



# PROBABILITÉ & STATISTIQUES

*INSA de Lyon*  
**4GMD**  
**2ème Semestre**

**Cours 10 Séances**  
**TD 12 Séances**

**David Dureisseix** – Pr  
Bâtiment S. Germain – 3<sup>ème</sup> Étage

*Équipe de Recherche:*  
LaMCoS / Systèmes Mécaniques et Contacts (SMC)

**Groupes TD 1**

**Jarir Mahfoud** – MdC – Responsable UV  
Bâtiment S. Germain – 3<sup>ème</sup> Étage

*Équipe de Recherche:*  
LaMCoS / Dynamique et Contrôle des Structures (DCS)

**Groupe TD 2,3,4**



# PROBABILITÉ & STATISTIQUES

## Objectifs du cours:

- *Avoir une démarche plus complète qu'une démarche déterministe,*
- *Donner un sens à une série de mesures ou d'observations,*
- *Pouvoir donner un avis sur la qualité d'une production ou d'un processus,*
- *Introduction & initiation à la Fiabilité des systèmes mécaniques,*
- *Introduction & initiation à l'approche simulation numérique (Monte Carlo)*

Example1

Example2



# PROBABILITÉ & STATISTIQUES

## Plan du cours

**I Statistiques descriptive (JM)**

**II Probabilité (DD)**

Expérimentaux  
Espace de probabilité  
Fiabilité

**III Variables Aléatoires (DD)**

**VI Lois de Probabilité**

Fonctions caractéristiques (DD)  
Lois de type discret (DD)  
Lois de type continue (JM)  
Lois fondamentales en Statistiques (JM)

**V Estimation (JM)**

**VI Test d'hypothèse (JM)**

**VII Régression & corrélation linéaires**

**VIII Introduction aux méthodes de simulation numériques**



# PROBABILITÉ & STATISTIQUES

## TD

1. *Statistiques descriptives*
2. *Fréquences et Probabilités conditionnelles,*
3. *Fiabilité des Systèmes*
4. *Formule de Bayes*
5. *Application à la Loi de Poisson*
6. *Applications à la Loi de Normal*
7. *Etudes et Comparaison de Lois*
8. *Application de la loi Log Normal*
9. *Estimation de  $m$  et  $\sigma$  dans PN -  $L(X)=N(m, \sigma)$*
10. *Introduction Aux Tests Graphiques –  $H_0 : L(X) = N ( m ; \sigma )$*
11. *Tests Graphiques Modèle de Weibull –  $L(T) = W(\beta ; \eta ; \delta)$*
12. *Régression – Corrélation linéaire*
13. *Test du  $\chi^2$  – Test de KOLMOGOROV - SMIRNOV - (LILLIEFORS)*



# PROBABILITÉ & STATISTIQUES

## Évaluation

- o **1 Contrôle Intermédiaire (1h) – 31 Mars, 16h30-17h30**
  - VA,
  - Prob & Proba conditionnelle,
  - Modèles de type discret,
  
- o **1 Devoir Terminal (3h) 16 Mai 9h-12h**
  - Ensemble du programme Cours et TD

***Tous les documents sont autorisés***



# PROBABILITÉ & STATISTIQUES

**Probabilité:** L'ensemble des développements conduisant à l'élaboration des modèles liés au raisonnement prenant en compte le caractère **aléatoire** des phénomènes.

***Modélisation***

***Statistiques:*** *l'ensemble de méthodes et de données dont le comportement ne peut être décrit avec certitude.*

***Traitement de données***



# PROBABILITÉ & STATISTIQUES

- On souhaite caractériser un processus, un phénomène (une Population de taille  $N$ ). N'ayant pas la possibilité de traiter toute la  $P_N$ , on le fait sur une petite partie représentative (un échantillon de taille  $n$ ). La réalisation de l'  $\varepsilon_n$  est l'échantillonnage,
- On s'intéresse exclusivement à une série de mesure (un  $\varepsilon_n$  ou une  $P_N$ ), c'est la **Statistique descriptive**,
- Depuis les informations fournies par l'  $\varepsilon_n$ , on souhaite en déduire des renseignements sur la  $P_N$ , c'est l'**estimation**.



