Loi Normale

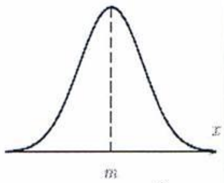
Dr : RAIAH. M
Service de Biostatistique, Faculté de Médecine d'Oran

I- Introduction

C'est la plus importante des lois utilisées en statistique. La loi normale s'applique aux variables quantitatives continues.

En biologie, on constate souvent que la distribution des valeurs d'une variables s'agglutine autour de la moyenne. En suite ces valeurs décroissent systématiquement de part et d'autre de cette moyenne.

C'est le cas par exemple de la distribution de la taille des individus dans une population.



II- Définition (1/2)

La distribution normale, appelée aussi gaussienne est une distribution continue qui dépend de 2 paramètres μ et σ.

Une variable aléatoire X suit une loi normale si sa densité est :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Où :

$x \in R$, $\mu \in R$ et $\sigma > 0$

II- Définition (2/2)

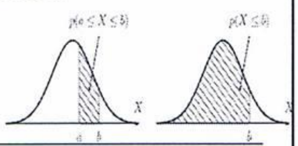
La variable aléatoire X suit une loi normale de moyenne μ et d'écart type σ (on note : $X \sim N(\mu ; \sigma)$) signifie que :

- L'ensemble des valeurs possibles de X est l'ensemble de tous les nombres réels : $X \in]-\infty ; +\infty[$
- Quels que soient les deux nombres a et b ($a \leq b$), la probabilité que X soit compris entre a et b est égale à "l'aire sous la courbe" en cloche N(μ ; σ) entre a et b.

$p(a \leq X \leq b)$ = aire sous la courbe entre a et b

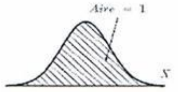
et aussi

$p(X \leq b)$ = aire sous la courbe de $-\infty$ jusqu'à b



Remarques :

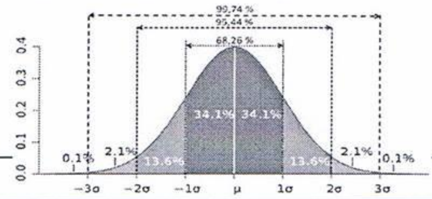
- Quels que soient $\mu \in R$ et $\sigma > 0$, l'aire "totale" sous la courbe vaut 1. (pour x allant de $-\infty$ à $+\infty$)



- Quel que soit $a \in R$ on a : $p(X = a) = 0$ la probabilité d'une "valeur isolée" est nulle
- Quels que soient les deux nombres a et b avec $a \leq b$ on a : $p(a \leq X \leq b) = p(a < X < b)$ (pour une loi normale, $<$ et $\leq$ sont "équivalents", ainsi que $>$ et $\geq$)
- $p(a \leq X \leq b) = \int_a^b f(x)dx$ (en terme d'intégrale)

Propriétés de la loi normale :

- 68,26% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu - \sigma; \mu + \sigma]$
- 95,44% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu -2\sigma; \mu +2\sigma]$
- 99,74% des valeurs d'une variable aléatoire normale sont comprises dans l'intervalle $[\mu -3\sigma; \mu +3\sigma]$



III - Loi normale centrée réduite

On dit que la distribution est centrée si E(X)=0 et réduite si V(X)=1.

La distribution normale centrée réduite N(0 ; 1) est définie par :

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$$

Transformation d'une loi normale quelconque en loi N(0 ; 1)

Soit X une variable aléatoire continue suivant une loi normale de moyenne μ et d'écart-type σ.

Si on applique le changement de variable :

$$t = \frac{X - \mu}{\sigma}$$

la variable t suit une loi normale centrée réduite :

$$g(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}$$

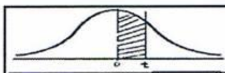
Remarques :

- Un calcul de probabilité sur une loi normale quelconque revient un calcul de probabilité sur une loi normale centrée réduite.
- Si X suit une loi normale N(μ ; σ), alors, pour calculer $p(X \leq a)$, il suffit de calculer :

$$p(t \leq \frac{a-\mu}{\sigma}) = G(\frac{a-\mu}{\sigma})$$

Où :

$G(\frac{a-\mu}{\sigma})$ est estimé avec la table de la loi N(0 ; 1)



Aire limitée par la courbe de la loi normale centrée réduite, l'axe des ordonnées et la droite d'équation $T = G(t) = \int_0^t g(t)dt$

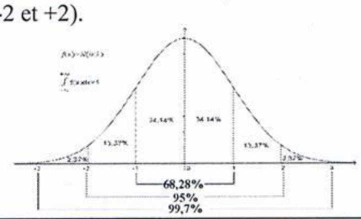
Avec $g(t) = (1/\sqrt{2\pi}) \cdot e^{(-1/2) t^2}$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	[illegible]	[illegible]	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4994	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Exemple : G (0,92) = 0, 3212. La valeur 0, 3212 est l'intersection de 0,9 et 0,02.
0,92 = 0,9 + 0, 02.

