

Médecine factuelle : une introduction à la statistique

Yohann.Foucher@univ-nantes.fr

Equipe d'Accueil 4275 "Biostatistique, recherche clinique et mesures subjectives en santé", Université de Nantes

Odontologie - Cours #1

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine factuelle

Concept EBM
L'échantillonnage
Les fluctuations
d'échantillonnage

Les variables aléatoires

Définitions
Loi de probabilité
Loi Normale

1. Médecine factuelle

Concept EBM
L'échantillonnage
Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions
Loi de probabilité
Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

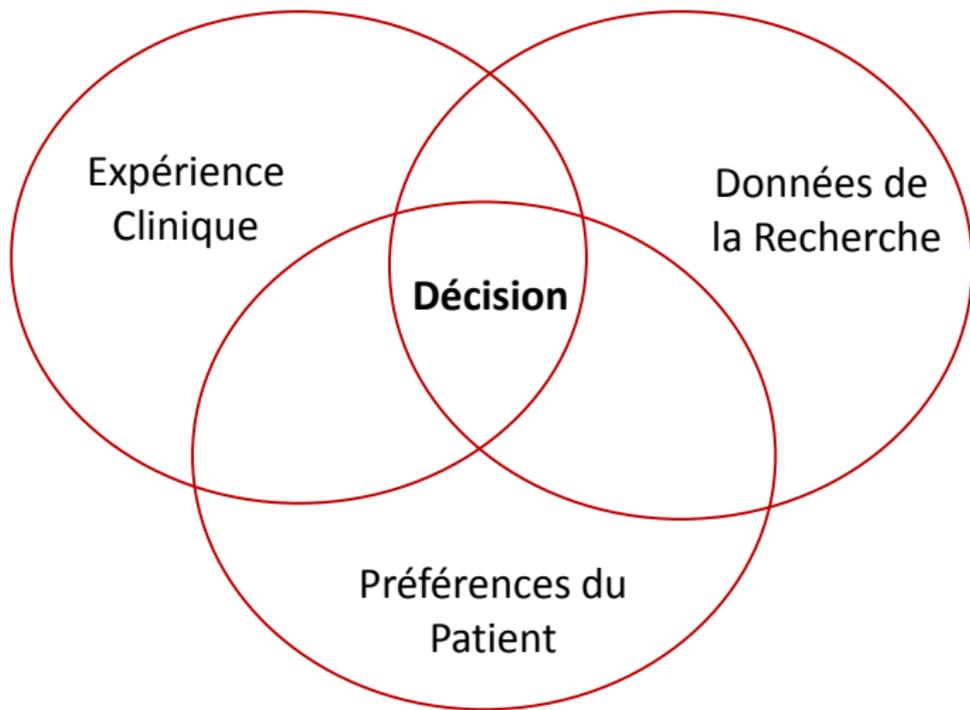
Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Les déterminants des pratiques médicales n'ont pas toujours été basés sur les faits
- Jusqu'aux années 60, force de la routine, des traditions...
- Notion d'expérience professionnelle.
- Depuis, structuration de l'*Evidence-Based-Medicine*.
- Comité Consultatif National d'Ethique :
 - *"Il n'est pas conforme à l'éthique d'administrer un traitement dont on ne sait pas, alors qu'on pourrait le savoir, s'il est le meilleur des traitements disponibles [...]"*