# STATISTIQUE DESCRIPTIVE

*Faire parler les mesures*

*Est-ce que l'  $\epsilon_n$  est représentatif?*

*Est-ce que les mesures sont justes?*

*La taille des chercheurs dans un des laboratoire de GMD en cm*

177	162	171	164	167
167	169	162	168	173
168	177	170	177	176
162	175	163	173	182
166	178	180	192	178

*Calcul directe (petit  $\epsilon_n$ ),  
Calcul par classe (grand  $\epsilon_n$ )*

moyenne	H	G	ecart type	g1	g2
171,88	171,58	171,73	7,25	0,67	0,61
CV	0,04				

Médiane 171





# STATISTIQUE DESCRIPTIVE

## Caractéristiques de la position centrale

**Moyenne arithmétique**  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$        $\bar{x} = \frac{1}{n} \sum_{i=1}^k t_i n_i = \sum_{i=1}^k t_i f_i$   
 $t_i$  : milieu de la classe  $i$

**Moyenne quadratique**  $q = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$  ;  $\sqrt{\frac{1}{n} \sum_{i=1}^k n_i t_i^2}$

**Moyenne géométrique**  $g = \left\{ \prod_{i=1}^n x_i \right\}^{1/n}$  ;  $\left\{ \prod_{i=1}^k (t_i)^{n_i} \right\}^{1/n}$

**Moyenne harmonique**  $\frac{1}{h} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}$  ;  $\frac{1}{n} \sum_{i=1}^k \frac{n_i}{t_i}$

**Médiane**

**Mode**

$n_i$  = fréquence absolue nb de fois où l'événement  $i$  a été réalisé,

$N$  = nb total de mesures ;  $f$  = fréquence relative =  $\frac{n_i}{N}$



# STATISTIQUE DESCRIPTIVE

## Caractéristiques de la dispersion

**Étendue**

$$W = x_{\max} - x_{\min}$$

**Variance "V"**

$$V = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad ; \quad \frac{1}{n} \sum_{i=1}^k n_i (t_i - \bar{x})^2 = s^2$$

**Écart type "s"**

$$s = \sqrt{V}$$

**Coefficient de variation**

$$C_v = \frac{s}{\bar{x}} = \frac{\sigma}{m}$$



# STATISTIQUE DESCRIPTIVE

## Moments non centrés d'une Variable Aléatoire X

Moment non centré d'ordre k pour =  $m_k = \int_D x^k dF_X$        $m_k = \int_D x^k f_X(x)$

$$\frac{1}{n} \sum_{i=1}^n x_i^k = m_k = \sum_{i=1}^n t_i^k \Pr_i$$

---

$$m_0 = 1 \quad m_1 = E\{X\}$$

---

$$Y = aX + b$$

$$E\{Y\} = a E\{X\} + b$$



# STATISTIQUE DESCRIPTIVE

## Moments centrés d'une Variable Aléatoire X

**Moment centré d'ordre k =** 
$$\mu_k = \int_D (x - E\{X\})^k dF_X = E(X - E(X))^k$$

$$\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$$

---

$$\mu_1 = 0$$

$$\mu_2 = V(X)$$

$$\mu_r = \sum_{j=0}^r (-1)^j C_r^j (m_1)^j m_{r-j} \quad ; \quad C_r^j = \frac{r!}{j!(r-j)!}$$

---

Relation de Koenig 
$$\mu_2 = V(X) = E(X^2) - m_1^2$$

---

$$Y = aX + b$$

$$V\{Y\} = a^2 V\{X\}$$



# STATISTIQUE DESCRIPTIVE

## Coefficients de forme d'une distribution

### *aplatissement*

$$g_2 = \frac{\mu_4}{\mu_2^2} - 3$$

$$g_2 < 0$$

*distribution plus aplatie que la loi normale*

$$g_2 = 0$$

*aplatissement d'un phénomène normal*

$$g_2 > 0$$

*moins aplatie que la loi normale*

### *asymétrie*

$$g_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

$$g_1 < 0$$

*distribution étalée à gauche*

$$g_1 = 0$$

*distribution symétrique*

$$g_1 > 0$$

*distribution étalée à droite*



# STATISTIQUE DESCRIPTIVE

## Traitement par classes (il faut entre 6 et 15 classes)

**Etape 1**  
Ordonner l'  $\epsilon_n$

Ech_Ordon
162
162
162
162
163
164
166
167
167
168
168
169
170
171
173
173
175
176
177
177
177
178
178
180
182
192

