REVUE DE STATISTIQUE APPLIQUÉE

E. MORICE

Quelques modèles mathématiques de durée de vie

Revue de statistique appliquée, tome 14, nº 1 (1966), p. 45-126

<http://www.numdam.org/item?id=RSA_1966__14_1_45_0>

QUELQUES MODÈLES MATHÉMATIQUES DE DURÉE DE VIE

E. MORICE

SOMMAIRE

	Pages
GENERALITES SUR LES LOIS DE MORTALITE	46
I - MODELES THEORIQUES FONDAMENTAUX DE LA STATISTIQUE	49
I.1 - Distributions binomiale, binomiale négative et multinomiale	49
I.2 - Distribution de Poisson	52
I.3 - Distribution normale	54
I.4 - Distribution log-normale	57
I.5 - Distribution Gamma	58
I.6 - Distribution de Khi-carré (χ^2). Relations entre les lois Gamma, Poisson et χ^2	60
I.7 - Distribution Beta	63
I.8 - Distribution de F (Fisher-Snedecor) et de t (Shedent-Fisher). Relations entre les lois binomiale, Beta et F	66
II - ETUDE DE QUELQUES MODELES DE DUREE DE VIE	70
II.1 - Loi exponentielle	71
II.1.1 à 3 - Définition. Propriétés. Etude graphique.	71
II.1.4 - Plan d'échantillonnage (essai tronqué sans remplacement)	73
II.1.5 - Estimation de la vie moyenne (essai tronqué sans remplacement), d'après le nombre des défaillances	76
II.1.6 - Distribution de la plus petite durée de vie de k éléments	77
II.1.7 - Estimation de la vie moyenne (essai censuré avec ou sans remplacement)	77
II.1.8 - Plan d'échantillonnage (essai censuré avec ou sans remplacement)	83
II.1.9 - Estimation de la vie moyenne (essai tronqué, avec ou sans remplacement)	85
II.2 - Loi exponentielle et processus de Poisson	88
II.2.1 - Processus de Poisson	88
II.2.2. - Intervalle de temps entre arrivées	89

Pages

II.2.5	- Test de l'existence d'un processus de Poisson	92
II.2.6	- Plan d'échantillonnage (essai tronqué avec remplacement)	93
II.2.7	- Loi Gamma dans le cas de défaillance par usure	100
II.2.8	- Loi mixte de défaillance (usure et défaillance aléatoire)	101
II.3 - Loi de Weibull		103
II.3.1.3	- Définition. Propriétés. Etude graphique.	104
II.4 - Lois des valeurs extrêmes		105
II.4.1	- Loi des valeurs extrêmes. Exemple	108
II.4.2	- Lois limites de Gumbel et Fraudenthal .	109
III - Annexes		113
III.1	- Tableau récapitulatif	114
III.2	- Graphiques de f(t), R(t), et h(t) pour diverses lois	123
III.3.4	- Bibliographie (ouvrages et tables)	117

GENERALITES SUR LES LOIS DE MORTALITE.

Les premiers travaux sur les lois de survie - ou, si l'on préfère, sur les lois de mortalité - ont porté sur les populations humaines ; il a fallu plus d'un siècle pour que ces recherches s'étendent à d'autres populations statistiques, par exemple : éprouvette de métal dans les essais de fatigue, pour prendre une importance considérable dans les recherches modernes de fiabilité (composants électroniques, études de corrosion....).

C'est en 1825 qu'est apparue, dans les travaux actuariels, la loi de Gompertz.

Gompertz faisait l'hypothèse que la mort peut être provoquée par deux causes distinctes :

- l'une le hasard, l'accident qui peut priver de la vie une personne en parfaite santé, cause dont l'intensité est indépendante de l'âge,

- l'autre, qu'il nommait "force of mortality", est l'affaiblissement progressif de l'individu et correspond dans l'expression de la force de mortalité à une quantité φ qui s'accroît dans le temps dx d'une quantité $d\varphi$ proportionnelle à φ.

Négligeant la première cause, Gompertz était conduit à écrire que la force de mortalité à l'âge x, c'est-à-dire ce que nous appelons taux instantané de mortalité, satisfaisait à l'équation :

$$\frac{1}{\varphi}\frac{d\varphi}{dx} = q = Cte$$

d'où

$$\varphi = a\ e^{qx}$$

Soit V_x le nombre de vivants d'âge x à un instant donné, et V_{x+dx} le nombre des survivants au bout d'un temps dx, le quotient de mortalité relatif à l'intervalle de temps dx est alors :

$$\frac{V_x - V_{x+dx}}{V_x}$$

Le quotient moyen de mortalité par unité de temps :

$$\frac{1}{dx}\frac{V_x - V_{x+dx}}{V_x}$$

tend, lorsque dx tend vers zéro, vers la quantité : $-\frac{V'_x}{V_x}$ qui est le taux instantané de mortalité à l'âge x.

L'hypothèse de Gompertz donne alors :

$$-\frac{V'_x}{V_x} = a\ e^{qx}$$

d'où :

$$\text{Log } V_x = - \int a\ e^{qx}\ dx + \text{Cte}$$

$$V_x = k\ e^{-\int a e^{qx} dx}$$

qui détermine V_x à une constante près.

Si on pose

$$e^{-\frac{a}{q}} = g \qquad e^{q} = c$$

on obtient la loi de Gompertz sous sa forme habituelle :

$$V_x = k\ g^{c^x}$$

k, g, c étant trois constantes dont la première k sert à fixer le nombre de vivants à l'origine ; les deux autres déterminant l'allure de la courbe et caractérisant la loi de survie.

Si la loi de Gompertz - d'ailleurs modifiée par Makcham en 1849 sous la forme $V_x = k\ s^x\ g^{c^x}$ pour tenir compte de la première cause de mortalité, soit alors $\varphi = -\frac{V'_x}{V_x} = a\ e^{qx} + b$ – ne peut pas correspondre à la mortalité humaine au cours de la vie entière, ainsi que le montre une table des quotients de mortalité observés, le mécanisme employé par Gompertz reste l'un des moyens fondamentaux pour établir un modèle de loi de mortalité dès que l'on peut faire, soit en fonction d'observations, soit pour des raisons théoriques, une hypothèse plausible sur le taux instantané de mortalité et sa variation en fonction du temps.

Afin d'obtenir des formules générales, indépendantes de l'effectif des individus au début de la période des observations, on raisonnera non sur les effectifs mais sur les fréquences, c'est-à-dire sur les probabilités d'arrivée de l'évènement E auquel on s'intéresse (mort ou plus généralement "défaillance" de l'unité observée.

Soit à l'instant $T = t$; $h(t)$ le taux instantané de défaillance, ou encore $h(t)\,dt$ la probabilité d'arrivée de l'évènement E dans un court intervalle de temps $(t, t + dt)$, ou probabilité conditionnelle d'arrivée de E entre t et $t + dt$, E ne s'étant pas produit avant l'instant t.

$F(t) = Pr(T < t)$ la probabilité d'arrivée de E entre 0 et t (par exemple la probabilité de défaillance d'un élément ou d'un ensemble)

$1 - F(t)$ sera alors la probabilité pour que E ne se produise pas entre 0 et t (probabilité de survie à l'époque t).

Pour que l'évènement se produise précisément entre t et $t + dt$, [probabilité $F(t + dt) - F(t)$], il faut qu'il ne se soit pas produit entre 0 et t [probabilité $1 - F(t)$], et que ceci étant réalisé, l'évènement se produise entre t et $t + dt$, soit, d'après le théorème des probabilités composées :

$$F(t + dt) - F(t) = [1 - F(t)]\, h(t)\, dt$$

$$\frac{dF(t)}{1 - F(t)} = h(t)\, dt$$

$$\left\{\log_e [1 - F(t)]\right\}_0^t = \int_0^t - h(t)\, dt$$

$$1 - F(t) = [1 - F(0)] \exp\left[- \int_0^t h(t)\, dt\right]$$

Compte tenu de la condition initiale :

$$F(0) = 0$$

on aura

$$\boxed{F(t) = 1 - \exp - \left[\int_0^t h(t)\, dt\right]}$$

<u>(fonction de répartition de la variable aléatoire T)</u>

On en déduit :

$$\boxed{f(t) = \frac{dF(t)}{dt} = h(t) \exp\left[- \int_0^t h(t)\, dt\right]}$$

<u>(densité de probabilité de la variable T pour T = t)</u>

$f(t)\,dt$ représente la proportion des éléments qui, "vivant" à l'instant $t = 0$, "mourront" entre t et $t + dt$.

La variable aléatoire T, instant de la défaillance est donc telle que :

$$f(t)\, dt = Pr\,[t < T < t + dt] = h(t) \exp\left[- \int_0^t h(t)\, dt\right] dt$$

La loi des défaillances pourra donc être définie par l'une quelconque des fonctions $f(t)$, $F(t)$ et $h(t)$ liées par les relations :

$$\boxed{h(t) = \frac{f(t)}{1 - F(t)}} \qquad \boxed{F(t) = \int_0^t f(t)\, dt}$$

La fonction $R(t) = 1 - F(t)$ sera appelée <u>fiabilité</u> de l'unité considéré.

I - MODELES THEORIQUES FONDAMENTAUX DE LA STATISTIQUE

Avant d'étudier quelques lois de mortalité utilisées dans les études de fiabilité, rappelons d'abord les propriétés de quelques distributions classiques.

I.1 - DISTRIBUTION BINOMIALE

Soit E un évènement (par exemple rupture d'un fil sous une contrainte donnée) dont la probabilité d'arrivée au cours d'une épreuve est égale à p, la probabilité que l'évènement E se produise x fois au cours de n épreuves indépendantes, (par exemple n éléments supposés identiques, soumis à la même contrainte) est :

$$f(x) = C_n^x p^x q^{n-x} = \frac{n!}{x!\,(n-x)!} p^x q^{n-x} \quad , \; q = 1 - p$$

$$(x = 0, 1, \ldots\ldots, n)$$

La variable aléatoire X, variable discrète pouvant prendre les valeurs 0, 1, ..., n est appelée variable binomiale $\mathcal{B}(n, p)$; f(x) étant sa <u>fonction de distribution</u> de la variable X et :

$$F(x) = \Pr\,[X \leqslant x] = \sum_{0}^{x} C_n^x p^x q^{n-x},$$

sa <u>fonction de répartition</u>.

Deux caractéristiques fondamentales de cette distribution sont :

son espérance mathématique ou valeur moyenne théorique du nombre d'arrivées de l'évènement en n épreuves :

$$E(X) = m = np$$

et sa variance :

$$V(X) = \sigma^2 = npq$$

Si on considère la fréquence x/n du nombre d'arrivées dans une série de n épreuves on aura :

$$E(X/n) = p$$

$$V(X/n) = \frac{pq}{n}$$

L'expérience, c'est-à-dire l'observation de x arrivées en n épreuves, permettra une estimation f = x/n de p, estimation d'autant plus précise que n est plus grand.

Propriétés

Si X_1, ... X_i, ... X_k sont des variables binomiales indépendantes de même paramètre p, avec n = n_1, ... n_i, ... n_k la variable $Z = \sum_{i=1}^{k} X_i$ est une variable binomiale $\mathcal{B}(N = \sum n_i, p)$.

I.1.2 - Distribution géométrique et distribution binomiale négative (distribution de Pascal)

La distribution binomiale correspond au cas où le nombre d'épreuves n étant fixé à l'avance, la variable aléatoire X est le nombre d'arrivées de l'évènement E.

Considérons maintenant le cas où le nombre c des arrivées de E est fixé à l'avance et où l'on fait des épreuves successives indépendantes jusqu'à la $c^{ième}$ arrivée de E : le nombre d'épreuves nécessaires pour obtenir ce résultat est une variable aléatoire N telle que :

$$Pr(N = n) = C_{n-1}^{c-1} p^c q^{n-c} = C_{n-1}^{n-c} p^c q^{n-c}$$

$$n = c, c + 1, ...$$

(Probabilité de c − 1 arrivées en n − 1 épreuves indépendantes suivies d'une arrivée de E à la $n^{ième}$ épreuve).

On peut aussi considérer la variable aléatoire Z = N − c, nombre d'épreuves nécessaires en supplément du nombre minimum c pour obtenir la $c^{ième}$ arrivée :

$$Pr(Z = z) = Pr [N = z + c] = C_{z+c-1}^{c-1} p^c q^z = C_{z+c-1}^{z} p^c q^z$$

$$z = 0, 1, ...$$

(Probabilité pour que le $c^{ième}$ évènement E se produise à l'épreuve de rang z + c).