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

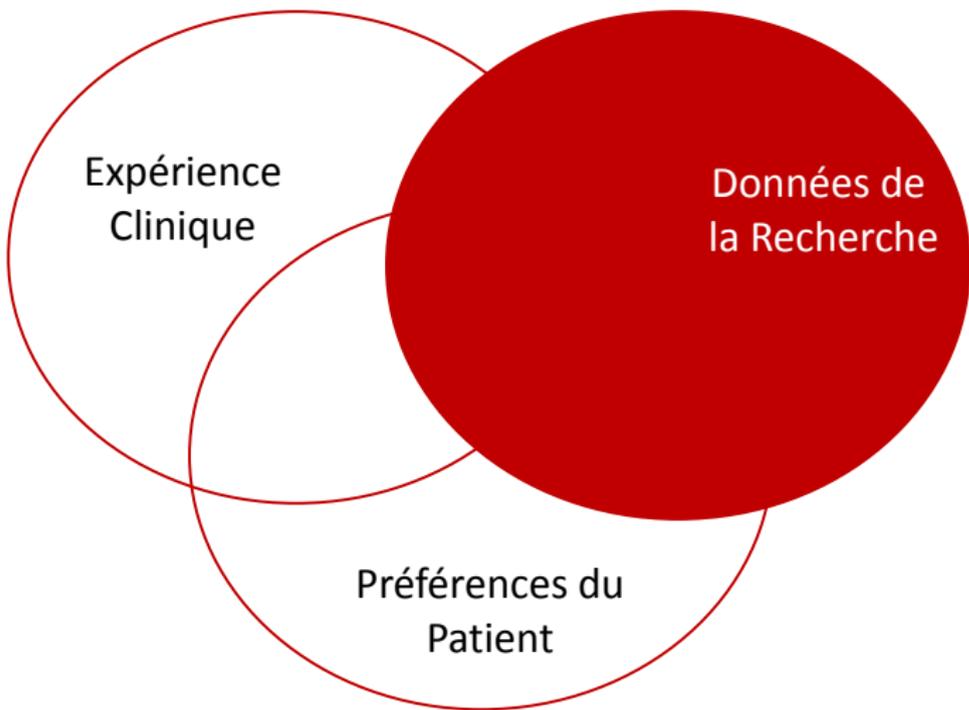
Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Médecine factuelle
- Concept EBM
- L'échantillonnage
- Les fluctuations d'échantillonnage
- Les variables aléatoires
- Définitions
- Loi de probabilité
- Loi Normale



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Le réel est monstrueux. Il est énorme, il est hors norme. *
- Etudier totalement une population reviendrait à vouloir percevoir de manière simultanée et continue, toutes les caractéristiques de tous les individus de cette population. †
- Pour approcher et distinguer les choses, pour s'en faire une idée tangible, il faut s'en tenir à l'appréhension d'un nombre limité de caractéristiques.

*. Edgar Morin

†. Y. Macé. Journal de la Société Française de Statistique, tome 147, 2006.

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- **La population cible :**
 - Population qui nous intéresse, celle qui est concernée par les résultats de l'étude.
- **La population source :**
 - Population pour laquelle on peut observer les patients.
 - Elle est incluse dans la population cible.
- **La population étudiée :**
 - Un faux ami : il s'agit le plus souvent de l'échantillon.
 - Un échantillon observable contrairement à une population.

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- 1 Essais cliniques contrôlés (études expérimentales).
- 2 Méta-analyses.
- 3 Cohortes (études observationnelles).
- 4 Etudes cas-témoins (études observationnelles).
- 5 Etudes transversales (études observationnelles).

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

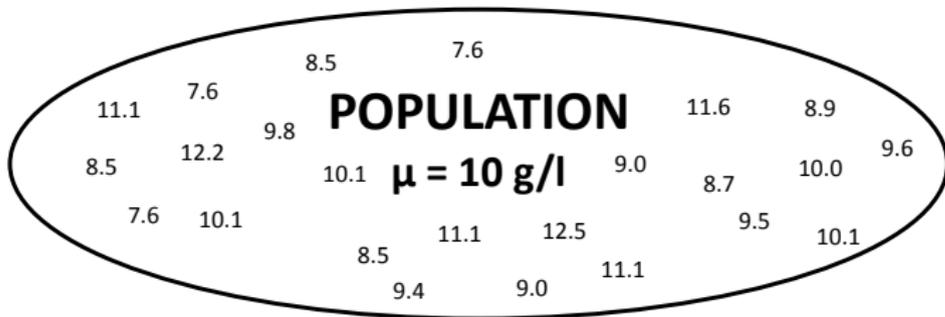
Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

