

Analyse de la mesure

Saïda Guellati

► **To cite this version:**

| Saïda Guellati. Analyse de la mesure. Master / M1. 2007. <sfo-00256974>

HAL Id: sfo-00256974

<https://hal-sfo.ccsd.cnrs.fr/sfo-00256974>

Submitted on 30 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse de la mesure

Méthodologie

Concepts fondamentaux

Mesurage  Valeur du mesurande (grandeur particulière)

 Définir la méthode et la procédure de mesure

 Analyse des résultats de mesures

Estimation de la grandeur

Incertitude de mesure

correction de la mesure (effet systématique)

 Comment présenter un résultat de mesure?

Evaluation de Type A de l'incertitude-Type

Mesure d'une grandeur \rightarrow n observations indépendantes X_k

Moyenne arithmétique des n observations $\rightarrow \bar{X} = \frac{1}{n} \sum_{k=1}^n x_k$

Variance expérimentale $\rightarrow S^2(X_k) = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$

Ecart type expérimental $\rightarrow S(X_k)$

Ecart type expérimental de la moyenne $\rightarrow s(\bar{X}) = \frac{s(X_k)}{\sqrt{n}}$

Modélisation d'un mesurage

Le mesurande Y n'est pas mesuré directement

$$Y = f(X_1, X_2, \dots, X_N)$$

X_i grandeur d'entrée

Exemple : puissance dissipée par une résistance à T

$$P = f(V, R_0, \alpha, t) = \frac{V^2}{R_0 [1 + \alpha(T - T_0)]}$$

Evaluation de Type B de l'incertitude-Type

Pour une estimation x_i d'une variable d'entrée X_i qui n'a pas été obtenue à partir d'observations répétées, la variance estimée associée $u^2(x_i)$ ou l'incertitude-type $u(x_i)$ est obtenue par un jugement scientifique fondé sur toutes informations disponibles sur la variabilité de X_i .



- * **Résultats de mesures antérieures**
- * **Spécification du constructeur**
- * **Données fournies par un certificat d'étalonnage ou autres certificats**
- * **L'incertitude assignée à des valeurs de références provenant d'ouvrages**
- * **L'expérience ou la connaissance générales du comportement et des Propriétés des matériaux et instruments utilisés.**

Evaluation de Type B de l'incertitude-Type

la variance estimée associée $u^2(x_i)$ ou l'incertitude-type $u(x_i)$
Évaluées de cette façon sont appelés respectivement variance de Type B et incertitude-type de Type B

Remarque

Une évaluation de type B d'une incertitude-type peut être aussi faible qu'une évaluation de type A

Determination de l'incertitude type-composée

I- Grandeurs d'entrée non corrélés

$$Y = g(X_1, X_2, \dots, X_N)$$

Toutes les grandeurs d'entrée sont indépendantes

L'incertitude de y estimation du mesurande Y donc le résultat de mesurage est

$$u_c^2 = \sum_{i=1}^N \left[\frac{\partial g}{\partial x_i} \right]^2 u^2(x_i) \quad u^2(x_i) \text{ est l'incertitude - type}$$

$$\frac{\partial g}{\partial x_i} = \frac{\partial g}{\partial X_i} \Big|_{x_1, x_2, x_3, \dots, x_N}$$

Remarque

Lorsque la linéarité de f devient trop significative, il faut inclure les termes d'ordre supérieures

Par exemple lorsque la loi de chaque X_i est symétrique autour de sa moyenne
Les termes non linéaire les plus importants sont:

$$\sum_{i=1}^N \sum_{j=1}^N \left[\frac{1}{2} \left\{ \frac{\partial^2 g}{\partial x_i \partial x_j} \right\}^2 + \frac{\partial g}{\partial x_j} \frac{\partial^3 g}{\partial x_i \partial x_j^2} \right] u^2(x_i) u^2(x_j)$$

Determination de l'incertitude type-composée (I)