(Ce résultat est quelquefois présenté sous la forme :

$$Pr(Z = z) = C_{-c}^{z} p^c (- q)^z$$

avec :

$$C_{-c}^{z} = \frac{(- c)(- c - 1) \ldots\ldots\ldots (- c - z + 1)}{Z!} = (- 1)^z C_{z+c-1}^{z}$$

d'où le nom de binomiale négative : la probabilité Pr(Z = z) est donnée en effet par le terme en q^z du développement illimité de $p^c(1 - q)^{-c}$

Les caractéristiques principales de ces distributions sont :

$$E(N) = \frac{c}{p} \qquad E(Z) = \frac{q^c}{p}$$

$$V(N) = V(Z) = \frac{q^c}{p^2}$$

Pour c = 1, on a :

$$\Pr(N = n|c = 1) = q^{n-1}p$$

$$\Pr(Z = z) = q^z p$$

(distributions géométriques, Z étant le nombre des épreuves avant la première arrivée de E).

La probabilité de faire au plus n épreuves pour voir se produire la $c^{ième}$ arrivée de E est :

$$\Pr[N \leqslant n] = \sum_{j=c}^{n} C_{j-1}^{c-1} p^c q^{j-c}$$

Cette probabilité est aussi égale à la probabilité pour que le nombre X d'arrivées de E en n épreuves soit égal ou supérieur à c, d'où :

$$\Pr[N < n/X = c] = \Pr[X \geqslant c/N = n] = \sum_{x=c}^{n} C_n^x p^x q^{n-x} ,$$

qui pourra être évaluée à l'aide des tables de la loi binomiale ou des approximations éventuellement utilisables.

Remarque

Si les épreuves sont faites à intervalles réguliers on peut considérer N = T comme la durée d'attente de la $c^{ième}$ arrivée de l'évènement E.

Si les intervalles de temps dt entre épreuves successives deviennent infiniment petits et si on pose $p = \lambda dt$, λ étant une constante positive, les lois de X et de N tendent respectivement vers des lois limites continues qu'on étudiera ultérieurement :

$$\Pr[X = x|T = t] = e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

(loi de Poisson, forme limite de la loi binomiale)

$$\Pr[t < T < t + dt|X = c] = \frac{1}{\Gamma(c)} \lambda e^{-\lambda t} (\lambda t)^{c-1} dt$$

(loi Gamma, forme limite de la loi binomiale négative, qui pour c = 1 se réduit à :

$$\Pr[t < T < t + dt|X = 1] = \lambda e^{-\lambda t} dt$$

(loi exponentielle, forme limite de la loi géométrique)

I.1.3 - Distribution multinomiale

Dans le cas précédent chacune des n épreuves indépendantes pouvait donner lieu à deux éventualités (arrivée ou non arrivée de l'évènement E).

Considérons maintenant le cas où chaque épreuve peut donner lieu à l'une où l'autre de k éventualités mutuellement exclusives E_1 ou E_2 ..... E_k de probabilités respectives p_1, p_2, ... p_k avec :

$$\sum_{i=1}^{k} p_i = 1$$

Si X_i est le nombre aléatoire d'arrivées de E_i en n épreuves indépendantes, on aura :

$$\Pr [X_1 = x_1 \ldots\ldots X_k = x_k] = \frac{n!}{x_1!\, x_2! \ldots\, x_k} p_1^{x_1} \ldots\ldots p_k^{x_k}$$

avec :

$$\sum_{i=1}^{k} x_i = n$$

Pour les diverses variables X_i, on aura :

$$E(X_i) = np_i$$
$$V(X_i) = np_i(1 - p_i)$$

L'observation de x_i arrivées de l'évènement E_i au cours de n épreuves permettra encore une estimation $f_i = \frac{x_i}{n}$ de p_i avec une variance

$$V\left(\frac{X_i}{n}\right) = \frac{p_i(1 - p_i)}{n}$$

I.2 - DISTRIBUTION DE POISSON (fig. 1)

Une variable aléatoire discrète X pouvant prendre les valeurs $x = 0, 1, \ldots\ldots \infty$ est distribuée suivant une loi de Poisson de paramètre m si l'on a :

$$\Pr(X = x) = f(x) = e^{-m} \frac{m^x}{x!}$$

$$\Pr(X < x) = F(x) = \sum_{0}^{x} e^{-m} \frac{m^x}{x!}$$

La variable X est appelée variable $\mathfrak{L}(m)$.

Les caractéristiques fondamentales de cette loi sont :

$$E(X) = m$$
$$V(X) = m$$

On peut considérer la loi de Poisson comme une forme limite de la loi binomiale, lorsque $n \longrightarrow \infty$, $p \longrightarrow o$, le produit np tendant vers une valeur finie $np = m$.

Pratiquement cette approximation est valable pour $p < 0,1$.

Propriétés

Si $X_1, \ldots. X_i, \ldots. X_k$ sont des variables de Poisson indépendantes de paramètres $m_1, \ldots. m_k$, la variable $Z = \sum_{i=1}^{k} X_i$ est une variable de Poisson de paramètre $M = \sum m_i$ variable $\mathfrak{L}(\sum m_i)$.

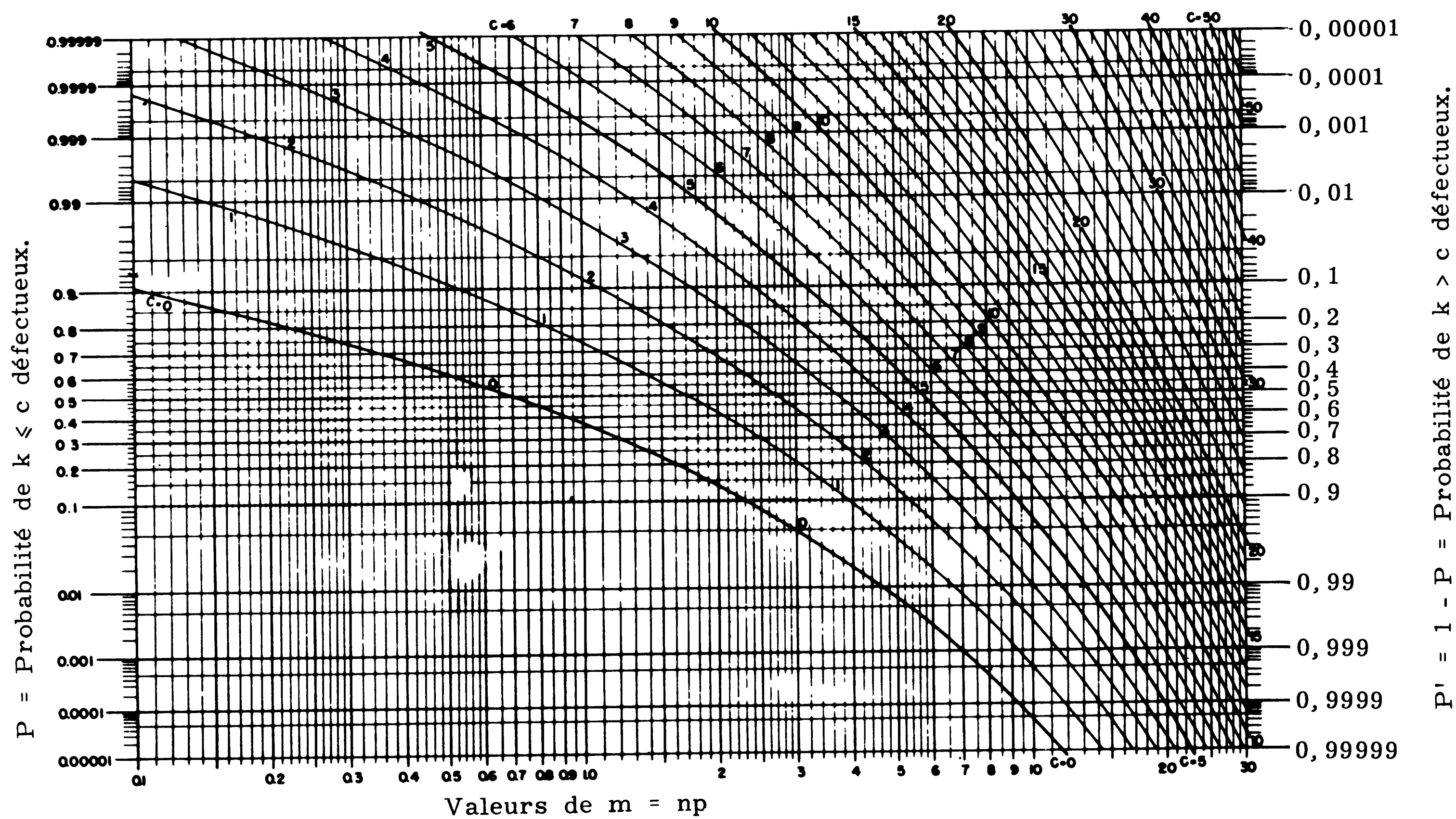


Figure 1 : Loi de Poisson (Probabilités cumulées,

(noter que ce résultat ne s'étend pas à une combinaison linéaire quelconque de variables de Poisson, ni en particulier à la différence de deux variables de Poisson).

I.3 - DISTRIBUTION NORMALE

Une variable continue X pouvant prendre toute valeur $x(-\infty < x < +\infty)$ est distribuée suivant la loi normale de paramètres m et σ, si l'on a :

$$f(x)\,dx = \Pr(x < X < x + dx) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx,$$

d'où :

$$F(x) = \Pr(X < x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx$$

On dira que X est une variable normale $\mathcal{N}(m, \sigma)$, $f(x)$ étant la densité de probabilité en $X = x$, et $F(x)$ sa fonction de répartition.

Si on pose :

$$U = \frac{X - m}{\sigma}$$

on dira que U est une variable normale réduite $\mathcal{N}(0, 1)$ dont les lois tabulées sont respectivement :

$$f(u)\,du = \Pr(u < U < u + du) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}\, du$$

$$F(u) = \Pr(U < u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}\, du$$

On désignera par u_α la valeur partuiculière de u telle que :

$$\alpha = \Pr(U < u_\alpha)$$

u_α est le quantile (ou fractile) d'ordre α de la distribution normale réduite.

Les caractéristiques fondamentales de la loi de la variable X sont précisément les paramètres qui servent à la définir :

$$\text{Moyenne : } E(X) = m = \text{Mode} = \text{Médiane}$$

$$\text{Variance : } V(X) = \sigma^2, \quad \text{Ecart-type : } \sigma(X) = \sigma.$$

m et σ^2 peuvent être estimés respectivement à partir des n observations d'un échantillon par :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad E(\bar{x}) = m$$

$$s^2 = \frac{1}{n-1}\sum_{i}^{n} (x_i - \bar{x})^2, \qquad E(s^2) = \sigma^2,$$

($\bar{x}$ et s^2 sont des estimations sans biais de m et σ^2).

I.3.1 - Etude graphique [11]

L'étude graphique d'une distribution normale à partir d'un échantillon se fait facilement à l'aide d'un graphique à échelle fonctionnelle dont l'axe des ordonnées est une échelle des probabilités cumulées F(u) de la loi normale (fig. 2).

Si F_i est la fréquence cumulée des observations inférieures à x_i dans l'échantillon, il suffira de placer sur le graphique les points (x_i, F_i) qui, dans le cas d'un échantillon provenant d'une population normale, devront se placer au voisinage d'une droite (droite de Henry) d'équation :

$$u = \frac{x - m}{\sigma}$$

avec :

$$F(x_i) = \int_{-\infty}^{u_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

Cette droite permettra :

1/ de constater, s'il y a lieu, le caractère normal de la distribution,

2/ s'il en est ainsi d'estimer m et σ

$$x = m \qquad \text{pour } u = 0, \ F(x) = 0{,}500$$

$$\sigma = x_1 - m \qquad \text{pour } u_1 = 1, \ F(x_1) = 0{,}841$$

$$\sigma = \frac{1}{x}(x_2 - m) \qquad \text{pour } u_2 = 2, \ F(x_2) = 0{,}977$$

I.3.2 - Propriétés

1/ Si X_1, ... X_i, ... X_k sont des variables normales indépendantes de paramètres respectifs (m_1, σ_1) ..... (m_i, σ_i) ..... (m_k, σ_k) la variable $Z = \sum_{i=1}^{k} X_i$ est aussi une variable normale de paramètres

$$m = m_1 + \dots + m_k \qquad \sigma = \sqrt{\sigma_1^2 + \dots + \sigma_k^2}$$

2/ Théorème limite centrale

D'une façon très générale, si X_1 ...... X_k sont des variables aléatoires indépendantes, de somme Z, telles que la variance σ_i^2 de chacune soit petite comparée à la variance de Z, la variable :

$$u = \frac{Z - E(Z)}{\sqrt{V(Z)}}$$

tend vers la variables normale réduite $\mathcal{N}(0, 1)$ lorsque $n \longrightarrow \infty$.

