelle de B sachant A est donnée par

$$\begin{aligned} d\mathcal{P}(B|A) &= \frac{d\mathcal{P}(A \cap B)}{d\mathcal{P}(A)} = \frac{f(x,y)dxdy}{f_x(x)dx} = \frac{f(x,y)}{f_x(x)}dy \\ d\mathcal{P}(A|B) &= \frac{d\mathcal{P}(A \cap B)}{d\mathcal{P}(B)} = \frac{f(x,y)dxdy}{f_y(y)dy} = \frac{f(x,y)}{f_y(y)}dx \end{aligned}$$

Aussi, les fonctions de densité conditionnelle $h(y|x)$ et $g(x|y)$ sont données par

$$\begin{aligned} h(y|x) &= \lim_{\Delta y \to 0} \frac{d\mathcal{P}(B|A)}{\Delta y} \\ &= \frac{f(x,y)}{f_x(x)}, \quad g(x|y) = \frac{f(x,y)}{f_y(y)} \end{aligned}$$

Chapitre 1 : Rappels Généreaux

Changement de variable

Cas d'une seule v.a X : $Y = h(X)$

- $h(X)$ est bijective

$f(X)$ étant la pdf de X. On cherche $g(Y)$ la pdf de Y.

Nous avons $\mathcal{P}(y \in [y, y+dy]) = \mathcal{P}(x \in [x+dx])$ ce qui donne $g(y)dy = f(x)dx$.

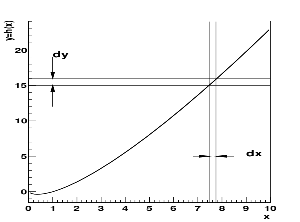
On déduit alors que

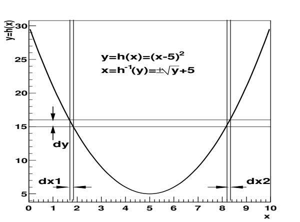
$$g(y) = f(h^{-1}(y))|\frac{1}{dy/dx}| = \frac{f(h^{-1}(y))}{|h'(x)|}$$

- $h(X)$ est non bijective. Considérons le cas illustré par la figure ci-contre où $h^{-1}(y) = \{x_1, x_2\}$ et $dx_1 = dx_2 = dx$. Nous obtenons

$$\begin{aligned} g(y)dy &= f(x_1)dx_1 + f(x_2)dx_2 \\ \Longrightarrow g(y) &= \frac{(f(x_1) + f(x_2))}{|h'(h^{-1}(y))|} \\ &= \frac{(f(-\sqrt{y}+5) + f(\sqrt{y}+5))}{2\sqrt{y}} \end{aligned}$$

Dans le cas où $h^{-1}(y) = \{x_1, \cdots, x_n\}$, alors $g(y) = \frac{\sum_{i=1}^{n} f(x_i)}{|h'(x)|}$

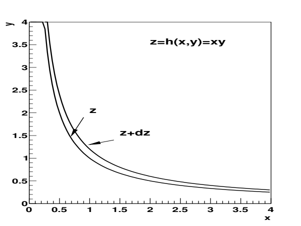




Changement de variable

Cas de plusieurs v.a $\vec{x} = (x_1, \cdots, x_n)$ et $z = h(\vec{x})$

Prenons le cas de $\vec{x} = (x, y)$ de pdf jointe $f(x, y)$. Soit dS l'élément de surface délimité par les hyper-plans d'équations z et $z + dz$. Ainsi la probabilité $\mathcal{P}(z \in [z, z + dz]) = \int\int_{dS} f(x, y)dxdy$ ce qui donne



$$\begin{aligned}
g(z)dz &= \int_0^{+\infty} dx \int_{z/x}^{(z+dz)/x} f(x, y)dy = \int_0^{+\infty} dy \int_{z/y}^{(z+dz)/y} f(x, y)dx \\
&= \int_0^{+\infty} dx f(x, \frac{z}{x})\frac{dz}{x} = \int_0^{+\infty} dy f(\frac{z}{y}, y)\frac{dz}{y} \\
\Longrightarrow \quad & g(z) = \int_0^{+\infty} f(x, \frac{z}{x})\frac{dx}{x} = \int_0^{+\infty} f(\frac{z}{y}, y)\frac{dy}{y}
\end{aligned}$$

Dans le cas général où z est une fonction de n variables aléatoires, $g(z)dz = \int_{dS} f(x_1, ..., x_n)dx_1...dx_n$.