I- Grandeurs d'entrée corrélés

$$Y = f(X_1, X_2, \dots, X_N)$$

Si les X_i sont significativement corrélés

L'incertitude de y estimation du mesurande Y donc le résultat de mesurage est

$$u_c^2 = \sum_{i=1}^N \left[\frac{\partial f}{\partial x_i} \right]^2 u^2(x_i) + 2 \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j)$$

$u^2(x_i)$ est l'incertitude – type

$u(x_i, x_j) = u(x_j, x_i)$ est la covariance estimée associée à x_i et x_j

Determination de l'incertitude type-composée (II)

I- Grandeurs d'entrée corrélés

Le degré de corrélation est caractérisé par le coefficient de corrélation

$$\rho(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i) \cdot u(x_j)}$$

$$\rho(x_i, x_j) = \rho(x_j, x_i) \text{ et } -1 \leq \rho(x_i, x_j) \leq +1$$

Concepts et termes statistiques fondamentaux

Une expérience physique donne un nombre fini de mesures

L'ensemble résultat \mathcal{E} = Echantillon $\{x_i\}$

Si au cours de n répétitions de \mathcal{E} , le résultat x est apparu n_x la fréquence f_x

$$f_x = \frac{n_x}{n}$$

$$\lim_{n \rightarrow \infty} f_x = P(x)$$

$P(x)$ est la probabilité de trouver le résultat x

Déscription d'une variable aléatoire (I)

Variable aléatoire discrète



$$P_i = P(x = x_i)$$

avec $P_i \in [0,1]$ et $\sum_{i=1}^n P_i = 1$

La fonction de répartition F est définie par

$$F(x_i) = \sum_{j=1}^i P_j = P_1 + P_2 + \dots + P_i = P(x \leq x_i)$$

Quelques propriétés

- $F(x_i)$ n'est jamais décroissant
- $P(a < x < b) = F(b) - F(a)$

Déscription d'une variable aléatoire (II)

Variable aléatoire continue \longrightarrow Densité de probabilité $f_X(u)$

$$P(X \in A) = \int_A f_X(u) du$$

La fonction de répartition F est définie par

$$F_X(u) = P(X \leq u)$$

Quelques propriétés

- $F_X(x)$ n'est jamais décroissant
- $F_X(x)$ varie de 0 à 1 quand x varie de $-\infty$ à $+\infty$ ($F_X(-\infty) = 0$ et $F_X(+\infty) = 1$)
- $P(a < X < b) = F_X(b) - F_X(a)$

Notions associés à la variable aléatoire (III)

Espérance mathématiques d'une variable aléatoire X

v.a. discrète $\longrightarrow E(X) = \mu = \sum_{i=1}^n x_i P(x_i)$

v.a. continue $\longrightarrow E(X) = \mu = \int_D x f_X(x) dx$

Variance

v.a. discrète $\longrightarrow V(X) = E((X - \mu)^2) = \sum_{i=1}^n (x_i - \mu)^2 P(x_i)$

v.a. continue $\longrightarrow V(X) = E((X - \mu)^2) = \int_D (x - \mu)^2 f_X(x) dx$

Ecart type $\longrightarrow \sigma = \sqrt{V(X)}$

Ecart type relative $\longrightarrow \sigma / \mu$

Variance relative $\longrightarrow \sigma^2 / \mu^2$

Notions associés à la variable aléatoires (IV)

Ecart type (d'une variable aléatoire ou d'une loi de probabilité)

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

Covariance de deux variables aléatoires

$$\text{cov}(y, z) = \text{cov}(z, y) = E \{ [y - E(y)][z - E(z)] \}$$

$$\text{cov}(y, z) = \iint (y - \mu_Y)(z - \mu_Z)P(y, z)dydz$$

P(y,z) est la densité de probabilité jointe des deux variables

La covariance peut être estimée par s(y_i, z_i) obtenue de n paires indépendantes D'observations simultanées y_i, z_i de y et z.

$$s(y_i, z_i) = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(z_i - \bar{z}) \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$$

Notions associées à la variable aléatoires (V)