Pratiquement, ceci revient à dire que si Z résulte d'un grand nombre de causes indépendantes produisant des effets additifs X_1 ..... X_k, tels que la variance due à chaque effet soit faible devant la variance résultante (chaque cause entrant pour une part négligeable dans la variabilité finale), Z suit une loi normale.

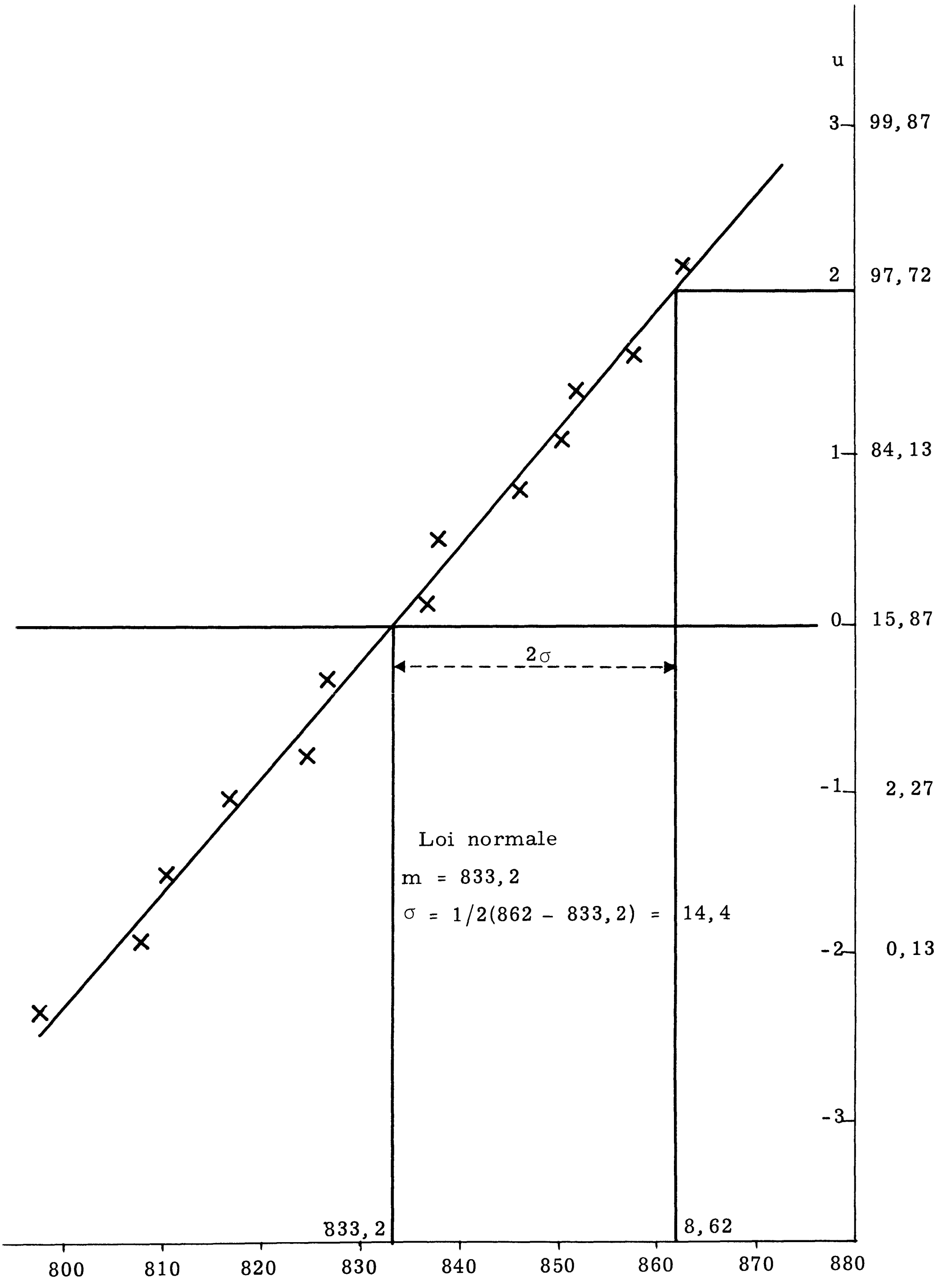


Fig. 2 - Loi normale. Droite de Henry

I.3.3 - La loi normale comme forme limite des lois précédentes

Les approximations ci-après, conséquences des propriétés d'additivité des lois binomiale et de Poisson, résultent directement de ce théorème général.

a) Loi binomiale

La variable binomiale $\mathcal{B}(n, p)$ étant la somme de n variables binomiales $\mathcal{B}(1, p)$, il en résulte que, pour $n \longrightarrow \infty$, la variable binomiale $X = \mathcal{B}(n, p)$ tend vers une variable normale $\mathcal{N}$ $(m = np, \sigma = \sqrt{npq})$ soit :

$$\frac{X - np}{\sqrt{npq}} \longrightarrow \mathcal{N}(0, 1)$$

Cette approximation, pratiquement valable pour $npq > 15$, sera améliorée en tenant compte de la correction de continuité soit :

$$Pr(a \leqslant X \leqslant b) = F\left(\frac{b + 1/2 - np}{\sqrt{npq}}\right) - F\left(\frac{a - 1/2 - np}{\sqrt{npq}}\right)$$

avec :

$$F(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

b) Loi de Poisson

De même la variable de Poisson $X = \mathcal{P}(m)$ étant la somme de m variables $\mathcal{P}(1)$, il en résulte que pour $m \longrightarrow \infty$, la variable X tend vers la variable normale $\mathcal{N}(m, m)$, soit :

$$\frac{X - m}{\sqrt{m}} \sim \mathcal{N}(0, 1)$$

Cette approximation pratiquement valable pour $m > 15$ sera améliorée comme ci-dessus en utilisant :

$$Pr(a \leqslant X \leqslant b) = F\left(\frac{b + 1/2 - m}{\sqrt{m}}\right) - F\left(\frac{a - 1/2 - m}{\sqrt{m}}\right)$$

c) Remarque

Le fait que dans une distribution normale $\mathcal{N}(m, \sigma)$, la variable X peut varier de $-\infty$ à $+\infty$, n'a pratiquement aucune importance dans l'étude de la distribution d'une variable essentiellement positive dès que le rapport m/σ est suffisamment grand, la probabilité pour que $x < m - 4\sigma$, par exemple étant négligeable.

I.4 - DISTRIBUTION LOG-NORMALE

Une variable aléatoire X est distribuée suivant une loi log-normale si la variable $Y = \log_{10} X$, ou, plus généralement, $Y = \log_{10}(X - a)$, pour $x > a$, est distribuée suivant une loi normale $\mathcal{N}(m, \sigma)$, m et σ étant respectivement la moyenne et l'écart-type de $\log_{10}(X - a)$.

Pour une loi donnée, m, σ et a étant connus, on aura, après avoir calculé $y = \log(x - a)$

$$\Pr(X < x) = \Pr(Y < y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}} du, \quad \left(u = \frac{y - m}{\sigma} \right)$$

Les caractéristiques de la loi de la variable aléatoire X sont :

$$E(X) = a + e^{m' + \frac{1}{2}\sigma'^2}$$

$$V(X) = e^{2m' + \sigma'^2}(e^{\sigma'^2} - 1)$$

$$\text{Mode} = a + e^{m' - \sigma'^2}$$

$$\text{Médiane} = a + e^{m'}$$

avec $m' = 2{,}3026\ m$, $\sigma' = 2{,}3026\ \sigma$

L'estimation de σ et σ^2 à partir d'un échantillon pourra être faite en calculant :

$$\bar{y} = \frac{\Sigma y}{n} \quad \text{et} \quad s^2 = \frac{1}{n - 1} \Sigma (y - \bar{y})^2$$

I.4.1 - Etude graphique (11)

L'emploi d'un papier pour graphique de Henry permettra encore d'étudier le caractère log-normal d'un distribution à partir d'un échantillon et d'estimer la moyenne et l'écart-type de log X ou de log(X - a) : on utilisera de préférence un graphique à échelle logarithmique en abscisse, évitant ainsi la transformation Y = log(X - a) et facilitant, s'il y a lieu, l'estimation graphique du troisième paramètre a, si c'est log(X - a) et non log X qui est distribuée normalement (fig. 3).

I.5 - DISTRIBUTION GAMMA

I.5.1 - Fonction Gamma

La fonction Gamma de paramètre k est définie par :

$$\Gamma(k) = \int_0^{\infty} e^{-x} x^{k-1} dx, \quad k > 0$$

D'une manière générale on a :

$$\Gamma(k + 1) = k\,\Gamma(k)$$

d'où, en particulier pour k entier :

$$\Gamma(k) = (k - 1)!$$

on a aussi $\Gamma(1/2) = \sqrt{\pi}$.

I.5.2 - Distribution Gamma

On dit qu'une variable X ($0 < x < \infty$) est une variable $\gamma(k)$, de paramètre k si :

$$\Pr(x < X < x + dx) = f(x)\,dx = \frac{e^{-x}\,x^{k-1}\,dx}{\Gamma(k)}$$

d'où :

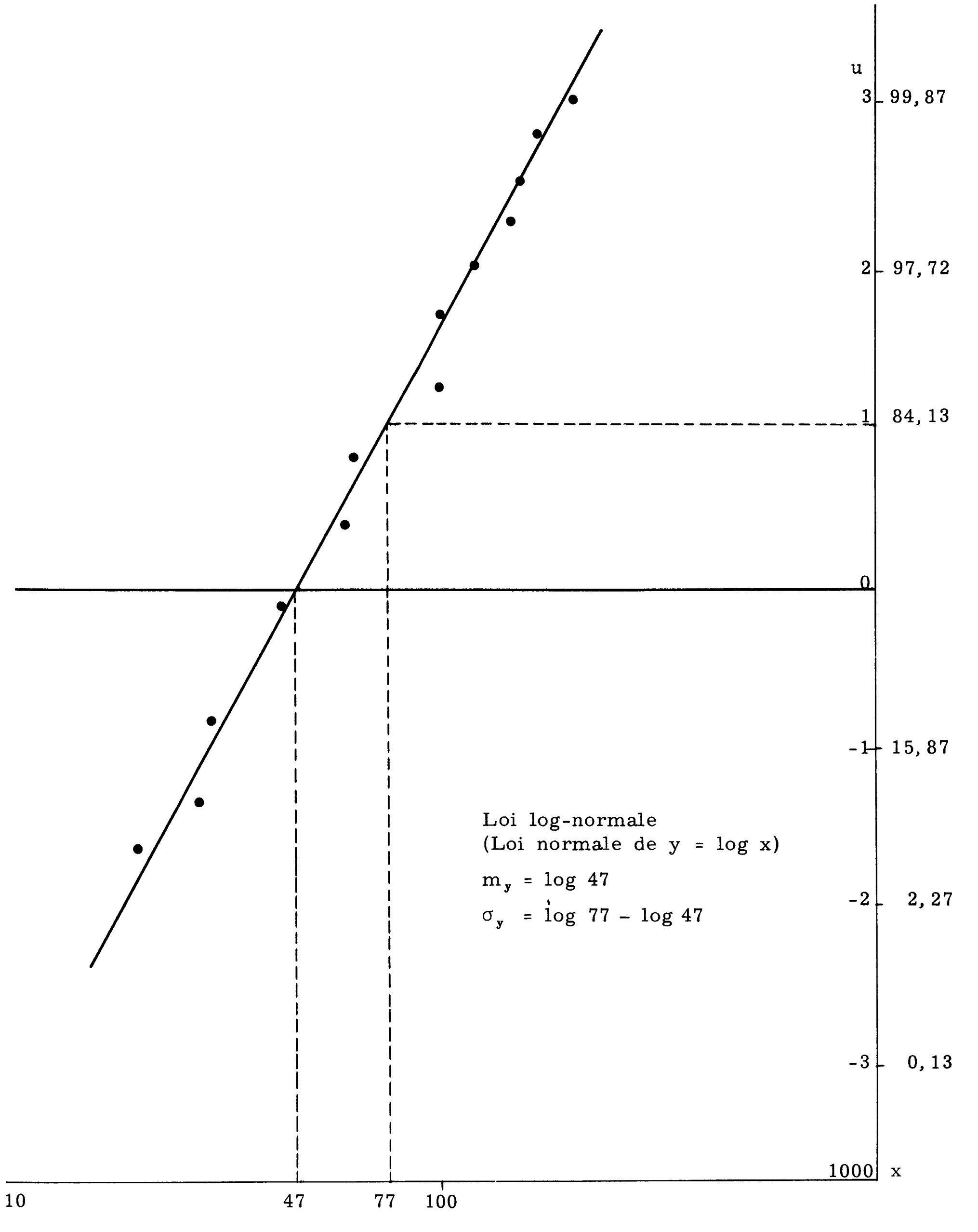


Fig. 3 - Loi log-normale.

$$F(x) = Pr(X \leqslant x) = \frac{1}{\Gamma(k)} \int_0^x e^{-x} x^{k-1} dx = \frac{\Gamma_x(k)}{\Gamma(k)},$$

$\Gamma_x(k)$ désignons la fonction gamma incomplète.