Changement de variable

Cas de plusieurs v.a $\vec{x} = (x_1, \cdots, x_n)$ et $\vec{z} = \vec{z}(\vec{x})$

Soit $\vec{x}(x_1, \cdots, x_n) \Longrightarrow f(\vec{x})$ et $\vec{z} = (z_1(\vec{x}), \cdots, z_n(\vec{x}))$. Chaque $z_i(\vec{x})$ a une fonction réciproque $x_i(\vec{z})$.
On définit la pdf jointe $g(\vec{z})$. Le but est de chercher les pdf marginales $g_i(z_i)$

Première étape : déterminer $g(\vec{z})$:

$$
\begin{aligned}
g(\vec{z}) &= |J| f(\vec{x}(\vec{z})) \\
&= \begin{vmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} & \cdots & \frac{\partial x_1}{\partial z_n} \\ \vdots & & & \vdots \\ \frac{\partial x_n}{\partial z_1} & \frac{\partial x_n}{\partial z_2} & \cdots & \frac{\partial x_n}{\partial z_n} \end{vmatrix} f(x_1(\vec{z}), ..., x_n(\vec{z}))
\end{aligned}
$$

où $|J|$ est le jacobien du changement de variables.
Exemple : $(x_1 = x, x_2 = y) \rightarrow (z_1 = x, z_2 = xy)$:
Le jacobien est alors égal à :

$$
|J| = \begin{vmatrix} \frac{\partial x}{\partial z_1} & \frac{\partial x}{\partial z_2} \\ \frac{\partial y}{\partial z_1} & \frac{\partial y}{\partial z_2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -\frac{y}{x} & \frac{1}{x} \end{vmatrix} = \frac{1}{z_1} \Longrightarrow g(z_1, z_2) = \frac{1}{z_1} f(z_1, \frac{z_2}{z_1}).
$$

Changement de variable

Cas de plusieurs v.a $\vec{x} = (x_1, \cdots, x_n)$ et $z = h(\vec{x})$

Deuxième étape :

On détermine $g_i(z_i)$ en intégrant sur toutes les variables z_j pour $i \neq j$:

$$g_i(z_i) = \int |J| f(x_1(\vec{z}), \cdots, x_n(\vec{z})) \prod_{j \neq i}^{n} dz_j$$

Reprenons l'exemple précédent $(x_1 = x, x_2 = y) \to (z_1 = x, z_2 = xy)$:

Cherchons à calculer $g_2(z_2)$:

$$g_2(z_2) dz_2 = dz_2 \int |J| f(x, y) dz_1$$

$$\Longrightarrow \quad g_2(z_2) = \int f(z_1, \frac{z_2}{z_1}) \frac{dz_1}{z_1} = \int f(x, \frac{z_2}{x}) \frac{dx}{x}$$

ce qui n'est d'autre que le résultat obtenu précédamment.

Chapitre 1 : Rappels Génèreaux

Propriétés des distributions

Espérance mathématique : cas de v.a. continue

Soit X une variable aléatoire de fdp $f(X)$.

La densité de probabilité est utilisée comme une fonction de pondération dans le calcul de la valeur moyenne d'une fonction $g(X)$. En effet

$$E[g(X)] = \int_{D(X)} g(X) f(X) dx$$

L'espérance mathématique est un opérateur linéaire :

$$E[\alpha g(X) + \beta] = \alpha E[g(X)] + \beta.$$

Si $g(X) = X$, il s'agit alors de l'espérance ou de la valeur moyenne de X qui est définie par

$$\mu = \int_{D(X)} X f(X) dx.$$

Exception : la loi de Cauchy $f(x) = \frac{1}{\pi(1+x^2)}$ n'admet pas d'espérance mathématique.

Propriétés des distributions

Espérance mathématique : cas de v.a. discrète

Soit X_i une variable aléatoire discrète dont la loi de probabilité est $\mathcal{P}(x_i)$.

La densité de probabilité est utilisée comme une fonction de pondération dans le calcul de la valeur moyenne d'une fonction $g(X)$. En effet

$$E[g(X)] = \sum_i g(x_i)\mathcal{P}(x_i)$$

Si $g(X) = X$, il s'agit alors de l'espérance ou de la valeur moyenne de X qui est définie par

$$\mu = \sum_i x_i \mathcal{P}(x_i).$$

L'expérance mathématique ou valeur moyenne est notée par $\overline{X}$ ou $< X >$.

