GCI 102. Méthodes probabilistes en génie civil

Chapitre 6

Distributions de probabilité continues

par

Pierre F. Lemieux, ing., professeur
Département de génie civil
Université de Sherbrooke
Tél. : (819) 821-8000 (poste 2938)
Télécopieur : (819) 821-7974
Courriel : Pierre.F.Lemieux@USherbrooke.ca

Révision : 25 juin 2003

Table des matières

1. Introduction

[LRS, p. 181] [B, p. 218-219]

En général, pour une variable aléatoire continue X,

$$-\infty < X < \infty$$

En génie, ce n'est pas toujours le cas. Les débits , les vitesses du vent, …, ne prennent pas de valeurs négatives.

Pour obtenir la probabilité dans un intervalle variant de a à b, il faut trouver la fonction de densité de probabilité f(X) de la VAC. Pour qu'une telle fonction f(X) puisse servir, il faut qu'elle vérifie les propriétés suivantes :

$$(1)\ f(X) \geq 0 \qquad -\infty < X < \infty$$

$$(2)\ \int_{-\infty}^{\infty} f(x)dx = 1$$

Quelques caractéristiques :

1. La fonction de distribution cumulative :

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

2. La probabilité de se trouver entre a et b :

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$$

3. L'espérance mathématique de la loi de probabilité :

$$\mu = E(X) = \int_{-\infty}^{\infty} x\ f(x)dx$$

4. La variance de la loi de probabilité :

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2\ f(x)dx$$

2. Distribution uniforme

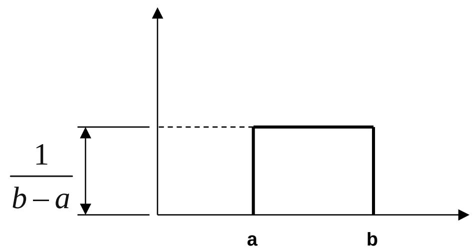
[LRS, p.181] [B, p. 220]

La valeur de la VAC a une chance égale de se produire dans l'intervalle défini entre a et b. La fonction de densité de probabilité est donnée par

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \text{ailleurs} \end{cases}$$

$$\mu = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$



3. Distribution normale

[LRS, p. 184] [B, 221- 236]

Importance de la loi normale :

1. **Plusieurs phénomènes mesurés avec une VAC peuvent être représentés exactement ou approximativement par la loi normale.**
2. **On peut utiliser la loi normale pour obtenir une approximation de certaines distributions de probabilité discrètes.**
3. **La loi normale fournit une base pour l'inférence statistique classique.**

Fonction de densité de probabilité :

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{1}{2}\left[\frac{x-\mu_X}{\sigma_X}\right]^2}$$

Normalisation de la loi normale : [LRS, p. 186] [B, p. 223]

Posons que

$$\boxed{z = \frac{x - \mu_X}{\sigma_X}}$$

Cote Z
ou
variable centrée réduite

Dès lors,

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

$$P(X \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_a} e^{-\frac{t^2}{2}}\, dt \qquad z_a = \frac{a - \mu_X}{\sigma_X}$$

Rappel :

$$\int_0^{\infty} e^{-a^2x^2} dx = \frac{\sqrt{\pi}}{2a} \quad \Rightarrow \quad \int_{-\infty}^{\infty} \boxed{e^{-a^2x^2}} dx = 2\int_0^{\infty} e^{-a^2x^2} dx$$

Fonction paire

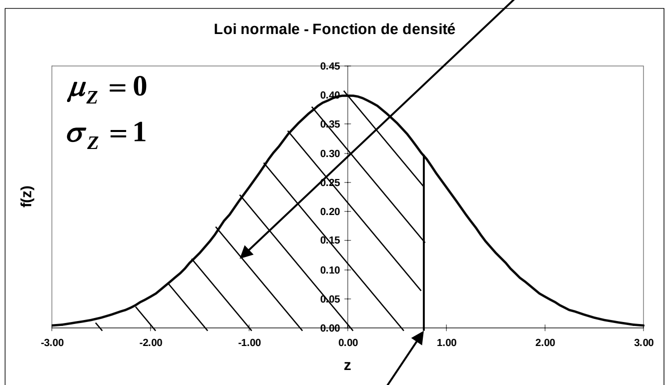
$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \frac{2\sqrt{2\pi}}{2} = \sqrt{2\pi}$$

Donc

$$\boxed{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = 1}$$

Graphique de la loi normale : [LRS, p. 187] [B, p. 222]

$$P\left(Z \le z\right) = F\left(z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$



z

Moyenne = médiane = mode

Calcul d'une probabilité avec la table : [LRS, p. A-4] [B, p. 224]

Trouver P(Z ≤ 1,00).