Les valeurs caractéristiques de la loi Gamma sont :

Mode : $k - 1$

Médiane : M telle que $\int_0^M e^{-x} x^{k-1} = \frac{1}{2} \Gamma(k)$

Moyenne : $E(X) = k$

Variance : $V(X) = k$

I.5.2.1 - Propriétés

Si $Z = \sum_1^n X_i$ est une somme de variables $\gamma(k_i)$ indépendantes, Z est aussi une variable $\gamma(k)$ de paramètre $k = \sum k_i$.

Cas particulier : Pour $k = 1$, on obtient la loi exponentielle particulière : $f(x) = e^{-x}$.

I.5.2.2 - Approximation normale de la loi Gamma

Si $k \nearrow \infty$, la distribution de la variable $X = \gamma(k)$ tend vers la distribution normale de moyenne k et d'écart-type $\sqrt{k}$

$$\frac{X - k}{\sqrt{k}} \rightsquigarrow \mathcal{N}(0, 1)$$

Cette approximation est pratiquement valable pour $k > 50$, soit :

$$F(x) = \int_{-\infty}^{\frac{x-k}{\sqrt{k}}} \frac{1}{\sqrt{2n}} e^{\frac{-u^2}{2}} du$$

Pour $15 < k < 50$ on peut utiliser l'approximation plus précise

$$\sqrt{4X} - \sqrt{4k - 1} \rightsquigarrow \mathcal{N}(0, 1)$$

I.6 - DISTRIBUTION DE KHI-CARRE : $\chi^2(\nu)$

La distribution de χ^2 n'intervient pas directement comme loi de durée de vie, mais elle est étroitement liée à la distribution Gamma et à la distribution de Poisson.

Par définition, c'est la loi de la distribution de la somme des carrés de ν variables normales réduites indépendantes :

$$\chi^2(\nu) = \sum_1^{\nu} u_i^2 \qquad \text{avec } u_i = \mathcal{N}(0, 1)$$

Pour $\nu = 1$; la loi de la variable $V = \frac{1}{2} u^2 = \frac{1}{2} \chi^2(1)$ est une loi Gamma de paramètre $\frac{1}{2}$ (V est une variable $\gamma(1/2)$.

On a, en effet, par changement de variable, à partir de la loi normale de u :

$$f(V)\ dV = \frac{V^{-1/2}\ e^{-V}}{\Gamma(1/2)}\ dV$$

Compte tenu de la propriété d'additivité des variables Gamma, la variable $1/2\ \chi^2(\nu) = \sum_1^{\nu} V_i$ est une variable Gamma de paramètre $\nu/2$, soit :

$$f\left(\frac{\chi^2}{2}\right)\ d\left(\frac{\chi^2}{2}\right) = \frac{1}{\Gamma(\nu/2)}\ e^{-\frac{\chi^2}{2}} \left(\frac{\chi^2}{2}\right)^{\frac{\nu}{2}-1}\ d\left(\frac{\chi^2}{2}\right)$$

d'où :

$$f(\chi^2)\ d(\chi^2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\ e^{-\frac{\chi^2}{2}} (\chi^2)^{\frac{\nu}{2}-1}\ d(\chi^2)$$

Les valeurs caractéristiques de la loi de χ^2 sont :

Mode : $\nu - 2$

Médiane : M telle que $\int_0^M f(\chi^2)\ d(\chi^2) = 1/2$

Moyenne : $E(\chi^2) = \nu$

Variance : $V(\chi^2) = 2\nu$

Cas particulier : pour $\nu = 2$, χ^2 est distribué suivant la loi exponentielle $f(\chi^2) = \frac{1}{2}\ e^{-\frac{1}{2}\chi^2}$

De l'additivité des variables Gamma, il résulte aussi que si $Z = \sum_{i=1}^{n} \chi^2(\nu_i)$ est une somme de variables $\chi^2(\nu_i)$, indépendantes. Z est aussi une variable $\chi^2(\nu)$ de paramètre $\nu = \sum_1^n \nu_i$

(ν est appelé nombre de degrés de liberté de la loi de χ^2).

I. 6. 1 - Approximation par la loi normale

Si $\nu \nearrow \infty$, la loi de $\chi^2(\nu)$ tend vers la loi normale de moyenne ν et d'écart-type $\sqrt{2\nu}$:

$$\frac{\chi^2(\nu) - \nu}{\sqrt{2\nu}} \rightsquigarrow \quad (0,\ 1)$$

Cette approximation est pratiquement valable pour $\nu > 100$

Pour $30 < \nu < 100$, on peut utiliser l'approximation plus précise :

$$\sqrt{2\ \chi^2(\nu)} - \sqrt{2\ \nu - 1} \rightsquigarrow \mathcal{N}(0,\ 1)$$

Le fractile d'ordre α de la distribution de χ^2 défini par :

$$\alpha = \Pr [\chi^2 < \chi^2_\alpha]^{(1)}$$

a pour valeur approchée,

$$\chi^2_\alpha = \frac{1}{2} [\sqrt{2\nu - 1} + u_\alpha]^2,$$

(u_α valeur de la variable normale réduite u telle que :

$$\alpha = \Pr(u < u_\alpha)$$

Une approximation plus précise serait donnée par :

$$\chi^2_\alpha = \nu \left[1 - \frac{2}{9\nu} + u_\alpha \sqrt{2/9\nu}\right]^3$$

I. 6. 2 - Relations entre les lois de χ^2, Poisson et Gamma

Soit Y une variable $\gamma(c)$, on a :

$$\Pr [Y < m] = \frac{1}{\Gamma(c)} \int_0^m e^{-y} y^{c-1} dy$$

$$= 1 - \frac{1}{\Gamma(c)} \int_m^\infty e^{-y} y^{c-1} dy$$

Par intégration par parties, on a :

$$\int_m^\infty e^{-y} y^{c-1} dy = e^{-m} m^{c-1} + (c-1) \int_m^\infty e^{-y} y^{c-2} dy$$

On en tire, de proche en proche, si c est entier :

$$\frac{1}{\Gamma(c)} \int_m^\infty e^{-y} y^{c-1} dy = e^{-m} \left[\frac{m^{c-1}}{(c-1)!} + \frac{m^{c-2}}{(c-2)!} + \dots + \frac{m}{1} + 1\right]$$

Soit :

$$1 - \Pr [Y < m] = \Pr [X < c]$$

X étant une variable de Poisson de moyenne m.

D'autre part, on a vu (§ I. 6) qu'une variable Gamma de paramètre c est aussi une variable $1/2\ \chi^2(\nu)$ avec $\nu = 2c$ degrés de liberté.

En résumé, si X est une variable de Poisson de moyenne m, Y une variable Gamma de paramètre c et χ^2 une variable Khi-carré avec $\nu = 2c$ degrés de liberté, on a :

$$\Pr [Y < m] = \Pr [\chi^2(\nu) < 2m,\ \nu = 2c]$$

$$= \Pr [X > c - 1]$$

----

(1) D'une manière générale, dans tout ce qui suit on désignera par W_α la valeur de la variable aléatoire W telle que :

$$\alpha = \Pr [W < W_\alpha]$$

Si on envisage la loi Gamma sous la forme :

$$f(t)\,dt = \frac{\lambda e^{-\lambda t}\,(\lambda t)^{c-1}}{\Gamma(c)}\,dt,$$

on aura, pour c entier :

$$\Pr\,[T < t] = \int_0^t f(t)\,dt = \frac{1}{\Gamma(c)}\int_0^{\lambda t} e^{-y}\,y^{c-1}\,dy$$

$$= \Pr\,[\chi^2 < 2\,\lambda t,\ \nu = 2\,c]$$

$$= 1 - \sum_{x=0}^{c-1} e^{-\lambda t}\,\frac{(\lambda t)^x}{x!}$$

L'ensemble de ces résultats peut être représenté par l'abaque ci-joint (fig. 4).

I.7 - DISTRIBUTION BETA

I.7.1 - Fonction Beta

La fonction Beta de paramètres m et n est définie par :

$$B_1\,(m,\ n) = \int_0^1 x^{m-1}\,(1 - x)^{n-1}\,dx\ ,\quad m,\ n > 0$$

Si on pose $x = \dfrac{1}{1 + y}$, $\quad dx = -\,(1 + y)^{-2}\,dy$

on peut encore écrire :

$$B_2\,(m,\ n) = \int_0^{\infty} y^{n-1}\,(1 + y)^{-m-n}\,dy$$

La fonction Beta est liée à la fonction Gamma par la relation

$$B(m,\ n) = B(n,\ m) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m + n)}$$

$$= \frac{(m - 1)!\ (n - 1)!}{(m + n - 1)!}$$

si m et n sont entiers.

I.7.2 - Distribution BETA

On dit qu'une variable aléatoire X $(0 \leqslant x \leqslant 1)$ est une variable $\beta_1(m,\ n)$ si sa loi de probabilité élémentaire est :

$$f(x)\,dx = \frac{x^{m-1}\,(1 - x)^{n-1}}{B(m,\ n)}\,dx,\qquad 0 \leqslant x \leqslant 1$$

d'où :

$$F(x) = \Pr\,(X < x) = \frac{\int_0^x x^{m-1}\,(1 - x)^{n-1}\,dx}{B(m,\ n)} = \frac{B_x\,(m,\ n)}{B(m,\ n)} = I_x\,(m,\ n)$$

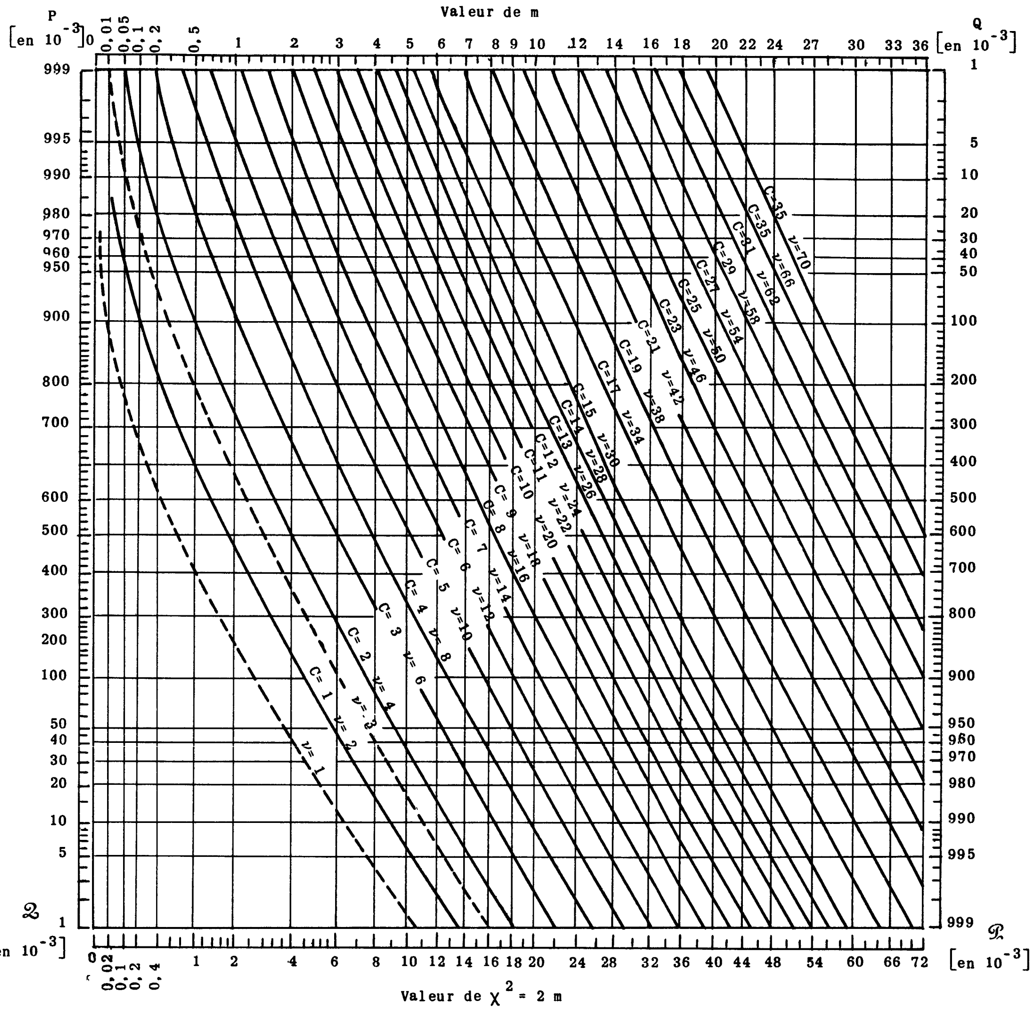


Fig. 4 - Loi de Poisson, χ^2 et Γ

Loi de Poisson : $P = Pr[\chi < c \mid m] = \sum_{x=0}^{c-1} e^{-m} \frac{m^x}{x!}$ $\quad Q = Pr.[\chi \geqslant c \mid m] = 1 - P$