**Etape 2**  
Calculer le nb des classes

$$K \cong \sqrt{n}$$

$$K > 100 \Rightarrow K = 5 \log(n)$$

**Etape 3**  
Calculer l'amplitude des classes

$$Ampl \cong \frac{X_{\max} - X_{\min}}{K}$$

**Etape 4**  
Compter les valeurs par classe

Classe	1	2	3	4	5	6
Amplitude	$160 \leq < 166$	$166 \leq < 172$	$172 \leq < 178$	$178 \leq < 184$	$184 \leq < 190$	$190 \leq < 196$
$n_i$	5	8	7	4	0	1

**La taille des chercheurs dans un des laboratoire de GMD en cm**

177	162	171	164	167
167	169	162	168	173
168	177	170	177	176
162	175	163	173	182
166	178	180	192	178

**Est-ce que j'ai raison?**  
**Est-ce que j'ai fait les bons choix?**



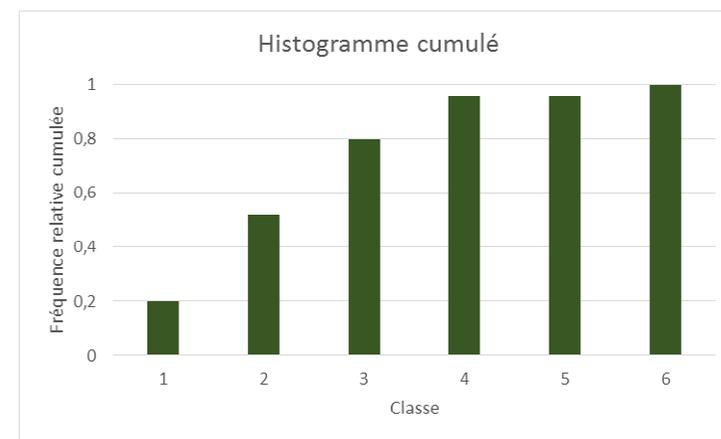
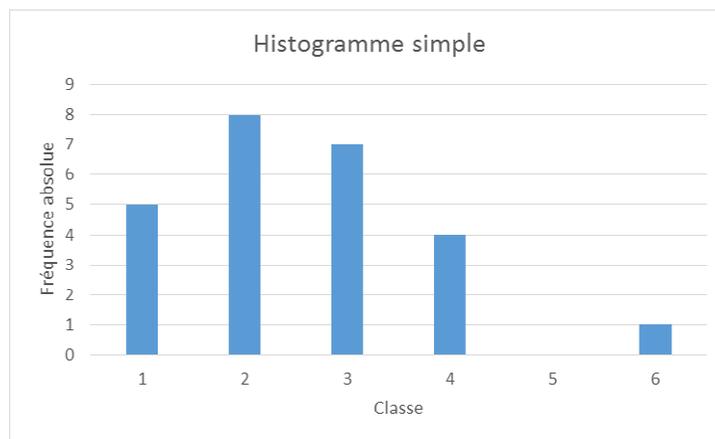
# STATISTIQUE DESCRIPTIVE

**Traitement par classes**  
(il faut entre 6 et 15 classes)

**La taille des chercheurs dans un des laboratoire de GMD en cm**

177	162	171	164	167
167	169	162	168	173
168	177	170	177	176
162	175	163	173	182
166	178	180	192	178

## HISTOGRAMMES





# STATISTIQUE DESCRIPTIVE

**Traitement par classes**  
(il faut entre 6 et 15 classes)

Classe	Amplitude	milieu	ni	fi	niXti	niXti2	niXti3	niXti4
1	160-166	163	5	0,2	815	132845	21653735	3529558805
2	166-172	169	8	0,32	1352	228488	38614472	6525845768
3	172-178	175	7	0,28	1225	214375	37515625	6565234375
4	178-184	181	4	0,16	724	131044	23718964	4293132484
5	184-190	187	0	0	0	0	0	0
6	190-196	193	1	0,04	193	37249	7189057	1387488001