Propriétés des distributions

Autres moyennes autres définitions ...

La moyenne définie auparavant est appelée la moyenne arithmétique.

D'autres moyennes sont définies mais de moindre importance :

$\Longrightarrow$ **Moyenne géométrique :**

$$\overline{X}_G = \left(\prod_i x_i^{n_i}\right)^{\frac{1}{N}} = \prod_i x_i^{p_i} \text{ telle que } p_i = \mathcal{P}(x_i) = n_i/N.$$

$$\ln(\overline{X}_G) = \sum_i p_i \ln x_i$$

$\Longrightarrow$ **Moyenne harmonique :**

$$\frac{1}{\overline{X}_H} = \sum_i \frac{p_i}{x_i} \text{ et en cas de v.a. continue } \frac{1}{\overline{X}_H} = \int \frac{f(x)}{x} dx.$$

Propriétés des distributions

Autres propriétés pour localiser la distribution des valeurs d'une v.a ...

En plus de la moyenne, D'autres grandeurs permettent de décrire la localisation de la distribution des valeurs d'une v.a.

Médiane x_M : définie comme la valeur pour laquelle la moitié des valeurs de X est située à gauche de x_M et l'autre moitié à sa droite

$$\int_{-\infty}^{x_M} f(x)dx = \frac{1}{2} \quad \textbf{et pour une v.a discrète} \quad \sum_{i=0}^{i=i_m} \mathcal{P}(x_i) = \frac{1}{2}.$$

MPV x_{MPV} : est la valeur de X qui présente la probabilité la plus élevée. Dans le cas d'une v.a. continue, elle corespond au pic de la distribution.

$\Longrightarrow$ **Les quantiles** $\cdots$

Propriétés des distributions

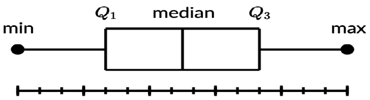
Quantiles, centiles, quartiles : diagramme en boite

Quantiles : valeur de la v.a. x_i pour une probabilité donnée p_i : $p_i = \int_{x_{min}}^{x_i} f(x)dx = F(x_i)$;

Centiles : quantiles de $p_i = n\%$ (multiples de $\%$.) ;

Quartiles : quantiles de $p_i = n \times 25\%$ (multiples de 25%) : $F(Q_1) = 0.25$, $F(Q_2) = 0.5$ et $F(Q_3) = 0.75$.

Dans un diagramme en boite, voir figure ci-contre, on représente cinq paramètre décrivant l'échantillon de données : le minimum, le premier quartile Q_1, la médiane, le troisième quartile Q_3 et le maximum.



Propriétés des distributions

Moments d'ordre supérieur

On appelle moment d'ordre r, s'il existe, et noté μ'_r par $\mu'_r = E[X^r]$.

Le moment d'ordre r est dit centré, noté par μ_r, si $\mu_r = E\left[(X-\mu)^r\right]$.

Variance d'une distribution : elle mesure la dispersion des valeurs autour de la valeur moyenne :

$$V[X] \quad = \quad \sigma^2 = E\left[(X-\mu)^2)\right] = E[X^2] - \mu^2.$$

$V[aX+b] = a^2 V[X]$.

Coefficient d'asymétrie : déduit à partir du moment centré d'ordre 3 par

$$\beta_1 \quad = \quad \frac{E\left[(X-\mu)^3\right]}{\sigma^3}$$

et il est nul lorsque la distribution est symétrique autour de μ.

Curtos : déduit à partir du moment centré d'ordre 4 par

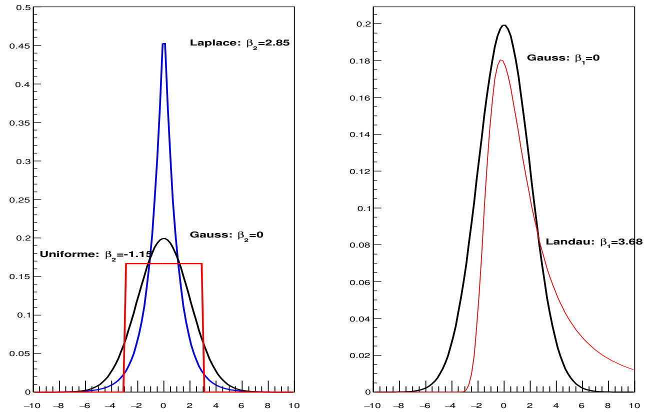
$$\beta_2 \quad = \quad \frac{E\left[(X-\mu)^4\right]}{\sigma^4} - 3$$

et mesure comment sont dispersées les valeurs de X autour de μ.