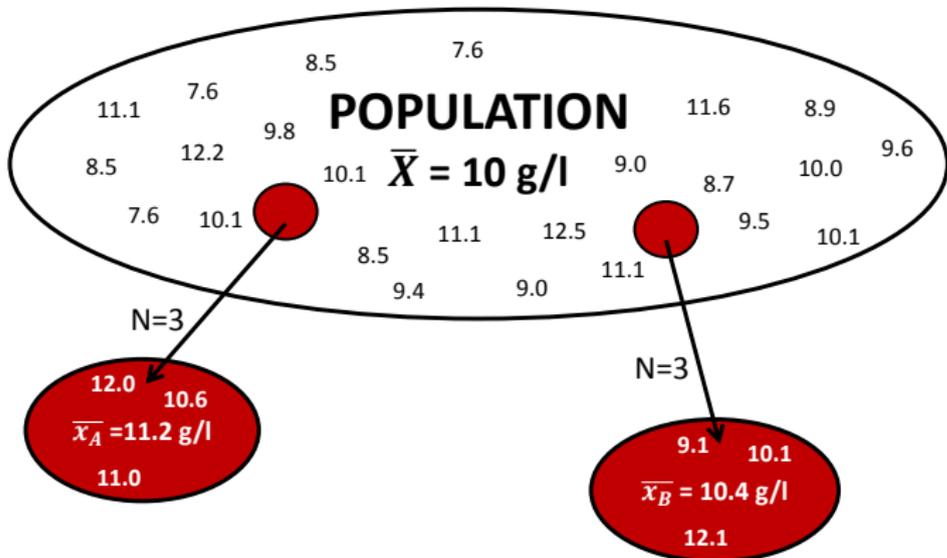
Définitions

Loi de probabilité

Loi Normale

- Soit \mathcal{P} la population cible.
Ex : $\mathcal{P} = \text{L'ensemble des patients anémiques.}$
- Soit μ le taux moyen d'hémoglobine
Ex : $\bar{X} = 10 \text{ g/l.}$
- Il n'est pas possible de mesurer \bar{X} à partir de tous les patients de la population.
- On réalise un échantillon de N patients à partir desquels on observe une moyenne \bar{x} .
Ex : $N = 3.$
- Problème : si plusieurs échantillons sont réalisés, on observera autant de moyennes.

La fluctuation d'échantillonnage



Médecine
 factuelle

Concept EBM

L'échantillonnage

Les fluctuations
 d'échantillonnage

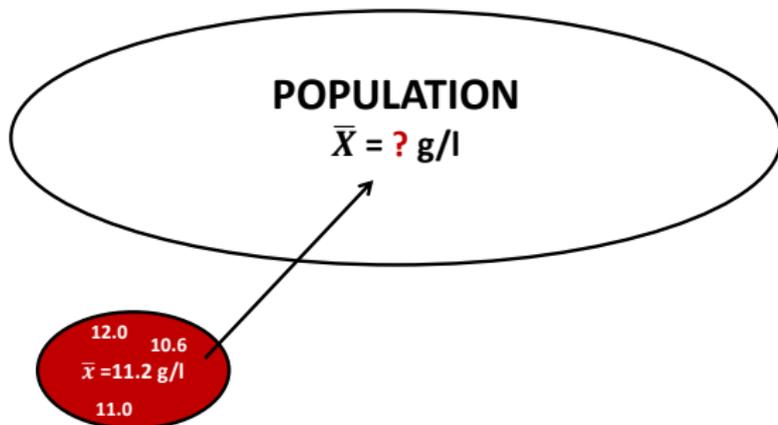
Les variables
 aléatoires

Définitions

Loi de probabilité

Loi Normale

- On espère que l'estimation observée est proche de la vraie valeur.
- On tente de quantifier le niveau d'incertitude..



Ex : J'estime que la taux moyen d'hémoglobine est proche de 11.2 g/l. Avec une confiance de 95%, j'estime que le taux moyen d'hémoglobine des patients anémiques est compris entre 7.5 et 14.9 g/l.