Exemple : pour m = 5, $P = Pr.[\chi < 3/5] = 0,125$ $\quad Q = Pr.[\chi \geqslant 3/5] = 0,875$

Loi de χ^2 : $\mathcal{P} = Pr.[\chi^2 < 2\,m \mid \nu = 2\,c\ d.l] = Pr[\gamma(c) < m]$

et Loi γ $\quad \mathcal{Q} = Pr.[\chi^2 > 2\,m \mid \nu = 2\,c\ d.l] = Pr[\gamma(c) > m]$

Prob. $\mathcal{P}$ | Prob. $\mathcal{Q}$
0 2m ∞
$0 < \chi^2 < 2\,m$ $\quad 2\,m < \chi^2 < \infty$

1er exemple : pour $\nu = 6$ degrés de liberté, le 9ème décide de χ^2, $\chi^2_{0,9}$ tel que

$Pr.[\chi^2 < \chi^2_{0,9}] = 0,9$, est 10,6,

valeur de 2 m correspondant à $\mathcal{P} = 900.10^{-3}$

2ème exemple : pour $\nu = 6$ degrés de liberté, la probabilité que χ^2 soit supérieur à 12 est $62,10^{-3}$, valeurs de $\mathcal{Q}$ correspondant à 2 m = 12.

$B_x(m, n)$ étant la fonction B incomplète qui est tabulée.

Les principales caractéristiques de la loi Beta sont :

Mode $\dfrac{m-1}{m+n-2}$

Moyenne $E(X) = \dfrac{m}{m+n}$

Variance $V(X)$ $\dfrac{mn}{(m+n)^2\,(m+n+1)}$

Compte tenu du changement de variable défini ci-dessus, on dira qu'une variable Y ($0 \leqslant y \leqslant \infty$) est une variable $\beta_2(m, n)$ si sa loi de probabilité élémentaire est :

$$f(y)\,dy = \frac{y^{n-1}(1+y)^{-m-n}}{B(m, n)}\,dy, \qquad 0 \leqslant y \leqslant \infty$$

loi dont les principales caractéristiques sont :

Mode $\dfrac{n-1}{m+1}$

Moyenne $E(Y) = \dfrac{m}{n-1}$

Variance $V(Y)$ $\dfrac{n(m+n-1)}{(n-1)^2\,(m-2)}$

I.7.3 - Propriétés

1/ Si x et y sont des variables Gamma indépendantes de paramètres respectifs m et n, la variable $\dfrac{x}{x+y}$ est une variable β_1 de paramètres m, n.

En effet, si on considère la loi du couple (x, y)

$$f(x, y) = \frac{1}{\Gamma(m)\,\Gamma(n)}\,e^{-(x+y)}\,x^{m-1}\,y^{n-1} \qquad 0 \leqslant x,\ y < \infty$$

le changement de variables :

$$u = x + y\ ,\quad v = \frac{x}{x+y}\ ,\quad \begin{matrix} 0 \leqslant u < \infty \\ 0 \leqslant v < 1 \end{matrix}$$

donne :

$$g(u, v) = \frac{1}{\Gamma(m)\,\Gamma(n)}\,e^{-u}\,u^{m+n-1}\,v^{m-1}(1-v)^{n-1}$$

qui montre que les variables u et v sont indépendantes.

On en déduit :

$$f(u) = \int_0^1 g(u, v)\,dv = \frac{e^{-u}\,u^{m+n-1}}{\Gamma(m+n)}\ ,$$

La variable u est une variable Gamma de paramètre m + n comme on le savait déjà, et :

$$h(v) = \int_0^\infty g(u, v)\, du = \frac{v^{m-1}(1-v)^{n-1}}{B(m, n)},$$

La variable v est une variable β_1 de paramètres m, n.

2/ Si x et y sont des variables Gamma indépendantes de paramètres respectifs m, n, la variable x/y est une variable β_2 de paramètres m, n.

En effet, le changement de variables

$$u = x + y \qquad v = x/y, \qquad 0 < u, v < \infty$$

donnera :

$$g(u, v) = \frac{1}{\Gamma(m)\,\Gamma(n)}\, e^{-u}\, u^{m+n-1}\, v^{m-1}\,(1+v)^{-m-n}$$

qui montre que les deux variables u et v sont indépendantes.

On en tire :

$$h(v) = \int_0^\infty g(u, v)\, du = \frac{v^{m-1}(1+v)^{-m-n}}{B(m, n)},$$

La variable v est une variable β_2 de paramètres m, n.

Ceci s'applique immédiatement à la loi de rapport de deux variables indépendantes $\chi^2(\nu_1)$ et $\chi^2(\nu_2)$

$$w = \frac{\chi^2(\nu_1)}{\chi^2(\nu_2)} \text{ est une variable } \beta_2\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)$$

$$f(w)\, dw = \frac{w^{\frac{\nu_1}{2}-1}\,(1+w)^{-\frac{\nu_1+\nu_2}{2}}}{B(\nu_1/2, \nu_2/2)}\, dw$$

I.8 - DISTRIBUTION DE F (FISHER-SNEDECOR)

Plus connue au titre de distribution du rapport des variances estimées à partir de deux échantillons indépendants tirés d'une population normale, la loi de la variable F(1) est celle du rapport de deux variables χ^2 indépendantes respectivement divisées par leurs degrés de liberté.

$$F(\nu_1, \nu_2) = \frac{1/\nu_1\ \chi^2(\nu_1)}{1/\nu_2\ \chi^2(\nu_2)}$$

Moyennant le changement de variables

$$w = \frac{\nu_1 F}{\nu_2}, \qquad dw = \frac{\nu_1}{\nu_2}\, dF$$

----

(1) Conformément aux notations usuelles, F désigne ici la variable aléatoire F.

la loi de F se déduit immédiatement de la loi de w, soit :

$$f(F)\, dF = \frac{\frac{\nu_1}{\nu_2}\left(\frac{\nu_1 F}{\nu_2}\right)^{\frac{\nu_1}{2}-1}\left(1 + \frac{\nu_1 F}{\nu_2}\right)^{-\frac{\nu_1+\nu_2}{2}}}{B(\nu_1/2,\ \nu_2/2)}\, dF$$

$$0 \leqslant F < \infty$$

En raison du rôle dissymétrique des degrés de liberté ν_1 et ν_2, on notera, pour l'emploi des tables qui donnent divers fractiles de la loi de $F(\nu_1, \nu_2)$, que ν_1 et ν_2 désignent respectivement la colonne et la ligne à l'intersection desquelles on lira la valeur du fractile cherché.

Les principales valeurs caractéristiques de la distribution de $F(\nu_1, \nu_2)$ sont :

Mode $\dfrac{\nu_2(\nu_1 - 2)}{\nu_1(\nu_2 + 2)}$

Moyenne $E(F) = \dfrac{\nu_2}{\nu_2 - 2}$, $\nu_2 > 2$

Variance $V(F) = \left(\dfrac{\nu_2}{\nu_2 - 2}\right)^2 \times \dfrac{2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}$

De plus, on notera que :

$$F_\alpha(\nu_1, \nu_2) = \frac{1}{F_{1-\alpha}(\nu_2, \nu_1)}$$

les tables ne donnant généralement que les fractiles correspondant à des valeurs de α supérieures à 0,50.

I.8.1 - Distribution de t(Student-Fisher)

Plus connue au titre de distribution de la différence entre une moyenne hypothétique m et son estimation à partir d'un échantillon de n observations tiré d'une population normale de variance estimée à partir de l'échantillon, ou de la différence entre les moyennes m_1 et m_2 de deux populations estimées à partir de deux échantillons tirés de deux populations normales de même variance, la distribution de Student-Fisher est plus généralement celle d'une variable t telle que t^2/ν soit le quotient de deux variables χ^2 respectivement distribués avec 1 et ν degrés de liberté

$$\frac{1}{\nu}\, t^2(\nu) = \frac{\chi^2(1)}{\chi^2(\nu)}, \quad t(\nu) = \sqrt{F(1,\ \nu)}$$

La loi de t :

$$f(t)\, dt = \frac{\Gamma\left(\frac{\nu + 1}{2}\right)}{\sqrt{\nu\pi}\, \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

dépend du paramètre ν nombre de degrés de liberté de la variance (ou de la variance commune) estimée. C'est une loi symétrique dont les principales caractéristiques sont :

Mode = Médiane = Moyenne $E(t) = 0$

Variance $V(t) = \dfrac{\nu}{\nu - 2}$, $\nu > 2$

Elle tend vers la loi normale longue ν augmente indéfiniment (approximation pratiquement valable pour $\nu > 30$)

I.8.2 - Relations entre les lois binomiale, Beta et F

Considérons l'intégrale :

$$\int_p^1 y^c(1 - y)^{n-c-1}\, dy = B(c + 1, n - c) - B_p(c + 1, n - c)$$

$$0 < p < 1, \quad n \text{ et } c \text{ entiers.}$$

L'intégration par parties, conduite de proche en proche, donne :

$$\sum_{x=0}^{c} C_n^x p^x(1 - p)^{n-x} = \frac{1}{c!\,(n - c - 1)!}[B(c + 1, n - c) - B_p(c + 1, n - c)]$$

$$= \frac{1}{B(c + 1, n - c)}[B(c + 1, n - c) - B_p(c + 1, n - c)]$$

$$= 1 - I_p(c + 1, n - c)$$

La probabilité qu'une variable binomiale de paramètre p soit inférieure ou égale à c, est égale à la probabilité qu'une variable Beta de paramètres $c + 1$, $n - c$, soit supérieure à p.

D'autre part, considérons le changement de variable défini à partir de la variable F par :

$$F = \frac{\nu_2}{\nu_1}\,\frac{u}{1 - u}, \quad 0 \leqslant u < 1$$

soit :

$$u = \frac{\nu_1 F}{\nu_2 + \nu_1 F}, \qquad 1 - u = \frac{\nu_2}{\nu_2 + \nu_1 F}$$

La distribution de la variable u est définie par :

$$g(u) = \frac{u^{\frac{\nu_1}{2}-1}(1 - u)^{\frac{\nu_2}{2}-1}}{B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)}$$

La variable u est une variable β_1 de paramètres $\nu_1/2$, $\nu_2/2$.

La probabilité que la variable u soit inférieure à une valeur donnée p est :

$$\int_0^p g(u)\, du = \frac{B_p(\nu_1/2, \nu_2/2)}{B(\nu_1/2, \nu_2/2)} = I_p(\nu_1/2, \nu_2/2)$$

on a donc, pour $\nu_1/2 = c + 1$, $\nu_2/2 = n - c$:

$$\Pr[X \leqslant c] = \sum_{x=c}^{c} C_n^x p^x(1 - p)^{n-x} = 1 - \Pr(u < p)$$

$$= 1 - \Pr\left[\frac{(c + 1)\,F}{n - c + (c + 1)\,F} < p\right]$$

$$= 1 - \Pr\left[F < \frac{n - c}{c + 1}\,\frac{p}{1 - p}\right]$$

avec :

$$F\left[\nu_1 = 2(c+1), \quad \nu_2 = 2(n-c)\right]$$

Pour une valeur donnée α de $\sum_{x=0}^{c} C_n^x\, p^x\,(1-p)^{n-x}$

on aura :

$$\frac{n-c}{c+1}\,\frac{p}{1-p} = F_{1-\alpha} \quad \begin{cases} \nu_1 = 2(c+1) \\ \nu_2 = 2(n-c) \end{cases}$$

II - ETUDE DE QUELQUES MODELES MATHEMATIQUES DE DUREE DE VIE

Dans le chapitre précédent nous avons rappelé brièvement quels étaient les outils fondamentaux de la statistique théorique, leurs propriétés, leurs liaisons.