P(Z ≤ 1,00) =

0,5 + 0,3413

= 0,8413

Aire = 0,5

TABLE A.2

THE STANDARD NORMAL DISTRIBUTION

ENTRY REPRESENTS AREA UNDER THE STANDARD NORMAL DISTRIBUTION FROM THE MEAN TO Z

Z	$\boxed{.00}$	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	[illegible]	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
$\boxed{1.0}$	$\boxed{.3413}$	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767

Exemples de calcul de probabilités :

1. Trouver la probabilité associée à des valeurs : **[LRS, p. 189]**

$x = 0,500$
$\mu_X = 0,503$
$\sigma_X = 0,004$

$$\Rightarrow \quad z = \frac{0,500 - 0,503}{0,004} = -0.750$$

Variable centrée réduite

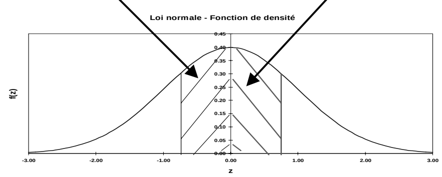
P(0,500 ≤ X ≤ 0,503) = ?

Solution :

En variable réduite, il faut trouver (- 0,750 ≤ Z ≤ 0,000)

symétrie

P(- 0,750 ≤ Z ≤ 0,000) **P(0,000 ≤ Z ≤ 0,750)**



Par la symétrie de la fonction,
P(- 0,750 ≤ Z ≤ 0,000) =
P(0,000 ≤ Z ≤ 0,750)

0,2734
d'après la table

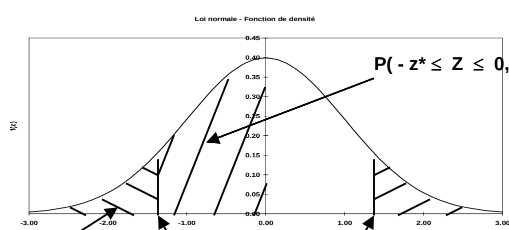
2. Trouver une valeur correspondant à une probabilité : [LRS, 191] [B, p. 225]

$\mu_X = 0,503$
$\sigma_X = 0,004$
x qui sera dépassée par 93 % des valeurs de X = ?

Solution :

Alors il faut trouver la valeur x de X pour laquelle 7 % des valeurs lui sont inférieures. Il faut donc trouver la valeur de z* telle que

P(- ∞ ≤ Z ≤ z*) = 0,07



P(- z* ≤ Z ≤ 0,00) = 0,43 = P(0,00 ≤ Z ≤ z*)

P(- ∞ ≤ Z ≤ z*) = 0,07

symétrie

[LRS, p. 192]

Pour P = 0,43

Z = 1,4 + 0,08 = 1,48

Par la symétrie,
$Z^* = -1{,}48$

Or

$$z^* = \frac{x^* - \mu_X}{\sigma_X}$$

$$\Rightarrow \quad x^* = \mu_X + z^* \sigma_X$$

$X^* = 0{,}503 - 1{,}48 \times 0{,}004$
i.e.
$x^* = 0{,}49708$

TABLE A.2

THE STANDARD NORMAL DISTRIBUTION

ENTRY REPRESENTS AREA UNDER THE STANDARD NORMAL DISTRIBUTION FROM THE MEAN TO Z

Z	.00	.01	.02	.03	.04	.05	.06	.07	$\boxed{.08}$	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	[illegible]	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
$\boxed{1.4}$	[illegible]	.4207	.4222	.4236	.4251	.4265	.4279	.4292	$\boxed{.4306}$	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	[illegible]	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	[illegible]	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767

P ≅ 0,43

Quelques valeurs de probabilité d'intérêt : [B, p. 230]

Quelques probabilités associées à des intervalles typiques		
X dans l'intervalle	**Intervalle centré réduit**	**Probabilité (aire sous la courbe normale)**
$\mu \pm \sigma$	± 1,0	**0,6826 ou 68,26 %**
$\mu \pm 1{,}96\ \sigma$	± 1,96	**0,9500 ou 95,00 %**
$\mu \pm 2\ \sigma$	± 2,0	**0,9544 ou 95,44 %**
$\mu \pm 2{,}58\ \sigma$	± 2,58	**0,9902 ou 99,02 %**
$\mu \pm 3\ \sigma$	± 3,0	**0,9974 ou 99,74 %**