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Les v.a. ont été définies à l'origine pour représenter un gain.
- Ex : Lancer d'une pièce de monnaie. X peut prendre les valeurs :
 - 1 euro si pile
 - -1 euro si face
- De façon plus générale une variable aléatoire est le résultat inconnu d'une expérience aléatoire.
- Une fois l'expérience réalisé, on observe la réalisation de la v.a. X . Cette valeur n'est plus aléatoire mais est certaine. On la note souvent x . x_1 pour la 1ère expérience, x_2 pour la 2nde expérience, etc.

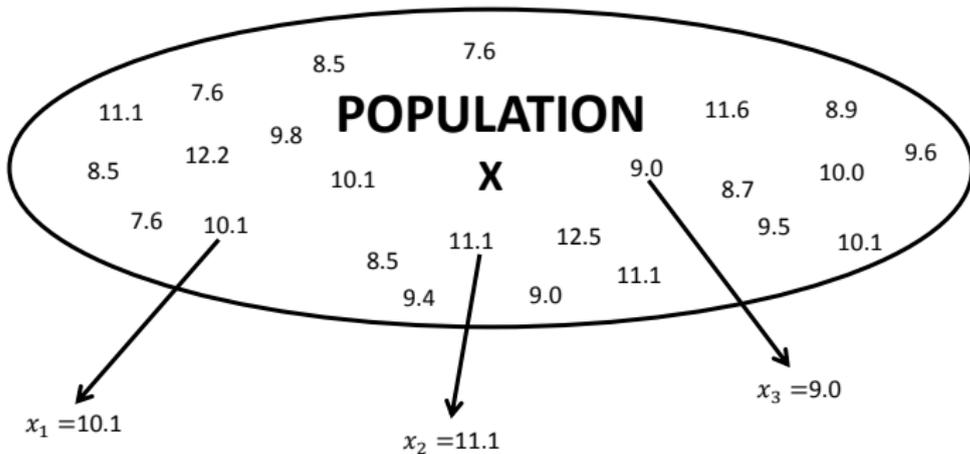
La notion de variable aléatoire

Médecine
factuelle

Concept EBM
L'échantillonnage
Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions
Loi de probabilité
Loi Normale



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- On parle de variable aléatoire discrète :
 - Ex1** : Jeu de dés après un lancer : résultats \rightarrow 1, 2, 3, 4, 5, 6.
 - Ex2** : Réponse thérapeutique : résultats \rightarrow guérison oui/non.
- On parle de variable aléatoire continue :
 - Ex1** : Taux d'hémoglobine : résultats \rightarrow tous les réels positifs...

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Une variable aléatoire est définie par sa loi de probabilité.
- Ex : Jeu de dés après un lancé

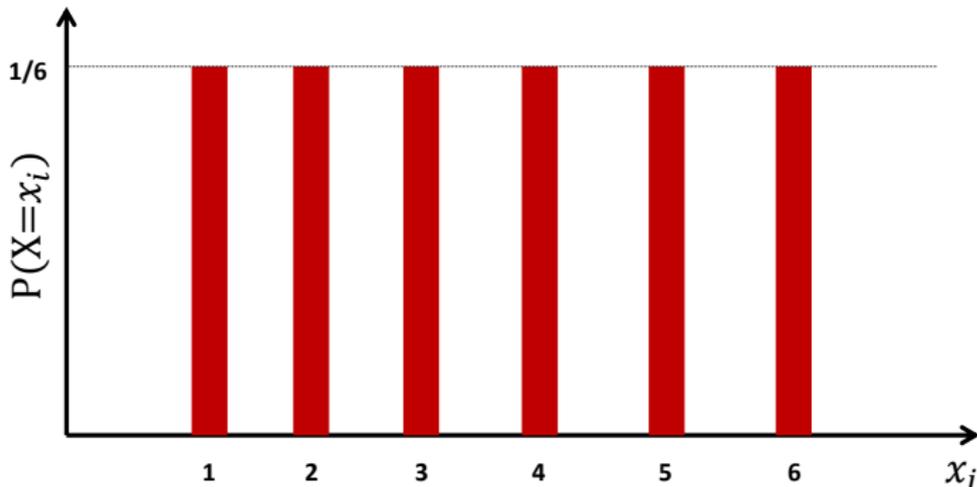
Résultats	1	2	3	4	5	6
Probabilités	1/6	1/6	1/6	1/6	1/6	1/6