Le curtos est nul pour la distribution normale.

Propriétés des distributions

Moments d'ordre supérieur : Quelques exemples de distributions



Propriétés des distributions

Covariance et coefficient de corrélation de distributions

La définition de l'espérance peut être généralisée à une fonction jointe de densité de plusieurs variables. En effet

$$E[g(X,Y)] = \int_{D(X,Y)} g(X,Y)f(X,Y)dx.$$

Et on définit les propriétés de chacune des variables comme suit

$$\begin{aligned}
\mu_X &= \iint_{D(X,Y)} Xf(X,Y)dXdY \\
\mu_Y &= \iint_{D(X,Y)} Yf(X,Y)dXdY \\
V[X] &= \iint_{D(X,Y)} (X-\mu_X)^2 f(X,Y)dXdY \\
V[Y] &= \iint_{D(X,Y)} (Y-\mu_Y)^2 f(X,Y)dXdY
\end{aligned}$$

Propriétés des distributions

Covariance et coefficient de corrélation de distributions

Définition

La covariance entre deux v.a. X et Y de densité jointe $f(X,Y)$ est définie comme suit

$$cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$$

Dans le cas d'une variable aléatoire discrète, la covariance est donnée par

$$cov(X,Y) = \sum_{ij}(x_i-\mu_x)(y_j-\mu_y)\mathcal{P}(x_i,y_j)$$

Définition

Le coefficient de corrélation est défini par

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X\sigma_Y}$$

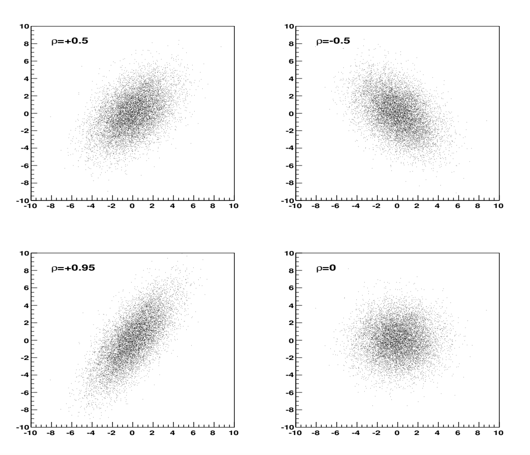
On note que $cov(X,X) = V[X]$ et $|\rho(X,Y)| \leq 1$.

Montrer que $|\rho(X,Y)| \leq 1$.

Indication : Prendre $V(\alpha X + Y) \geq 0$ et de le développer en fonction de α et de prende $\alpha = \sigma_y/\sigma_x$

Propriétés des distributions

Exemples de corrélations linéaires



Propriétés des distributions

Quelques remarques

- **Deux variables aléatoires sont indépendantes si $f(X,Y) = f_X(X) \times f_Y(Y)$ alors**

$$\begin{aligned}
E[XY] &= \iint_{D(X,Y)} XYf(X,Y)dXdY \\
&= \iint_{D(X,Y)} XYf_X(X) \times f_Y(Y)dXdY \\
&= \int_{D_X} Xf_X(X)dX \times \int_{D_Y} Yf_Y(Y)dY \\
&= E[X] \times E[Y] \Longrightarrow cov(X,Y) = 0.
\end{aligned}$$

Remarques

Deux v.a. indépendantes ont une covariance nulle. Attention : la réciproque n'est pas vraie.

Exemple : Soit Z et X deux v.a indépendantes avec $Z \in \{-1,+1\}$ avec la même probabilité $\Longrightarrow E[Z] = 0$. On pose $Y = ZX$. Nous avons $cov(X,Y) = cov(X,ZX) = E[ZX^2] - E[ZX]E[X] = E[Z]\left(E[X^2] - (E[X])^2\right) = 0$ alors que X et Y ne sont pas indépendantes.

Propriétés des distributions

Quelques remarques

Dans le cas de plusieures variables aléatoires de densité jointe $f(X_1, ..., X_n)$, on définit la matrice de covariance V associée à $f(X_1, ..., X_n)$ dont les éléments sont données par

$$V_{ij} = cov(X_i, X_j)$$

V est appelée aussi matrice de variance ou matrice d'erreurs.