4. Distribution normale p/r aux distributions binomiales et de Poisson

[LRS, p.196] [B, p. 232]

(a) La loi normale peut servir à obtenir une valeur approximative de la distribution binomiale lorsque

$$\sigma_X^2 = np\left(1-p\right) \geq 10$$

[LRS, p. 196]

si $p \leq 0.5$ il faut $np \geq 5$

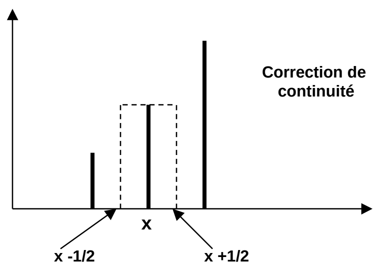
si $p \geq 0.5$ il faut $n\left(1-p\right) \geq 5$

[B, p.232]

Dès lors,

$$\mu_X = np$$

$$z = \frac{x-np}{\sqrt{np\left(1-p\right)}}$$



$$P\left(X = x \middle| n;p\right) = P\left(x - \tfrac{1}{2} \leq X \leq x + \tfrac{1}{2}\right)$$

(b) La loi normale peut servir à obtenir une valeur approximative de la distribution de Poisson lorsque le paramètre **[LRS, p.197]**

$$\lambda \geq 5$$

Dès lors,

$$\mu_X = \lambda$$

$$\sigma_X = \sqrt{\lambda}$$

$$z = \frac{x-\lambda}{\sqrt{\lambda}}$$

5. La distribution log-normale

[LRS, p. 206] [B, p. 256]

Utilisation de ln X au lieu de X. Posons que

$$Y = \ln X \qquad X > 0$$

La distribution log-normale devient :

$$f(x) = \frac{1}{\sqrt{2\pi}\, e^{y}\sigma_Y} e^{-\frac{1}{2}\left[\frac{y-\mu_Y}{\sigma_Y}\right]^2} \qquad x > 0$$

$$E(X) = \mu_X = e^{\mu_Y + \sigma_Y^2/2}$$

$$\sigma_X = \mu_X \sqrt{\left(e^{\sigma_Y^2} - 1\right)}$$

$$z = \frac{\ln(x) - \mu_Y}{\sigma_Y}$$ pour utiliser la table classique de la loi normale

6. La distribution exponentielle

[LRS, p.209] [B, p. 238]

La distribution exponentielle s'utilise surtout dans la théorie des files d'attente pour modéliser la durée de temps entre les arrivées à un poste de service, par exemple un poste de contrôle à un pont. Elle est caractérisée par un seul paramètre, λ = le nombre moyen des arrivées par unité de temps :

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0$$

$$\mu = \frac{1}{\lambda}$$

$$\sigma = \frac{1}{\lambda}$$

$$P(\text{temps d'arrivée} \leq X) = 1 - e^{-\lambda t}$$

Exemple d'application : [LRS, p. 210]

Soit une moyenne d'arrivées de 20 clients à l'heure à un guichet automatique d'une banque. Quelle est la probabilité que le prochain client arrive dans les prochaines 6 min ?

$$\lambda = 20$$

$$x = 0,1\,\text{h}$$

$$P\left(\text{temps d'arrivée} \le 0,1\right) = 1 - e^{-(20\times 0,1)} = 1 - 0,1353 = 0,8647$$

De plus,

$$\mu = \sigma = \frac{1}{20} = 0,05\,\text{h ou } 3\,\text{min}$$

7. La distribution de Gumbel pour les valeurs extrêmes

Valeur extrême : valeur min ou max choisie dans un ensemble de données.

Exemple :

Le débit de crue annuel max est le plus gros débit obtenu pendant une année spécifiée.

Pour plusieurs années consécutives d'enregistrement, l'ensemble des débits annuels max peut s'étudier statistiquement.