$\Sigma$  25 1 4309 744001 128691853 2,23E+10

moyenne	H	G	ecart type	g1	g2
171,88	171,58	171,73	7,25	0,67	0,61
CV	0,04				

Médiane 171  
Mode

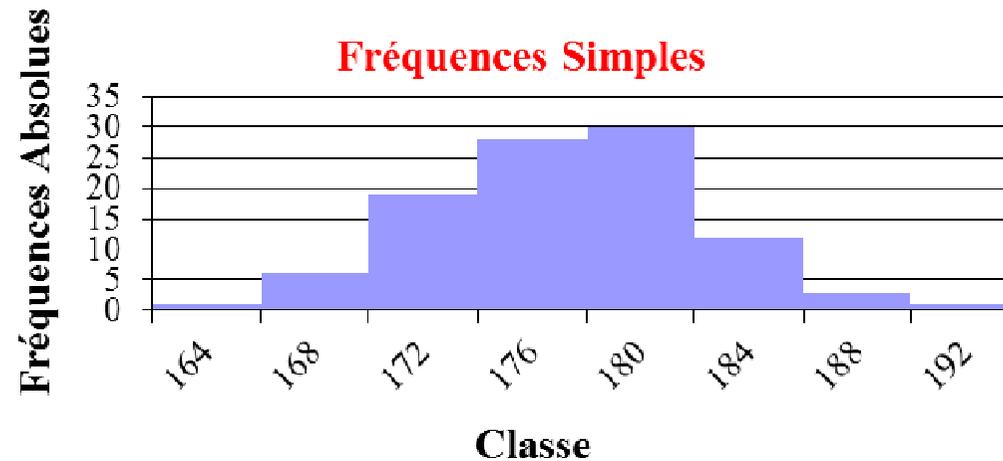
$\Sigma$	25	4309	744001	128691853	22301259433
mi		1,724E+02	2,976E+04	5,148E+06	8,921E+08
$\mu$			7,216E+00	2,839E+02	9,740E+03
g1				7,556E-01	
g2					5,925E-01



# STATISTIQUE DESCRIPTIVE

*La taille des chercheurs dans un des laboratoire de l'INSA en cm*

173	177	175	179	175	172	173	178	186	174
181	177	172	160	173	176	181	180	177	182
179	167	176	169	173	183	174	167	183	168
174	178	171	177	176	179	178	186	182	184
177	180	176	171	177	192	183	171	178	179
176	179	178	174	170	178	173	170	182	174
170	169	169	181	180	170	178	173	173	172
179	175	177	169	168	174	170	175	177	170
186	170	181	176	171	176	166	175	170	180
183	167	179	175	180	174	177	178	174	178



<b>m</b>	175,68	<b>g1</b>	0,07
<b><math>\sigma</math></b>	5,16	<b>g2</b>	0,45



# STATISTIQUE DESCRIPTIVE

*La taille des chercheurs dans un des laboratoire de l'INSA en cm*

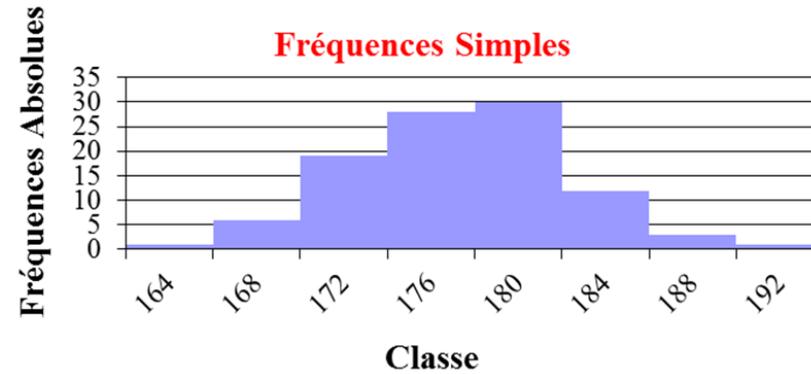
173	177	175	179	175	172	173	178	186	174
181	177	172	160	173	176	181	180	177	182
179	167	176	169	173	183	174	167	183	168
174	178	171	177	176	179	178	186	182	184
177	180	176	171	177	192	183	171	178	179
176	179	178	174	170	178	173	170	182	174
170	169	169	181	180	170	178	173	173	172
179	175	177	169	168	174	170	175	177	170
186	170	181	176	171	176	166	175	170	180
183	167	179	175	180	174	177	178	174	178

Classe	milieu	ni	ni*Ti	ni*Ti2	ni*Ti3	ni*Ti4
<=164,00000	1,620E+02	1	1,620E+02	2,624E+04	4,252E+06	6,887E+08
<=168,00000	1,660E+02	6	9,960E+02	1,653E+05	2,745E+07	4,556E+09
<=172,00000	1,700E+02	19	3,230E+03	5,491E+05	9,335E+07	1,587E+10
<=176,00000	1,740E+02	28	4,872E+03	8,477E+05	1,475E+08	2,567E+10
<=180,00000	1,780E+02	30	5,340E+03	9,505E+05	1,692E+08	3,012E+10
<=184,00000	1,820E+02	12	2,184E+03	3,975E+05	7,234E+07	1,317E+10
<=188,00000	1,860E+02	3	5,580E+02	1,038E+05	1,930E+07	3,591E+09
<=192,00000	1,900E+02	1	1,900E+02	3,610E+04	6,859E+06	1,303E+09