Variable aléatoire centrée réduite

X v.a. d'espérance $E(X) = \mu$ et de variance $V(X) = \sigma^2$

→ $X - \mu$ est la v.a centrée d' espérance nulle

X/σ est la v.a réduite de variance unité

$(X-\mu) / \sigma$ est la v.a centrée réduite d'espérance 0 et de variance 1

Covariance X_1 et $X_2 \rightarrow E(X_{1/2}) = \mu_{1/2}$ et $V(X_{1/2}) = \sigma_{1/2}$

→

$$\begin{aligned} E(Y) &= \mu_1 \pm \mu_2 \\ V(Y) &= \sigma_1^2 + \sigma_2^2 \pm 2\sigma_{12} \end{aligned}$$

$$\sigma_{12} = \text{cov}(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)]$$

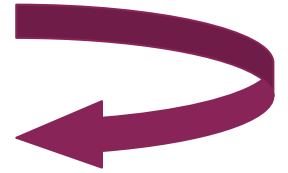
Notions associés à la variable aléatoires (VI)

Généralisation au cas de k variables aléatoires X_1, X_2, \dots, X_k :



$$\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \text{ alors } \underline{E}(\underline{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} \text{ et } \underline{V}(\underline{X}) = \begin{pmatrix} \sigma_1^2 & \sigma_{21} & \dots & \sigma_{k1} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1k} & \sigma_{2k} & \dots & \sigma_k^2 \end{pmatrix}$$

Matrice de Variance-covariance



$$\underline{V}(\underline{X}) = \underline{E} \left\{ [\underline{X} - \underline{E}(\underline{X})] [\underline{X} - \underline{E}(\underline{X})]^T \right\}$$

Quelques lois de distribution de variables aléatoires

I- Loi normale (loi de Gauss)

Une v.a X suit une loi de Gauss ou loi normale notée $\mathcal{N}(\mu, \sigma^2)$ si sa densité de probabilité est

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right\}$$

La fonction de répartition associée s'écrit

$$F(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^x \exp\left\{-\frac{1}{2} \frac{(t - \mu)^2}{\sigma^2}\right\} dt$$

Loi fondamentale de la statistique → théorème de la limite centrale

Remarque


On appelle **quantile (fractile)** d'ordre q de la variable X , où $q \in \{0,1\}$, la valeur x_q telle que $P(X < x_q) = q$ ou de même $F_X(x_q) = q$.


Lois fondamentales de l'échantillonnage

Observations répétées  Echantillon aléatoire de taille n

 Suite de n v.a indépendante régies par la même loi mère

Moyenne de l'échantillon ou moyenne empirique  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

Variance empirique  $\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

Variance de l'échantillon  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Lois fondamentales de l'échantillonnage

Loi du Khi-deux:

Soit Z_1, Z_2, \dots, Z_v , une suite de v.a.i, de loi $\mathcal{N}(0;1)$. Alors la variable $\sum_{i=1}^v Z_i^2$ suit une loi appelée loi du Khi-deux à v degrés de libertés, noté $X^2(v)$.

Densité de probabilité associée est

$$f_v(x) = \frac{1}{2^{\frac{v}{2}} \Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} \text{ pour } x > 0 \text{ (sinon } 0)$$

La fonction Gamma est définie par

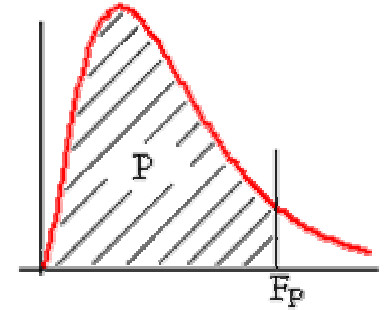
$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

Propriétés élémentaires de la loi du khi-deux

- l'espérance de la loi $\chi^2(v)$ est égale à v
- la variance de la loi $\chi^2(v)$ est égale à $2v$
- $\chi^2(2) = \text{exponentielle}(1/2)$

Fractiles de la loi du χ^2 (ν)

Cette table donne les fractiles F_P de la loi de khi-deux à ν degrés de liberté : $P = \text{Probabilité} (\chi^2 < F_P)$