La distribution des valeurs extrêmes peut se représenter en particulier par la distribution de Gumbel.

$$f\left(x\right) = \frac{1}{\alpha}\exp\left[-\frac{x-u}{\alpha} - \exp\left(-\frac{x-u}{\alpha}\right)\right] \quad -\infty < x < \infty$$

$$\alpha = \frac{\sqrt{6}s_x}{\pi} \quad \text{et} \quad u = \bar{x} - 0,5772\alpha$$

8. La distribution gamma [B, p. 252]

La distribution gamma à 2 paramètres a comme fonction de densité de probabilité

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha} \boxed{\Gamma(\alpha)}} x^{\alpha-1} e^{-x/\beta} & x > 0, \alpha > 0, \beta > 0 \\ 0 & \text{ailleurs} \end{cases}$$

Fonction gamma (fonction récurrente)

$$\Gamma(\alpha) = (\alpha - 1)\ \Gamma(\alpha - 1)$$

$$\mu = \alpha\,\beta \quad \text{et} \quad \sigma^2 = \alpha\,\beta^2$$

$$\alpha_3 = 2/\sqrt{\alpha} \quad \text{et} \quad \alpha_4 = 3\left(1 + \frac{2}{\alpha}\right)$$

9. La distribution Pearson III et Log-Pearson III

La distribution Log-Pearson III est principalement utilisée en hydrologie pour l'analyse fréquentielle des débits de crue annuels max.

Il s'agit d'une distribution gamma à 3 paramètres : la moyenne, l'écart-type et le coefficient d'asymétrie sont alors incorporés aux paramètres λ, β et ε de cette distribution.

Pearson III :

$$f(x) = \frac{\lambda^{\beta} (x - \varepsilon)^{\beta-1} e^{-\lambda(x-\varepsilon)}}{\boxed{\Gamma(\beta)}} \quad x \geq \varepsilon$$

Fonction gamma

$$\lambda = \frac{s_x}{\sqrt{\beta}}$$

$$\beta = \left(\frac{2}{C_s}\right)^2 \qquad \varepsilon = \bar{x} - s_x\sqrt{\beta}$$

Log-Pearson III :

$$f(x) = \frac{\lambda^{\beta}\left(y-\varepsilon\right)^{\beta-1} e^{-\lambda(y-\varepsilon)}}{x\,\Gamma(\beta)} \qquad y = \log x \geq \varepsilon$$

$$\beta = \left[\frac{2}{C_s(y)}\right]^2$$

$$\lambda = \frac{s_y}{\sqrt{\beta}}$$

$$\varepsilon = \overline{y} - s_y\sqrt{\beta}$$

Exemples d'utilisation : voir le chapitre sur l'analyse fréquentielle.

10. La distribution de Weibull

[LRS, p.211] [B, p. 254]

Cette distribution est définie avec 2 paramètres :

Paramètre de forme $\alpha > 0$

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-\left(x/\beta\right)^{\alpha}} \qquad x \geq 0$$

Paramètre d'échelle $\beta > 0$

$$P(X \geq x) = 1 - e^{-\left(x/\beta\right)^{\alpha}} \qquad X \geq 0$$

Pour $\alpha = 1$, on retrouve la loi exponentielle, i.e.

$$f(x) = \frac{1}{\beta} e^{-x/\beta}$$

À mesure que α augmente, la loi de Weibull tend vers la loi normale.

De plus,

$$\mu = \beta\Gamma\left(1+\frac{1}{\alpha}\right)$$

$$\sigma^2 = \beta^2\left\{\Gamma\left(1+\frac{2}{\alpha}\right) - \left[\Gamma\left(1+\frac{1}{\alpha}\right)\right]^2\right\}$$

Rappel :

$$\Gamma(n+1) = n\Gamma(n)$$

11. Exemples d'application

(1) Loi normale :

La mesure du diamètre intérieur d'un tuyau de fonte ductile suit une distribution normale avec

$$\mu_X = 5{,}01 \text{ po}$$
$$\sigma_X = 0{,}03 \text{ po}$$

Les limites de spécification pour ces tuyaux sont : 5,00 $\pm$ 0,05 po.

(avec la table)

$$z_{sup} = \frac{5{,}05-5{,}01}{0{,}03} = 1{,}33 \quad \Rightarrow \quad F(z) = 0{,}5 + 0{,}408241 = 0{,}908241$$

$$z_{inf} = \frac{4{,}95-5{,}01}{0{,}03} = -2{,}0 \quad \Rightarrow \quad F(z) = 0{,}5 - 0{,}477250 = 0{,}02275$$

Donc, pour trouver le % de tuyaux non acceptables :

$$P(\text{non acceptable}) = P(X \geq 5{,}05) + P(X \leq 4{,}95)$$
$$= (1 - 0{,}908241 + 0{,}02275) = 0{,}1145 \text{ ou } 11{,}45\%$$

(2) Loi log-normale :

On a remarqué que les hauteurs annuelles des précipitations à une station pluviométrique sur un bassin versant possède une distribution log-normale. La moyenne et l'écart-type des valeurs obtenues à cette station sont :

$$\mu_X = 1800\,\text{mm}$$

$$\sigma_X = 360\,\text{mm}$$

(1) Quelle est la probabilité que la hauteur annuelle des précipitations se situent entre 800 et 2000 mm ?