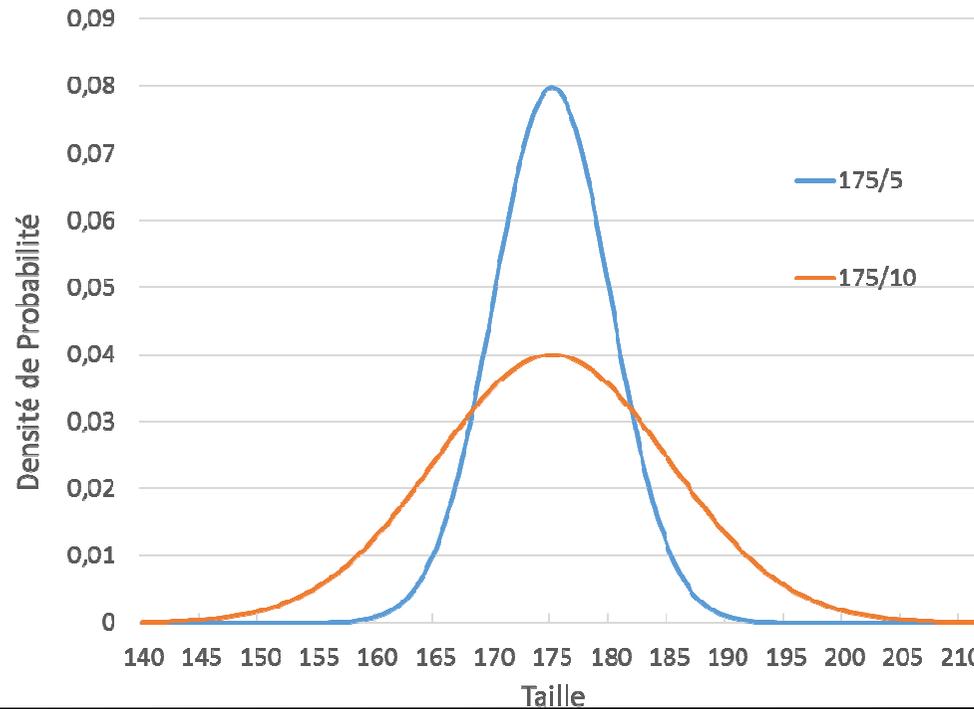
$\Sigma$	100	1,753E+04	3,076E+06	5,402E+08	9,496E+10
mi		1,753E+02	3,076E+04	5,402E+06	9,496E+08
$\mu$			5,093E+00	8,267E+00	2,030E+03
g1				6,258E-02	
g2					1,694E-02



# STATISTIQUE DESCRIPTIVE



*Comparaison de 2 distributions avec une taille moyenne de 175 cm et un écart type de 5 ou 10 cm*





# Exemple d'application

## Cryptanalyse

Dans, pratiquement, toutes les langues, les lettres n'ont pas la même fréquence d'apparition. Donc, dans un texte écrit en français, il aura forcément plus de E, A et S que de K ou W

*Fréquence d'apparition des lettres en langue française*

A	B	C	D	E	F	G	H	I	J	K	L	M
9,42	1,02	2,64	3,39	15,87	0,95	1,04	0,77	8,41	0,89	0	5,34	3,24
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
7,15	5,14	2,86	1,06	6,46	7,9	7,26	6,24	2,15	0	0,3	0,24	0,32

TIXIV KYWXEZ PINIYRI6HMVMGLPIX IWX RI PI 13 JIZVMIV 1805 E HYVIR, YRI ZMPPI  
H'EPPIQEKRI WMXYII E QM-GLIQMR IRXVI EEGLIR IX GSPSKRI. WSR TIVI C IXEMX VIGIZIYV  
HIW TSWXIW. HMVMGL-PIX IWX YR IPIZI FVMPPERX. UYM EGLIZI WIW IXYHIW  
WIGERHEMVIW E 16 ERW.

Limite de la méthode:

- Avoir assez de mots (un échantillon de taille suffisante),
- Texte écrit en français standard.