P
ν

	0.01	0.02	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.85	0.9
1	0.000	0.001	0.004	0.016	0.036	0.064	0.102	0.148	0.275	0.455	0.708	1.074	1.323	1.642	2.072	2.706
2	0.020	0.040	0.103	0.211	0.325	0.446	0.575	0.713	1.022	1.386	1.833	2.408	2.773	3.219	3.794	4.605
3	0.115	0.185	0.352	0.584	0.798	1.005	1.213	1.424	1.869	2.366	2.946	3.665	4.108	4.642	5.317	6.251
4	0.297	0.429	0.711	1.064	1.366	1.649	1.923	2.195	2.753	3.357	4.045	4.878	5.385	5.989	6.745	7.779
5	0.554	0.752	1.145	1.610	1.994	2.343	2.675	3.000	3.656	4.351	5.132	6.064	6.626	7.289	8.115	9.236
6	0.872	1.134	1.635	2.204	2.661	3.070	3.455	3.828	4.570	5.348	6.211	7.231	7.841	8.558	9.446	10.645
7	1.239	1.564	2.167	2.833	3.358	3.822	4.255	4.671	5.493	6.346	7.283	8.383	9.037	9.803	10.748	12.017
8	1.647	2.032	2.733	3.490	4.078	4.594	5.071	5.527	6.423	7.344	8.351	9.524	10.219	11.030	12.027	13.362
9	2.088	2.532	3.325	4.168	4.817	5.380	5.899	6.393	7.357	8.343	9.414	10.656	11.389	12.242	13.288	14.684
10	2.558	3.059	3.940	4.865	5.570	6.179	6.737	7.267	8.295	9.342	10.473	11.781	12.549	13.442	14.534	15.987
11	3.053	3.609	4.575	5.578	6.336	6.989	7.584	8.148	9.237	10.341	11.530	12.899	13.701	14.631	15.767	17.275
12	3.571	4.178	5.226	6.304	7.114	7.807	8.438	9.034	10.182	11.340	12.584	14.011	14.845	15.812	16.989	18.549
13	4.107	4.765	5.892	7.041	7.901	8.634	9.299	9.926	11.129	12.340	13.636	15.119	15.984	16.985	18.202	19.812
14	4.660	5.368	6.571	7.790	8.696	9.467	10.165	10.821	12.078	13.339	14.685	16.222	17.117	18.151	19.406	21.064
15	5.229	5.985	7.261	8.547	9.499	10.307	11.037	11.721	13.030	14.339	15.733	17.322	18.245	19.311	20.603	22.307

Lois fondamentales de l'échantillonnage

Loi de Student

Soit Z et Q deux v.a indépendantes telles que $Z \rightarrow \mathcal{N}(0;1)$, $Q \rightarrow X^2(\nu)$. Alors la v.a

$$T = \frac{Z}{\sqrt{\frac{Q}{\nu}}}$$

Suit une loi appelée loi de Student à ν degrés de liberté, notée $t(\nu)$

Densité de probabilité associée est

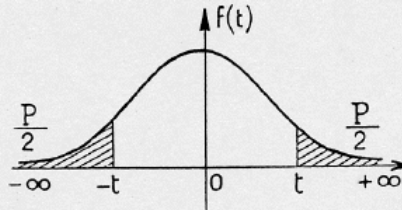
$$f_{\nu}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad x \in \mathcal{R}$$

Propriétés élémentaires de la loi du khi-deux

- l'espérance $E(T) = 0$ si $\nu \geq 2$
- la variance $V(T) = \nu/(\nu-2)$ si $\nu \geq 3$

Table de distribution de T (Loi de Student)

Valeurs de T ayant la probabilité P d'être dépassées en valeur absolue.