Nous allons maintenant examiner une application à l'étude de quelques modèles mathématiques de durée de vie.

Remarquons d'abord que la détermination d'un modèle mathématique, fonction du temps, capable de décrire convenablement à la fois une population et un phénomène évolutif relatif à cette population est beaucoup plus difficile a priori que celle d'un modèle pouvant décrire la situation moyenne et les variations de qualité de la production d'une machine, problème où des raisons techniques autant que statistiques conduisent à penser au rôle fondamental de la loi normale, ou de la loi binomiale et de ses approximations.

L'évolution, au cours du temps ou au cours d'une succession d'épreuves, du risque de défaillance d'un élément ou d'un ensemble ne peut évidemment pas être caractérisée de manière systématique par un simple paramètre (taux de défaillance constant).

L'expérience confirme généralement l'idée que la vie d'un organe qui vient d'être fabriqué comprend trois périodes successives qui ont quelque analogie avec celles de la vie d'un être vivant, mais aussi difficiles à délimiter.

a) période de jeunesse ou d'adaptation, généralement de courte durée, avec un taux instantané de défaillance décroissant.

b) période de vie utile ou de fonctionnement normal avec un taux sensiblement constant.

c) période d'usure avec un taux croissant.

La première intéresse surtout le producteur d'un point de vue technique : souci de ne mettre en service que des organes ayant dépassé ce stade.

La seconde, particulièrement importante en électronique (composants et ensembles), est de beaucoup celle qui se prête le mieux à l'étude statistique en raison de l'hypothèse $h(t) = Cte$; elle conduit à donner un rôle fondamental à la loi exponentielle et à la loi de Poisson.

La troisième, que l'on peut souhaiter éviter dans certains domaines, est aussi fort importante dans d'autres (résistance des matériaux, analyse des structures). Beaucoup plus difficile à étudier en raison des hypothèses diverses que l'on peut faire sur la variation du taux de défaillance, elle conduit à la recherche de modèles plus complexes (loi de Weibull, lois des valeurs extrêmes.....).

Dans ce chapitre nous étudierons seulement quelques uns de ces modèles d'un point de vue formel, mais on ne doit pas oublier que l'emploi de ces outils pour l'extension à une population des résultats fournis par un échantillon et ceci pour une durée supérieure à celle de l'expérimentation, exige beaucoup de précautions et d'études préalables.

C'est là un aspect fondamental des études de fiabilité, auxquelles les mathématiques n'apportent d'aide que si une base de départ suffisamment solide a été établie, base fondée à la fois sur l'étude technologique, la technique industrielle et l'observation statistique sans cesse poursuivie dans un domaine en évolution rapide.

II.1 - LOI EXPONENTIELLE

II.1.1 - Définition

La loi exponentielle des durées de vie correspond au cas particulier où l'on a fait l'hypothèse d'un taux instantané de défaillance constant, soit :

$$h(t) = \lambda = \text{Cte}$$

d'où :

$$\Pr(T < t) = F(t) = 1 - e^{-\lambda t}$$

(proportion des défaillants entre 0 et t pour $t \geqslant 0$, $e = 2,71828\ldots$)

$$f(t)\, dt = \lambda\, e^{-\lambda t}\, dt$$

(probabilité de défaillance entre t et t + dt)

$R(t) = e^{-\lambda t}$ = Probabilité de non défaillance entre 0 et t : (proportion des survivants à l'instant t).

On dira que la variable t est une variable $\mathcal{E}(\lambda)$, variable à distribution exponentielle de paramètre λ.

II.1.2 - La durée moyenne d'attente de l'évènement E (vie moyenne, ou temps moyen entre défaillances dans le cas de remplacement) est :

$$\theta = E(t) = \int_0^\infty t\, f(t)\, dt = 1/\lambda$$

La variance de cette durée d'attente est :

$$\sigma^2 = V(t) = 1/\lambda^2$$

II.1.3 - Etude graphique

Si les n individus d'un échantillon d'une population d'éléments considérés a priori comme identiques meurent respectivement aux instants t_1, t_2, ... t_i, ... t_n, t_i étant l'instant de la $i^{\text{ème}}$ défaillance, la probabilité $1 - F(t_i)$ de survie à l'époque t_i pourra être estimée par :

$$R_i = 1 - \frac{i}{n+1} = \frac{n+1-i}{n+1} \sim \frac{n-i}{n}, \text{ si n est grand}$$

Si la distribution est exponentielle de paramètre λ, les points (t_i, R_i) devront, dans un graphique semi-logarithmique, se placer au voisinage de la droite définie par :

$$\text{Log}_e R = -\lambda . t = -\frac{t}{\theta} \qquad \text{(fig. 5)}$$

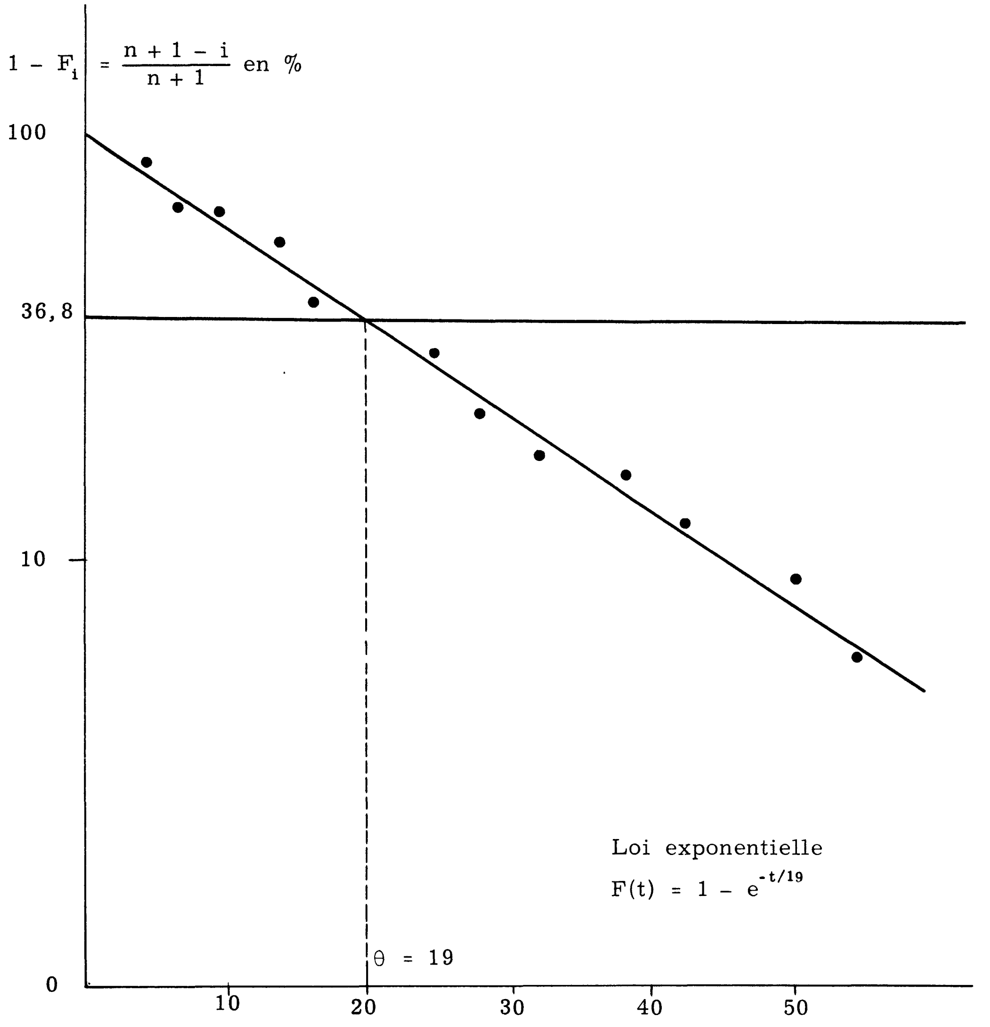


Fig. 5 - Loi exponentielle

La loi exponentielle est souvent présentée sous la forme :

$$f(t) = \frac{1}{\theta}\, e^{-t/\theta} \qquad F(t) = 1 - e^{-t/\theta}$$

mettant en évidence la vie moyenne θ.

Pour $t = \theta$ on a :

$$F(t) = 1 - \frac{1}{e} = 0,63212, \qquad 1 - F(t) = 0,36788,$$

ce qui permet une estimation graphique de la vie moyenne.

II.1.4 - Loi du nombre X de défaillants parmi n éléments indépendants au cours d'une durée T (non remplacement des défaillants).

Cette loi sera la loi binomiale :

$$f(x) = Pr(X = x) = C_n^x (1 - e^{-\lambda T})^x (e^{-\lambda T})^{n-x}$$

soit :

$$F(x) = Pr(X \leqslant x) = \sum_{x=o}^{x} C_n^x (1 - e^{-\lambda T})^x (e^{-\lambda T})^{n-x}$$

dont les valeurs caractéristiques sont :

$$E(X) = n(1 - e^{-\lambda T}) = n(1 - e^{-T/\theta})$$
$$V(X) = n(1 - e^{-\lambda T})\, e^{-\lambda T} = n(1 - e^{-T/\theta})\, e^{-T/\theta}$$

II.1.4.1 - Applications

1/ Plan d'échantillonnage pour acceptation d'un lot d'éléments appartenant à une population dont la durée de vie est distribuée exponentiellement (test sans remplacement des défaillants, de durée préfixée T ; essai tronqué).

Etant donné un lot d'éléments dont la durée de vie satisfait à une distribution exponentielle de vie moyenne θ, on se propose de déterminer un plan d'échantillonnage portant sur un échantillon de n éléments choisis au hasard dans le lot et conduisant à l'acceptation du lot si le nombre x des défaillants constatés au cours d'une durée T est inférieur ou égal à une valeur donnée c (critère d'acceptation).

La probabilité d'acceptation $P_a(\theta)$ devra satisfaire aux conditions suivantes qui définissent les risques encourus respectivement par le fournisseur et le client (fig. 6).

fig. 6

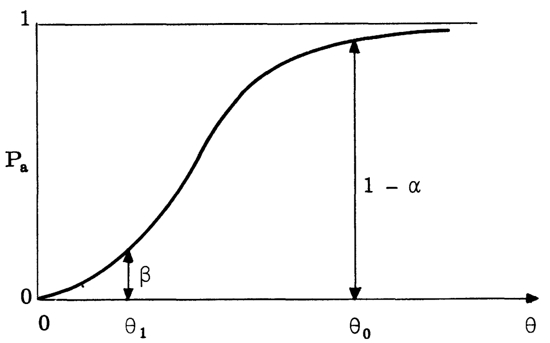


Fig. 6 - Courbe d'efficacité

$$P_a(\theta = \theta_o) = \sum_{x=o}^{c} C_n^x (1 - e^{-T/\theta_o})^x (e^{-T/\theta_o})^{n-x} = 1 - \alpha$$

$$P_a(\theta = \theta_1) = \sum_{x=c}^{c} C_n^x (1 - e^{-T/\theta_1}) (e^{-T/\theta_1})^{n-x} = \beta$$

α = risque du fournisseur : probabilité de se voir refuser un lot de qualité $\theta = \theta_o$

β : risque du client : probabilité d'être conduit à accepter un lot de qualité $\theta = \theta_1$

α, β généralement petits et $\theta_1 < \theta_0$ étant fixés, on veut déterminer un plan (n, c, T) satisfaisant aux deux conditions ci-dessus.

On remarquera que T, θ_0, θ_1, n'interviennent dans ces conditions que par les rapports T/θ_0 et T/θ_1.

Compte tenu de la relation entre la loi binomiale et la loi de F (cf. § I.8.1), ces conditions peuvent s'écrire :

$$(1)\quad F_{1-\beta} = \frac{n-c}{c+1}\,\frac{1-e^{-T/\theta_1}}{e^{-T/\theta_1}} \qquad \begin{cases} \nu_1 = 2(c+1) \\ \nu_2 = 2(n-c) \end{cases}$$

$$(2)\quad F_{\alpha} = \frac{n-c}{c+1}\,\frac{1-e^{-T/\theta_0}}{e^{-T/\theta_0}} ,$$

équations fonctionnelles en n et c, $F_{1-\beta}$ et F_α dépendant de n et de c.

n et c devant être entiers, le problème ne peut avoir qu'une solution approchée qui a été donnée par Epstein(1).

Le document américain H 108(2) donne, pour quelques valeurs des paramètres θ_1/θ_0 et T/θ_0, les valeurs de n (effectif de l'échantillon) et de $r = c + 1$ (critère de rejet). - (cf. tableau I), correspondant à la règle suivante :

Règle

n éléments étant soumis à l'épreuve, celle-ci conduira à

1/ Acceptation si au bout de la durée T, le nombre des défaillances est inférieur à r.