$\Longrightarrow$ V est symétrique, $V_{ij} = V_{ji}$ et $V_{ii} = \sigma^2_{X_i}$.

Variance et espérance d'une fonction lineaire de v.a. $f(X_i) = \sum_i \alpha_i X_i$. Nous avons

$$\begin{aligned} E[\sum_{i=1,N} \alpha_i X_i] &= \sum_{i=1,N} \alpha_i E[X_i] \\ V[\sum_{i=1,N} \alpha_i X_i] &= \sum_{i=1,N} \alpha_i^2 V[X_i] + 2\sum_{i=1,N} \sum_{j>i} \alpha_i \alpha_j cov(X_i, X_j). \end{aligned}$$

Si les v.a. X_i sont indépendantes $\Longrightarrow V[\sum_{i=1,N} \alpha_i X_i] = \sum_{i=1,N} \alpha_i^2 V[X_i]$

Propriétés des distributions

Application

On effectue une mesure N fois et l'on obtient $(X_1, \ldots, X_N)$. Toutes les mesures ont alors la même espérance μ et la même déviation standard σ.

La grandeur recherchée G est obtenue en effectuant une moyenne arithmétique $G = \overline{X} = \sum_i X_i / N$.$G$ est alors considérée comme v.a. .

$\Longrightarrow E[G] = ?$

$$E[G] = E[\overline{X}] = \frac{1}{N} \sum_{i=1,N} E[X_i] = \frac{1}{N} \sum_{i=1,N} \mu = \frac{1}{N} N\mu = \mu$$

$\Longrightarrow V(G) = ?$

$$\begin{aligned} V[\overline{X}] &= \frac{1}{N^2} \sum_{i=1,N} V[X_i] + \frac{2}{N^2} \sum_{i=1,N} \sum_{j>i} cov(X_i, X_j) \\ &= \frac{\sigma^2}{N} + \frac{2}{N^2} \sum_{i=1,N} \sum_{j>i} cov(X_i, X_j). \end{aligned}$$

Mesures indépendantes : $\sigma(\overline{X}) = \sqrt{V[\overline{X}]} = \frac{\sigma}{\sqrt{N}}$.

Mesures non indépendantes : $V[\overline{X}] = \frac{1}{N^2} \sum_{i=1,N} V[X_i] + \frac{2}{N^2} \sum_{i=1,N} \sum_{j>i} \rho_{ij} \sigma_i \sigma_j$.

Cas de deux mesures : $V[\overline{X}] = \frac{1}{2} \sigma^2 (1 + \rho)$

Fonction génératrice des moments

Introduction & définition

L'introduction de la fonction génératrice permet de calculer de manière aisée les moments d'ordre supérieur d'une v.a. et ce lorsque celle-ci est bien définie.

Lorsque la fdp de la v.a. décroit plus rapidement à l'infini qu'une exponentielle, la fonction génératrice existe. Aussi, nous allons l'utiliser seulement pour certaines fdp.

Définition

On apelle fonction génératrice des moments d'une v.a. continue X de fdp $f(X)$, si elle existe, la fonction définie par

$$\Psi_X(t) \quad = \quad E[e^{tX}] = \int_{-\infty}^{+\infty} f(x)e^{tx}dx.$$

Dans le cas d'une v.a. discrète r de loi de probabilité $P(r) = p_r$, la fonction génératrice est $\Psi_r(t) = \sum_r p_r e^{rt} = \sum_r p_r (e^t)^r$.

On pose $Z = e^{rt}$ et $\Psi_r(t) = E[Z] = G[Z]$.

Fonction génératrice des moments

Expression des moments

Proposition

Le moment d'ordre r de la v.a. X est donné par

$$\mu'_r \quad = \quad \Psi_X^{(r)}(0).$$

où $\Psi_X^{(r)}$ est la dérivée d'ordre r de Ψ_X. En particulier, l'espérance mathématique d'une v.a. n'est d'autre que la dérivée première Ψ'_X prise à $t = 0$

Faisons le développement limité de $E(e^{tX})$, en supposant que la fonction génératrice est bien définie,

$$E(e^{tX}) \quad = \quad \int_{\infty}^{+\infty} \left[\sum_{k=0}^{\infty} \frac{(tx)^k}{k!}\right] f_X(x)dx = \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_{\infty}^{+\infty} x^k f_X(x)dx = \sum_{k=0}^{\infty} \mu'_k \frac{t^k}{k!}$$

$$\Psi_X(t) \quad = \quad \sum_{k=0}^{\infty} \mu'_k \frac{t^k}{k!} \Longrightarrow \mu'_r = \Psi_X^{(r)}(0).$$

Chapitre 1 : Rappels Génréraux

Propagation des erreurs

Cas de $f(X_1, \cdots, X_N)$:

Supposons que nous avons réalisé les mesures X_i de N variables aléatoires, respectivement, de valeurs moyennes μ_i et d'erreurs σ_i.