Propriétés de la loi normale centrée réduite :

- La loi N(0;1) est centrée autour de la valeur zéro, et elle a pour un écart type la valeur 1.
- 95% des valeurs sont comprises entre -1,96 et +1,96 ≈ (-2 et +2).



IV - Exemple :

On suppose que la glycémie est distribuée normalement dans la population, avec une moyenne de 1 g/l et un écart-type de 0,03 g/l. On mesure la glycémie chez un individu.

Calculer la probabilité pour que sa glycémie soit:

a) inférieure à 1,06 ;

b) supérieure à 0,9985 ;

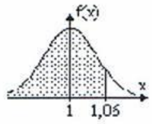
c) comprise entre 0,94 et 1,08.

Correction :

Notons X la glycémie mesurée sur un individu de la population. X suit une loi normale N(1 ; 0,03).

a) P (X < 1,06)

C'est la surface hachurée suivante :



Avant de chercher la probabilité demandée, il faut transformer la variable X en variable centrée et réduite :

$T = \frac{x-\mu}{\sigma} = \frac{1,06-1}{0,03} = 2$

Donc, P (x < 1,06) = P (t < 2)

$P(t<2) = \int_{-\infty}^{2} g(t)dt$

$= \int_{-\infty}^{0} g(t)dt + \int_{0}^{2} g(t)dt$

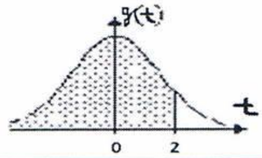
$= 0,5 + G(2)$

La valeur G (2) peut être lue dans la table

G (2) = 0,4772 $\Longrightarrow$ P (t < 2) = 0,5 + 0,4772

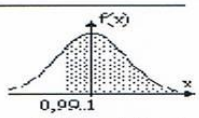
= 0,9772

$\boxed{P(x < 1,06) = 0,9772}$



b) P (X > 0,9985)

C'est la surface hachurée suivante :



Transformation de la variable X en variable centrée et réduite :

$T = \frac{x-\mu}{\sigma} = \frac{0,9985-1}{0,03} = -0,05$

P (X > 0,9985) = P (T > - 0,05)

$= \int_{-0,05}^{+\infty} g(t)dt$

$= \int_{-0,05}^{0} g(t)dt + \int_{0}^{+\infty} g(t)dt$

= G(-0,05) + 0,5

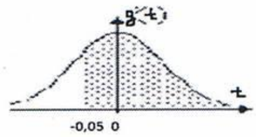
Par symétrie, G(-0,05) = G(0,05)

La valeur G (0,05) peut être lue dans la table :

G (0,05) = 0,0199 $\Longrightarrow$ P (T > - 0,05) = 0,5 + 0,0199

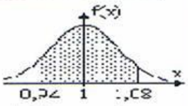
= 0,5119

$\boxed{P(X > 0,9985) = 0,5119}$



■) P (0,94 < X < 1,08)

C'est la surface hachurée suivante :



Transformation de la variable X en variable centrée et réduite :

$$T_1 = \frac{0{,}94-1}{0{,}03} = -2 \; ; \; T_2 = \frac{1{,}08-1}{0{,}03} = 2{,}66$$

P (0,94 < X < 1,08) = P (- 2 < T < 2,66)

$$= \int_{-2}^{2{,}66} g(t)dt$$

$$= \int_{-2}^{0} g(t)dt + \int_{0}^{2{,}66} g(t)dt$$

$$= G(-2) + G(2{,}66)$$

Par symétrie, G(-2) = G(2)

Les valeurs G (2) et G (2,66) peuvent être lues dans la table :

$$\left.\begin{array}{l} G(2) = 0{,}4772 \\ G(2{,}66) = 0{,}4961 \end{array}\right\} \; P(-2 < T < 2{,}66) = 0{,}4772+0{,}4961 = 0{,}9733$$

$$\boxed{P(0{,}94 < X < 1{,}08) = 0{,}9773}$$