$\frac{P}{v}$	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,785	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Lois fondamentales de l'échantillonnage

Théorème de la limite centrale 

Soit $\{X_n\}$ une suite de v.a indépendantes de même loi admettant une moyenne μ et une variance σ^2 . Alors la suite $\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}$ converge en loi vers la loi $\mathcal{N}(0,1)$, s'écrit:

$$\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0;1)$$

Remarque :

$n > 30$ suffit pour obtenir une approximation de probabilité à 0.02 près

Théorie de l'estimation

X une v.a à un certain phénomène aléatoire observable de façon répétée



**Estimer certaines caractéristique d'intérêt de sa loi
(moyenne, variance, fonction de répartition, densité de probabilité,..)
sur la base d'une série d'observation (x_1, x_2, \dots, x_n)**



Estimation ponctuelle



**Attribuer au mieux une valeur
unique à la caractéristique**



Estimation par intervavalle



**Attribuer un encadrement
plausible à la caractéristique**

Définition informelle d'un estimateur et d'une estimation:

Estimer une caractéristique c de la variable aléatoire X sur la base de la réalisation (x_1, x_2, \dots, x_n) de n -échantillon (X_1, X_2, \dots, X_n) . On appelle **estimateur** toute statistique donc toute fonction de X_1, X_2, \dots, X_n dont la réalisation après expérience est envisagée comme estimation de c .

Un **estimateur** se définit dans l'intention de fournir une estimation.

Un **estimateur** est une variable aléatoire

Une **estimation** est une valeur numérique prise par l'estimateur suite à la réalisation de n -échantillons.

Si un estimateur est déterminé par une fonction $h(X_1, X_2, \dots, X_n)$, l'estimation correspondante est $h(x_1, x_2, \dots, x_n)$.

Exemple

moyenne $E(X)$ de la loi X , un estimateur est la moyenne empirique \bar{X} qui produit une estimation \bar{x} moyenne descriptive de la série de valeur observée.

Cadre de l'estimation paramétrique:

En estimation paramétrique la loi de X appartient à une famille de loi décrite

 Forme fonctionnelle connue par fonction de répartition, densité de probabilité..

dépendant de un ou plusieurs paramètres inconnus réels, on note θ ce paramètre qui peut être un vecteur paramètre. L'ensemble des valeurs possibles θ appelés espace paramétrique sera noté Θ inclu dans \mathbb{R}^k , k dimension du paramètre.

Exemple:

La famille des lois de Gauss est décrite par la famille des densités de probabilités de la forme

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right\}$$

un paramètre (μ, σ^2)

L'espace paramétrique est $\{(\mu, \sigma^2) \mid \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}, \sigma^2 \geq 0\}$.


Notation pour les estimateurs et les estimations:

Pour un paramètre désigné par une certaine lettre θ on note

souvent un estimateur par $\hat{\theta}$

On rajoute un indice pour justifier la méthode d'estimation

Qualités des estimateurs

? Choisir l'estimateur  Critère de qualité pertinents

On note T_n l'estimateur de θ à étudier, comme θ est inconnue, il faut que le Comportement de T_n soit satisfaisant quelque soit θ c.a.d quelque soit la Loi mère.

Biais d'un estimateur

Soit une v.a. X de loi de densité (ou fonction de probabilité); $f(x;\theta)$ où $\theta \in \Theta \subseteq \mathbb{R}$.
Soit (X_1, X_2, \dots, X_n) . Un n -échantillon issu de cette loi et T_n un estimateur de θ .
On appelle biais de T_n pour θ la valeur:

$$b_{\theta}(T_n) = E_{\theta}(T_n) - \theta$$

Si $b_{\theta}(T_n) = 0 \forall \theta \in \Theta$, on dit que T_n est un estimateur sans biais

Exemple

Considérons la variance empirique \tilde{S}^2 comme estimateur de la variance où


$$\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$E_{\theta}(\tilde{S}^2) = \frac{n-1}{n} \sigma^2(\theta)$$

Le biais est
$$b_{\theta}(\tilde{S}^2) = \frac{n-1}{n} \sigma^2(\theta) - \sigma^2(\theta) = -\frac{1}{n} \sigma^2(\theta)$$

Il s'agit donc d'un estimateur biaisé

On prend souvent comme variance expérimentale


$$S^2(X_k) = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2$$