Loi de probabilité d'une v.a. discrète

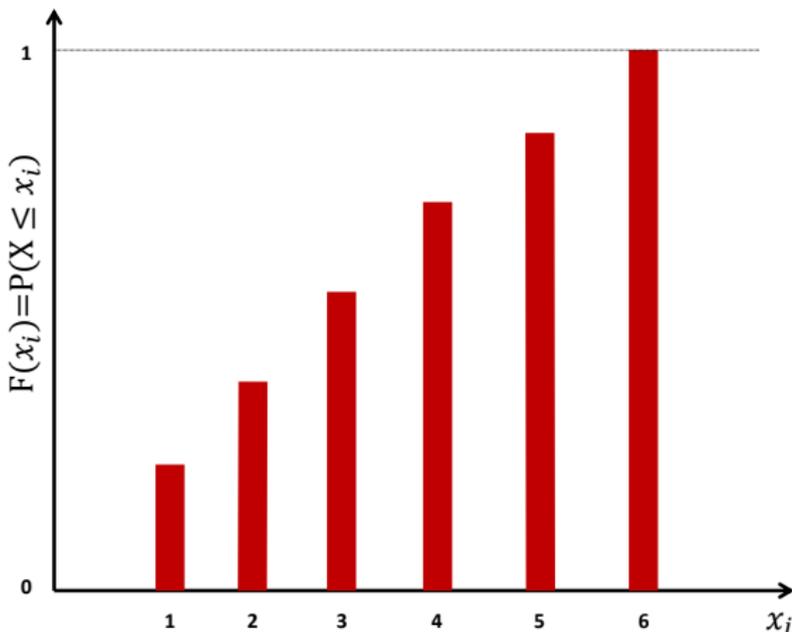
- Contrainte : $\sum_i P(X = x_i) = 1$
- Représentation par une table :

Résultats (x_i)	1	2	3	4	5	6
Probabilités ($P(X = x_i)$)	1/6	1/6	1/6	1/6	1/6	1/6

- Ou représentation par un graphique en bâtons :



- $F(x_i) = P(X \leq x_i) = \sum_{j: x_j \leq x_i} P(X = x_j)$



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

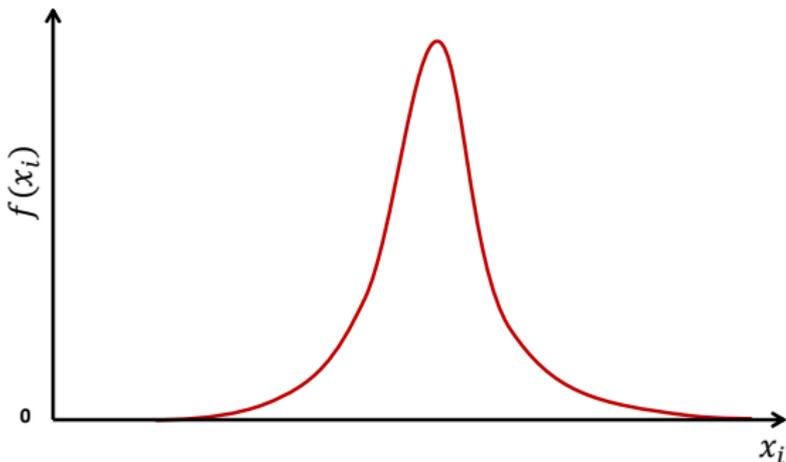
Loi de probabilité

Loi Normale

- Problème : la probabilité $P(X = x_i)$ est nulle.
- On travaille sur la limite :