2/ Rejet si le r^e défaillance se produit à un instant $t_r < T$.

Remarque

Les valeurs de r et de n qui figurent dans ces tableaux sont les valeurs entières qui correspondent au mieux aux risques α et β envisagés. Pour les petites valeurs de r et de n correspondant aux plans $\theta_1/\theta_0 = 1/10$, les risques réels α' et β' définis par les relations (1) et (2) peuvent différer notablement des valeurs α, β figurant dans les tableaux.

2/ Si on se fixe seulement l'un des risques, par exemple θ_0, α, on aura :

$$F_\alpha = \frac{n-c}{c+1}\left(e^{T/\theta_0} - 1\right) \qquad \begin{cases} \nu_1 = 2(c+1) \\ \nu_2 = 2(n-c) \end{cases}$$

qui permet de déterminer l'un quelconque des paramètres n, c, T, connaissant les deux autres.

---

(1) Epstein "Truncated life tests in the exponentiel case", Annals Mathematical Statistics 1954 (Vol. 25 N° 3 pp. 555-564

(2) H 108 "Quality Control and reliability handbook" Assistant Secretary of defense-Washington 25 D.C.

Tableau 1

Valeurs de r pour α, β, θ_1/θ_0 et T/θ_0 donnés

θ_1/θ_0	r	T/θ_0				r	T/θ_0			
		1/3	1/5	1/10	1/20		1/3	1/5	1/10	1/20
		n	n	n	n		n	n	n	n
		α = 0,01		β = 0,01			α = 0,01		β = 0,05	
2/3	136	403	622	1172	2275	101	291	448	842	1632
1/2	46	119	182	340	657	35	87	132	245	472
1/3	19	41	61	113	216	15	30	45	82	157
1/5	9	15	22	39	74	8	13	18	33	62
1/10	5	6	9	15	28	4	4	6	10	18
		α = 0,01		β = 0,10			α = 0,05		β = 0,01	
2/3	83	234	359	675	1307	95	289	447	843	1639
1/2	30	72	109	202	390	33	90	138	258	499
1/3	13	25	37	67	128	13	30	45	83	160
1/5	7	11	15	26	50	7	13	20	36	69
1/10	4	4	6	10	18	4	6	9	15	29
		α = 0,05		β = 0,05			α = 0,05		β = 0,10	
2/3	67	198	305	575	1116	55	159	245	462	895
1/2	23	59	90	168	326	19	47	72	134	258
1/3	10	21	32	59	113	8	16	24	43	83
1/5	5	8	12	22	41	4	6	9	15	29
1/10	3	4	5	9	17	3	4	5	9	17
		α = 0,10		β = 0,01			α = 0,10		β = 0,05	
2/3	77	238	369	699	1358	52	156	242	456	886
1/2	26	73	112	210	407	18	48	73	137	265
1/3	11	27	40	75	145	8	18	27	50	97
1/5	5	10	14	26	51	4	7	10	19	36
1/10	3	5	7	12	23	2	2	3	6	11
		α = 0,10		β = 0,10			α = 0,25		β = 0,25	
2/3	41	121	186	351	681	12	34	53	101	196
1/2	15	39	59	110	213	5	12	19	36	69
1/3	6	12	18	34	66	2	3	5	10	20
1/5	3	5	7	12	23	1	1	1	3	6
1/10	2	2	3	6	11	1	1	1	3	6

Exemple : $\alpha = 0{,}10, \quad \theta_0 = 1000, \quad c = 4, \quad n = 10$

$$F_{0,10} = \frac{6}{5}\,(e^{T/1000} - 1), \qquad \nu_1 = 10, \qquad \nu_2 = 12$$

soit :

$$\frac{6}{5}\,(e^{T/1000} - 1) = \frac{1}{F_{0,90}}, \quad \nu_1 = 12, \qquad \nu_2 = 10$$

$$e^{T/1000} = 1 + \frac{5}{6 \times 2{,}29}$$

$T = 314$

On acceptera le lot si au cours d'un essai de 314 heures, on a observé au plus 4 défaillances sur 10 éléments soumis à l'épreuve.

La brochure H. 108 donne les valeurs de T/θ_0 en fonction de $r = c + 1 = 1, 2, \ldots\ldots, 100$ et de $n = k\,r$ ($k = 2, \ldots\ldots 10$ et 20) pour $\alpha = 0{,}01 - 0{,}05 - 0{,}10 - 0{,}25 - 0{,}50$.

On pourra ensuite calculer l'un des éléments β, θ_1 à l'aide de

$$F_{1-\beta} = \frac{n - c}{c + 1}\,(e^{T/\theta_1} - 1) \qquad \begin{cases} \nu_1 = 2(c + 1) \\ \nu_2 = 2(n - c) \end{cases}$$

Dans l'exemple précédent, pour $\theta_1 = 200$ on a :

$$F_{1-\beta} = \frac{6}{5}\,(e^{1{,}57} - 1) = 4{,}57$$

d'où :

$$\beta \approx 0{,}01$$

II.1.5 - Estimation de la vie moyenne d'après le résultat global d'une épreuve de durée T (sans remplacement)

Connaissant seulement le nombre k des défaillants observés au cours de la durée T, on pourra déterminer un intervalle de confiance centré (θ_i, θ_s) à $1 - \alpha$ pour θ, à partir des conditions suivantes :

(1) $$\frac{\alpha}{2} = \Pr[\,x \leqslant k \,|\, \theta = \theta_i)] = \sum_{x=c}^{k} C_n^x\,(1 - e^{-T/\theta_i})^x\,(e^{-T/\theta_i})^{n-x}$$

(2) $$\frac{\alpha}{2} = \Pr[\,x \geqslant k \,|\, \theta = \theta_s)] = 1 - \sum_{x=0}^{k-1} C_n^x\,(1 - e^{-T/\theta_s})^x\,(e^{-T/\theta_s})^{n-x}$$

Compte tenu des relations entre la loi binomiale et la loi de F (§ I.8.1), la relation (1) peut s'écrire :

$$1 - \frac{\alpha}{2} = \Pr\left[F < \frac{n - k}{k + 1}\,\frac{1 - e^{-T/\theta_i}}{e^{-T/\theta_i}}\right], \quad \begin{cases} \nu_1 = 2\,k + 2 \\ \nu_2 = 2\,n - 2\,k \end{cases}$$

d'où :

$$\frac{n - k}{k + 1}\,\frac{1 - e^{-T/\theta_i}}{e^{-T/\theta_i}} = F_{1-\alpha/2}$$

et :

$$\theta_i = T\left[\operatorname{Log}_e\left(1 + \frac{k + 1}{n - k}\,F_{1-\alpha/2}\right)\right]^{-1}$$

$$\nu_1 = 2k + 2, \qquad \nu_2 = 2n - 2k$$

De même la relation (2) permettrait de calculer une limite supérieure :

$$\theta_s = T\left[\mathrm{Log}_e\left(1 + \frac{k}{n-k+1} F_{a/2}\right)\right]^{-1}$$

$$\nu_1 = 2k, \qquad \nu_2 = 2n - 2k + 2$$

II.1.6 - Distribution de la plus petite durée de vie de k éléments obéissant à des lois exponentielles indépendantes

Soient les k lois définies par :

$$\Pr(T_i < t) = 1 - e^{-\lambda_i t} \qquad i = 1, 2, \ldots\ldots k$$

et Z la variable aléatoire correspondant au minimum des variables T_i, on a :

$$\Pr(Z < t) = 1 - \Pr(Z > t) = 1 - \Pr(T_1 > t)\Pr(T_2 > t) \ldots\ldots \Pr(T_k > t)$$

$$= 1 - e^{-\lambda_1 t} e^{-\lambda_2 t} \ldots\ldots e^{-\lambda_k t} = 1 - e^{-(\lambda_1 + \ldots\ldots \lambda_k)t}$$

La variable Z est donc aussi une variable exponentielle de paramètre $\lambda = \lambda_1 + \lambda_2 + \ldots\ldots \lambda_k$.

Dans le cas particulier où les k éléments appartiennent à la même population de paramètre λ, l'espérance mathématique de Z sera :

$$E(Z) = 1/k\lambda = \frac{\theta}{k}$$

et sa variance :

$$V(Z) = 1/k^2\lambda^2 = \frac{\theta^2}{k^2}$$

II.1.7 - Estimation de la vie moyenne θ d'éléments obéissant à une même loi exponentielle

1/ Essai arrêté à la $k^{\text{ème}}$ défaillance, sans remplacement des défaillaillants (Essai censuré)

Considérons un échantillon de n éléments mis à l'épreuve et dont on a observé les durées de vie ordonnées des k premiers défaillants

$$t_1 < t_2 < \ldots\ldots\ldots < t_k, \quad 1 \leqslant k \leqslant n$$

Quelle est l'estimation optimale $\hat{\theta}$ de la vie moyenne θ ? Quelle confiance peut-on accorder à cette estimation ? (les éléments défaillants ne sont pas remplacés).

La réponse à cette double question - réponse qui va être justifiée ci-après - est la suivante[1]

----

(1) Epstein and Sobel "Life Testing" J.A.S.A. 1953 Vol 48 pp. 486-502

L'estimation de la vie moyenne θ est :

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n-k)\, t_k\right] = \frac{T}{k},$$

T étant la somme des durées de vie observées.

On a :

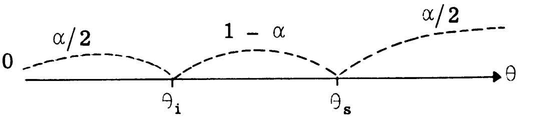
$$E(\hat{\theta}_k) = \theta \qquad V(\hat{\theta}_k) = \theta^2/k.$$

A cette estimation on pourra associer un intervalle de confiance

$$\left[\theta_i = \frac{2\,k\,\hat{\theta}}{\chi^2_{1-\alpha/2}}, \qquad \theta_s = \frac{2\,k\,\hat{\theta}}{\chi^2_{\alpha/2}}, \qquad \nu = 2\,k\right] \text{ tel que :}$$

$$1 - \alpha = \Pr\,[\theta_i < \hat{\theta} < \theta_s]$$

c'est-à-dire que pour une valeur de α, choisie à l'avance, il y a une probabilité $1 - \alpha$ que l'intervalle (θ_i, θ_s) contienne la vraie valeur inconnue θ.



Démonstration

1/ La loi de la $i^{ème}$ valeur observée est définie par la densité de probabilité(1)

$$f(t_i) = \frac{n!}{(i-1)!\;1!\;(n-i)!}\,[1 - e^{-t_i/\theta}]^{i-1}\,\frac{1}{\theta}\,(e^{-t_i/\theta})^{n-i}$$

(Probabilité pour que parmi les n éléments considérés, $i - 1$ quelconques d'entre eux soient défaillants avant l'instant t_i, l'un d'eux défaille entre t_i et dt_i, les $n - i$ autres continuant à fonctionner au-delà de $t = t_i$.

2/ La densité de probabilité de l'ensemble $(t_1, t_2, \ldots\ldots, t_k)$, c'est-à-dire la probabilité d'avoir k défaillances parmi les n éléments, aux instants $t_1, t_2, \ldots\ldots, t_k$, et $n - k$ éléments non défaillants à l'instant t_k est(1) :

$$g(t_1, t_2, \ldots\ldots, t_k) = \frac{n!}{(n-k)!}\,\frac{1}{\theta^k}\,e^{-\frac{1}{\theta}\left[\sum_{i=1}^{k} t_i + (n-k)\,t_k\right]}$$

----

(1) $\frac{n!}{(i-1)!\;1!\;(n-i)!}$ = nombre des dispositions ordonnées (permutations avec répétition) de n éléments appartenant respectivement à trois groupes d'effectifs $i - 1$, 1 et $n - i$.

$\frac{n!}{(n-k)!}$ = nombre des dispositions ordonnées (arrangements) de k éléments distincts pris parmi les n éléments.