On cherche une grandeur $\mathcal{G}$ définie en fonction des mesures effectuées par $\mathcal{G} = f(X_1, \cdots, X_N)$. **Quelle est l'erreur sur la grandeur ?**

$\longrightarrow$ Caluler l'espérance mathématique de $\mathcal{G}$: $E(\mathcal{G})$

Pour ce faire, on développe f autour de $(\mu_1, ..., \mu_2)$ ce qui donne

$$f(X_1, ..., X_N) = f(\mu_1, ..., \mu_N) + \sum_{i=1,N} (X_i - \mu_i) \frac{\partial f}{\partial X_i}|_{\mu_i} + \mathcal{O}[(X_i - \mu_i)^2].$$

Le terme quadratique est négligeable car il est en moyenne proportionnel au carré de l'erreur.
On pose

$$\frac{\partial f}{\partial \mu_i} = \frac{\partial f}{\partial X_i}|_{\mu_i}.$$

L'espérance de $\mathcal{G}$ est donnée par

$$E(\mathcal{G}) = E(f) \simeq f(\mu_1, ..., \mu_2)$$

Quant à l'erreur, on calcule la variance de $\mathcal{G}$:

$$V[G] \simeq V\left[f(\mu_1, \cdot, \mu_N) + \sum_{i=1}^{N} \frac{\partial f}{\partial \mu_i}(X_i - \mu_i)\right] = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial \mu_i}\right)^2 V[X_i] + 2\sum_{i,j>i} \frac{\partial f}{\partial \mu_i}\frac{\partial f}{\partial \mu_j} cov(X_i, X$$

$$\Longrightarrow \sigma_G^2 = \sum_{i=1}^{N}\left(\frac{\partial f}{\partial \mu_i}\right)^2 \sigma_i^2 + 2\sum_{i,j>i} \frac{\partial f}{\partial \mu_i}\frac{\partial f}{\partial \mu_j} \rho_{ij}\sigma_i\sigma_j$$ **Equation de propagation des erreurs**

Propagation des erreurs

Cas de $\vec{f}(f_1, ..., f_m)$

On établit dans ce paragraphe l'équation de la propagation des erreurs dans le cas où l'on a m fonctions f_l des variables aléatoires X_i. Les erreurs et les corrélations des X_i sont données par V_{ij} et l'on cherche la matrice de covariance des f_l, U_{lm}, qui détermine les erreurs et les corrélations sur f_l.

On adopte la notation, deux indices répétés correspond à une sommation sur cet indice. En partant de ce que l'on a défini auparavant, on peut écrire

$$
\begin{aligned}
U_{lm} &= U(f_l(X_1, \cdots, X_N), f_m(X_1, \cdots, X_N)) \\
&\simeq cov\left(f_l(\mu_1, ..., \mu_N) + \sum_{i=1,N}(X_i - \mu_i)\frac{\partial f_l}{\partial \mu_i}, f_m(\mu_1, ..., \mu_N) + \sum_{j=1,N}(X_j - \mu_j)\frac{\partial f_m}{\partial \mu_j}\right) \\
&\simeq \frac{\partial f_l}{\partial \mu_i}\frac{\partial f_m}{\partial \mu_j} cov(X_i, X_j) \\
&\simeq A_{li} V_{ij} A^T_{jm}
\end{aligned}
$$

avec $A_{ij} = \partial f_i / \partial \mu_j$. Aussi l'équation de propagation des erreurs dans ce cas peut se mettre sous forme matricielle comme suit

$$U = AVA^T.$$

Exemple : Les coordonnées cylindriques (r, θ, φ) d'une trace sont mesurées de manière indépendantes et ce avec $\sigma_r = 0.1cm$, $\sigma_\theta = 0.1radian$ et $\sigma_z = 3cm$. Retrouver les erreurs sur les coordonnées cartésiennes.