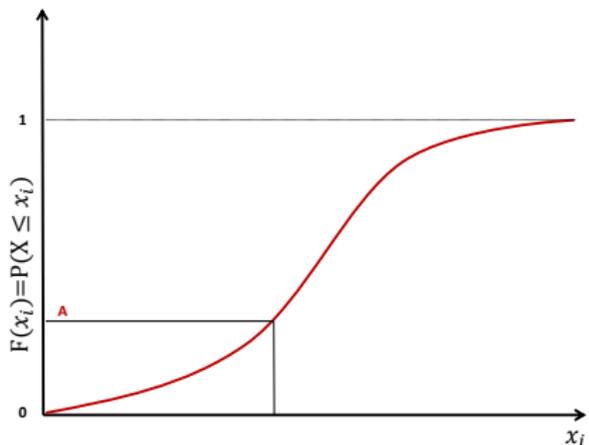
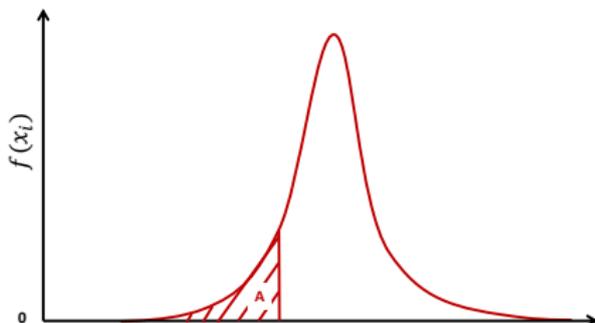
$$f(x_i) = \lim_{\Delta x \rightarrow 0} P(x_i < X < x_i + \Delta x) / \Delta x$$

- C'est la fonction de densité.
- Aucune possibilité de représentation sous forme de tableau.
- Représentation par un graphique :



La fonction de répartition d'une v.a. continues

- $F(x_i) = P(X \leq x_i) = \int_{-\infty}^{x_i} f(x) dx$



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

1. Médecine factuelle

Concept EBM

L'échantillonnage

Les fluctuations d'échantillonnage

2. Les variables aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

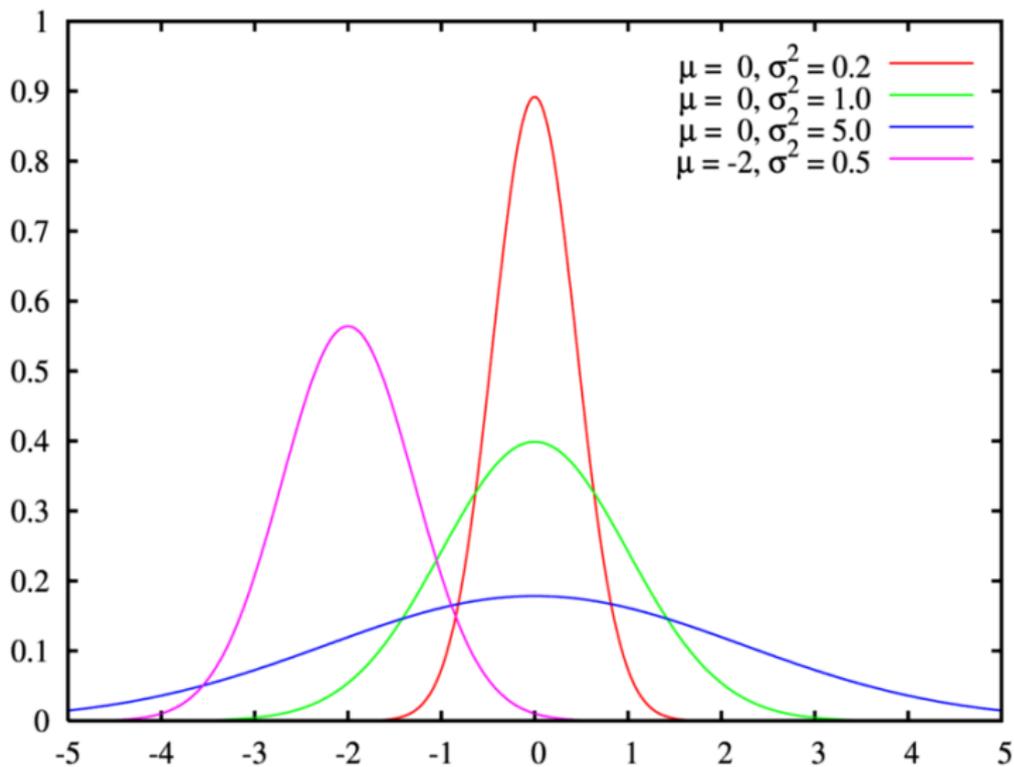
Loi Normale

- Loi normale ou loi de Laplace-Gauss
- Caractérisée par sa moyenne (μ) et son écart-type (σ) :