(2) Quelle est la probabilité que la hauteur annuelle des précipitations soit d'au moins 1000 mm ?

Solution :

(a) P(800 < X < 2000) :

$$\sigma_X = \mu_X \sqrt{\left(e^{\sigma_Y^2} - 1\right)} \quad \Rightarrow \quad \sigma_Y^2 = \ln\left(1 + \frac{\sigma_X^2}{\mu_X^2}\right)$$

$$\boxed{\sigma_Y = \sqrt{\ln\left[1 + \left(360/1800\right)^2\right]} = 0,198}$$

$$\mu_X = e^{\mu_Y + \sigma_Y^2/2} \quad \Rightarrow \quad \mu_Y = \ln \mu_X - \sigma_Y^2/2$$

$$\boxed{\mu_Y = \ln\left(1800\right) - 0,198^2 = 7,48}$$

$$z_2 = \frac{\ln(2000) - 7,48}{0,198} \cong 0,60 \quad \text{et} \quad z_1 = \frac{\ln(800) - 7,48}{0,198} \cong -4.0$$

D'après la table,

$$P(800 < X < 2000) = \left[0,5 + F(0,60)\right] - \left[0,5 - F(4,00)\right]$$
$$= 0,2257 + 0,5 = 0,7257$$

(b) P(X) ≥ 1000 :

$P(X < 1000)$ par la table [B, p. 697]

$$P(X \geq 1000) = 1,0 - \boxed{\left(0,5 + F\left(\frac{\ln(1000) - 7,48}{0,198}\right)\right)}$$
$$= 1,0 - \left(0,5 + F(2,85)\right) = 1,0 - \left(0,5 + 0,4978\right) = 0,0022$$

(3) La loi gamma :

Le débit annuel max à un endroit spécifique le long d'une rivière suit une loi gamma à 2 paramètres avec

$$\alpha = 1,56$$
$$\beta = 6,41$$

(1) Moyenne des débits annuels max :

$$\mu_X = \alpha\beta = 1,56 \times 6,41 = 9,9996 \cong 10$$

(2) L'écart-type de ces débits :

$$\sigma_X = \beta\sqrt{\alpha} \cong 8,0$$

(3) P(X ≤ 9 m³ / s) :

ou LOI.GAMMA(9;1.56;6.41;Vrai)

Avec EXCEL : =*GAMMADIST*(9;1.56;6.41;True)

P(X ≤ 9 m³ / s) = 0,5567

11. Distributions de probabilité et EXCEL [B, p. 259]

Dans <u>EXCEL,</u> nous retrouvons les distributions de probabilités suivantes :

1. La distribution binomiale : LOI.BINOMIALE (*BINOMDIST*)

2. La distribution exponentielle : LOI.EXPONENTIELLE (*EXPONDIST*)

3. La distribution gamma : LOI.GAMMA (*GAMMADIST*)

4. La distribution log-normale : LOI.LOGNORMALE (*LOGNORMDIST*)

5. La distribution binomiale négative : LOI.BINOMIALE.NEG (*NEGBINOMDIST*)

6. La distribution normale : LOI.NORMALE (*NORMDIST*)

7. La distribution normale centrée : LOI.NORMALE.STANDARD (*NORMSDIST*)

8. La distribution Student-t : LOI.STUDENT (*TDIST*)

9. La distribution Weibull : LOI.WEIBULL (*WEIBULL*)

Dans <u>EXCEL</u> :

1. Cliquer dans la cellule où la valeur de la probabilité doit se trouver.

2. Inscrire = (pour indiquer à EXCEL qu'il s'agit d'une formule de calcul)

3. Cliquer sur l'icône f_x

4. Dans Catégorie de fonctions (*Function category*), choisir Statistiques (*Statistical*)

5. Choisir la loi de distribution appropriée, puis cliquer sur OK.

6. Fournir les valeurs demandées, puis cliquer sur OK.

7. Pour obtenir de l'aide sur la fonction, cliquer sur le ?