$$X \sim \mathcal{N}(\mu, \sigma)$$

- La distribution est symétrique et centrée autour de la moyenne.
- L'écart-type représente la dispersion autour de la moyenne.

Loi normale



Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

Médecine
factuelle

Concept EBM
L'échantillonnage
Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions
Loi de probabilité

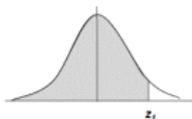
Loi Normale

- Si $X \sim \mathcal{N}(\mu, \sigma)$ alors :

$$U = (X - \mu) / \sigma \sim \mathcal{N}(0, 1)$$

- $\mathcal{N}(0, 1)$ est la loi normale centrée et réduite.
- Intérêt de la transformation : on peut calculer $P(X < x)$.

Loi normale



$$p(z \leq z_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_1} e^{-\frac{1}{2}z^2} dz$$

Médecine
factuelle

Concept EBМ

L'échantillonnage

Les fluctuations
d'échantillonnageLes variables
aléatoires

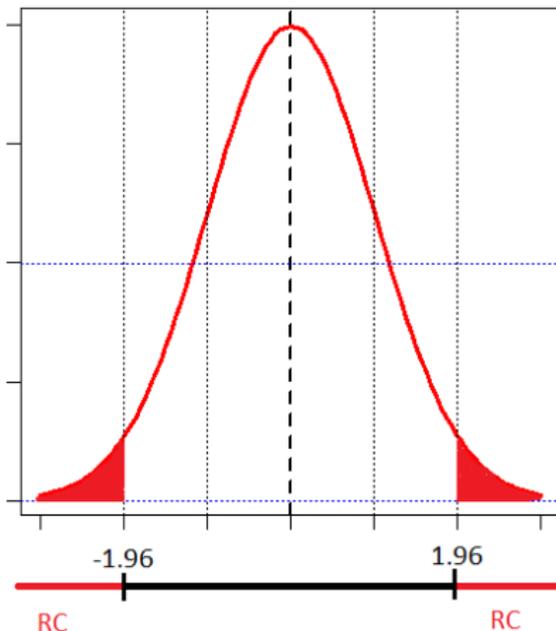
Définitions

Loi de probabilité

Loi Normale

z_1	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

- Valeur importante : $P(-1,96 < T < 1,96) = 0,95$.



- Soit un échantillon de taille N et X_1, X_2, \dots, X_N une suite de variables aléatoires indépendantes et identiquement distribuées.
- Supposons que cette loi possède une espérance μ et un écart-type σ (avec $\sigma \neq 0$).
- Considérons la somme $S_N = X_1 + X_2 + \dots + X_N = \sum_{i=1}^N X_i$. Alors l'espérance de S_N vaut $N\mu$ et son écart-type vaut $\sigma\sqrt{N}$.
- Quand N est assez grand ($N > 30$), la loi normale $\mathcal{N}(N\mu, N\sigma^2)$ est une bonne approximation de la loi de S_N .
- Cette approximation s'écrit :

$$U = \frac{S_N - N\mu}{\sigma\sqrt{N}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}} \sim \mathcal{N}(0, 1)$$

Médecine
factuelle

Concept EBM

L'échantillonnage

Les fluctuations
d'échantillonnage

Les variables
aléatoires

Définitions

Loi de probabilité

Loi Normale

- Conséquence : Une donnée influencée par une multitude de phénomènes aléatoires indépendants qui s'additionnent est approximativement décrite par une loi normale même si les phénomènes qui la composent ne suivent pas des lois normales.