3/ L'estimation de θ par le maximum de vraisemblance, c'est-à-dire celle qui rend maximale la probabilité correspondante à l'ensemble observé est définie par :

$$\frac{\partial g}{\partial \theta} = 0$$

d'où :

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n - k)\, t_k\right]$$

4/ Pour étudier la loi de la variable aléatoire $\hat{\theta}_k$, posons :

$$y_1 = t_1, \quad y_2 = t_2 - t_1, \ \ldots\ldots\ldots, \ y_k = t_k - k_{k-1}$$

soit :

$$t_1 = y_1, \quad t_2 = y_1 + y_2 \ \ldots\ldots\ldots \ t_k = y_1 + y_2 + \ldots\ldots y_k$$

d'où :

$$\hat{\theta}_k = \frac{1}{k}\sum_{i=1}^{k} (n - i + 1)\, y_i$$

Moyennant ce changement de variable, la densité de probabilité de l'ensemble $(y_1, y_2, \ldots\ldots y_k)$ est :

$$g(y_1, y_2, \ldots y_k) = \frac{n!}{(n-k)!}\,\frac{1}{\theta^k}\, e^{-\frac{1}{\theta}\sum_{i=1}^{k}(n-i+1)y_i}$$

qui peut s'écrire :

$$g(y_1, y_2, \ldots\ldots y_k) = \prod_{i=1}^{k}\left(\frac{n - i + 1}{\theta}\right) e^{-\frac{1}{\theta}\sum_{i=1}^{k}(n-i+1)y_i}$$

qui montre que les variables y_i sont indépendantes, la loi de l'une quelconque d'entre elles étant la loi exponentielle :

$$f(y_i) = \frac{n - i + 1}{\theta}\, e^{-\frac{n-i+1}{\theta}y_i},$$

d'où la valeur moyenne et la variance de y_i et de t_k :

$$E(y_i) = \frac{\theta}{n - i + 1}$$

$$E(t_k) = E\sum_{i=1}^{k} y_i = \theta\sum_{i=1}^{k}\frac{1}{n - i + 1}$$

$$V(y_i) = \frac{\theta^2}{(n - i + 1)^2}$$

$$V(t_k) = \theta^2\sum_{i=1}^{k}\frac{1}{(n - i + 1)^2}$$

Il en résulte encore que l'on a :

$$E(\hat{\theta}_k) = \frac{1}{k} E\left[\sum_{i=1}^{k} (n - i + 1)\, y_i\right] = \frac{1}{k} k\, \theta = \theta$$

$$V(\hat{\theta}_k) = \frac{1}{k^2} \sum (n - i + 1)^2\, V(y_i) = \frac{1}{k^2} k\, \theta^2 = \frac{\theta^2}{k}$$

Si on pose :

$$z_i = \frac{n - i + 1}{\theta} y_i$$

la loi de z_i est la loi Gamma de paramètre 1

$$f(z_i)\, dz_i = e^{-z_i}\, dz_i$$

Il résulte de la propriété d'additivité de la loi Gamma (§ I.5.2.1) que la loi de la variable $w = \sum_{i=1}^{k} \frac{n - i + 1}{\theta} y_i = k \frac{\hat{\theta}_k}{\theta}$ est la loi Gamma de paramètre k :

$$f(w)\, dw = \frac{1}{\Gamma(k)} e^{-w}\, w^{k-1}\, dw$$

d'où la loi de $\hat{\theta}_k = \theta \frac{w}{k}$

$$f(\hat{\theta}_k)\, d\hat{\theta}_k = \frac{1}{\Gamma(k)} e^{-k\frac{\hat{\theta}_k}{k}} \left(\frac{k\, \hat{\theta}_k}{\theta}\right)^{k-1} \frac{k}{\theta}\, d\hat{\theta}_k$$

D'autre part, de la relation entre la loi Gamma et la loi de χ^2 (§ I.6) il résulte que la variable $2\, k \frac{\hat{\theta}_k}{\theta}$ est une variable χ^2 avec $\nu = 2k$ degrés de liberté.

Pour un risque α donné (niveau de confiance $1 - \alpha$) les limites inférieure et supérieure de l'intervalle de confiance centré à $1 - \alpha$ pour θ seront définies par :

$$\frac{\alpha}{2} = \Pr\left[\chi^2 > \frac{2\, k\, \hat{\theta}_k}{\theta_i}\right] \qquad \nu = 2\, k$$

$$\frac{\alpha}{2} = \Pr\left[\chi^2 < \frac{2\, k\, \theta_k}{\theta_s}\right]$$

d'où l'intervalle de confiance centré (θ_i, θ_s) défini par :

$$\theta_i = \frac{2\, k\, \hat{\theta}_k}{\chi^2_{1-\alpha/2}} = \frac{2\, T}{\chi^2_{1-\alpha/2}}, \quad \theta_s = \frac{2\, k\, \hat{\theta}_k}{\chi^2_{\alpha/2}} \quad \frac{2\, T}{\chi^2_{\alpha/2}}, \quad \nu = 2\, k,$$

T étant la somme des durées de vie observées.

Le choix entre les différents estimateurs possibles $\hat{\theta}_k$, c'est-à-dire le choix de k = 1, 2, ... n, dépend du degré de confiance que l'on désire accorder au résultat et du coût de l'opération (coût unitaire des éléments et durée d'attente du résultat caractérisé par sa valeur moyenne :

$$E(t_k) = \theta \sum_{i=1}^{k} \frac{1}{n - i + 1}, \quad (k = 1, 2, \ldots\ldots, n)$$

2/ Essai avec remplacement des défaillants

Pour chaque groupe constitué par l'un des n éléments placés au début de l'essai et les remplaçants éventuels qui lui sont associés en cas de défaillance avant l'instant t_k de la $k^{ième}$ défaillance de l'ensemble complet, la densité de probabilité de ce groupe est :

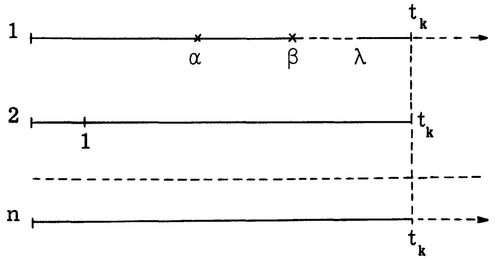
$$\frac{1}{\theta} e^{-t_\alpha/\theta} \times \frac{1}{\theta} e^{\frac{-(t_\beta - t_\alpha)}{\theta}}, \ldots\ldots\ldots e^{\frac{-(t_k - t_\lambda)}{\theta}} = \frac{1}{\theta} e^{-t_k/\theta}$$

α, β, ... λ étant les rangs des u défaillances constatées dans ce groupe, ces rangs étant ordonnés dans l'ensemble complet des k défaillances.

$$1 < \ldots\ldots < \alpha < \ldots\ldots < \beta < \ldots\ldots < \lambda < \ldots\ldots < k$$

La densité de probabilité de l'ensemble (t_1, t_2, ... t_k) est :

$$g(t_1, t_2, \ldots\ldots t_k) = n^k \frac{1}{\theta^k} e^{-nt_k/\theta} \qquad (t_1 < t_2 < \ldots\ldots < t_k)$$



n^k étant le nombre des dispositions ordonnées de k éléments parmi les n, chacun des n éléments pouvant figurer jusqu'à k fois dans une même disposition.

L'estimation de θ par le maximum de vraisemblance est alors

$$\hat{\theta}_k = \frac{n\, t_k}{k} = \frac{T}{k}$$

T étant encore la somme des durées de vie observées.

Si l'on pose :

$$y_1 = t_1, \; y_2 = t_2 - t_1, \; \ldots\ldots\; y_k = t_k - t_{k-1}, \; t_k = \sum_{i=1}^{k} y_i$$

d'où :

$$\hat{\theta}_k = \frac{n}{k} \sum_{i=1}^{k} y_i$$

la densité de probabilité de l'ensemble (y_1, ... y_k) est :

$$g(y_1, y_2 \ldots\ldots y_k) = \left(\frac{n}{\theta}\right)^k e^{-\frac{n}{\theta}\sum_{i=1}^{k} y_i}$$

$$= \prod_{i=1}^{k} \frac{n}{\theta} e^{-\frac{n}{\theta}y_i}$$

Les variables y_i sont des variables indépendantes de loi élémentaire

$$f(y_i)\, dy_i = \frac{n}{\theta} e^{-\frac{n}{\theta}y_i} dy_i$$

d'où :

$$E(y_i) = \frac{\theta}{n} \qquad V(y_i) = \frac{\theta^2}{n^2}$$

$$E(t_k) = k\frac{\theta}{n} \qquad V(t_k) = k\frac{\theta^2}{n^2}$$

$$E(\hat{\theta}_k) = \hat{\theta} \qquad V(\hat{\theta}_k) = \frac{\theta^2}{k}$$

Si l'on pose :

$$z_i = \frac{n}{\theta} y_i$$

la loi de z_i est la loi Gamma de paramètre 1 :

$$f(z_i)\, dz_i = e^{-z_i}\, dz_i$$

La loi de la variable $W = \sum_{i=1}^{k} z_i$ est la loi Gamma de paramètre k.

Il en résulte que la loi de $2\,W = 2\,k\frac{\hat{\theta}_k}{\theta}$ est encore une loi de χ^2 avec $\nu = 2\,k$ degrés de liberté.

La loi de la variable $\hat{\theta}_k$ est alors :

$$f(\hat{\theta}_k)\, d\hat{\theta}_k = \frac{1}{(k-1)!}\left(\frac{k}{\theta}\right)^k \hat{\theta}_k^{k-1} e^{-k\hat{\theta}_k/\theta} d\hat{\theta}_k$$

Pour un risque α donné (niveau de confiance $1-\alpha$), les limites de l'intervalle de confiance à $1-\alpha$ pour θ sont définies par :

$$\frac{\alpha}{2} = \Pr\left[\chi^2 > 2\,k\frac{\hat{\theta}_k}{\theta_i}\right]$$

$$\frac{\alpha}{2} = \Pr\left[\chi^2 < 2\,k\frac{\hat{\theta}_k}{\theta_s}\right] \qquad \nu = 2\,k$$

d'où l'intervalle de confiance (θ_i, θ_s), centré à $1-\alpha$:

$$\left(\theta_i = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{1-\alpha/2}}, \quad \theta_s = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{\alpha/2}}\right), \qquad \nu = 2\,k$$

Nota

Les essais précédents arrêtés à la $k^{\text{ième}}$ défaillance (k fixé, t_k aléatoire) sont appelés essais "censurés", par opposition à ceux qui sont arrêtés à un instant t préfixé (k aléatoire) appelés essais "tronqués".

II.1.8 - Plan d'échantillonnage pour acceptation d'un lot d'éléments dont la loi de durée de vie est exponentielle : essai basé sur les k premières défaillances observées : (essai censuré)

L'information fournie par l'essai permet d'estimer la vie moyenne (§ II.1.6) :

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n - k)\, t_k\right] \quad \text{(sans remplacement)}$$

$$\hat{\theta}_k = \frac{n\, t_k}{k} \quad \text{(avec remplacement des défaillants)}$$

La décision d'acceptation ou de rejet devra donc être basée sur $\hat{\theta}_k$ et k, compte tenu des conditions fixées a priori pour l'efficacité du plan $(\theta_1\ \beta\ \theta_0\ \alpha)$.

Si l'on envisage une condition de la forme $\hat{\theta}_k > C$ pour l'acceptation on devra d'abord avoir :

$$\Pr[\hat{\theta}_k > C/\theta = \theta_0] = 1 - \alpha$$

soit :

$$\Pr\left[\frac{2\,k\,\hat{\theta}_k}{\theta_0} > \frac{2\,k\,C}{\theta_0}\right] = 1 - \alpha$$

La variable $\frac{2\,k\,\hat{\theta}_k}{\theta_0}$ étant une variable $\chi^2(\nu = 2\,k)$, cette condition s'écrit (§ I.6.1) :

$$\Pr\left[\chi^2(\nu = 2\,k) > \frac{2\,k\,C}{\theta_0}\right] = 1 - \alpha$$

d'où :

$$C = \theta_0 \frac{\chi^2_\alpha}{2\,k}, \qquad \nu = 2\,k$$

Pour une valeur quelconque de θ, la probabilité d'acceptation sera

$$\Pr\left[2\,k \frac{\hat{\theta}_k}{\theta} > \frac{\theta_0}{\theta}\chi^2_\alpha\right],$$

soit :

$$\Pr\left[\chi^2 > \frac{\theta_0}{\theta}\chi^2_\alpha\right]$$

En particulier pour $\theta = \theta_1$, on devra avoir :

$$\Pr\left[\chi^2 > \frac{\theta_0}{\theta_1}\chi^2_\alpha\right] = \beta$$

Tableau 2

Essais censurés

Valeurs de k et C/θ_0 en fonction de θ_1/θ_0, α et β

θ_1/θ_0	$\alpha = 0,01$ $\beta = 0,01$		$\alpha = 0,01$ $\beta = 0,05$		$\alpha = 0,01$ $\beta = 0,10$		$\alpha = 0,01$ $\beta = 0,25$	
	k	C/θ_0	k	C/θ_0	k	C/θ_0	k	C/θ_0
2/3	136	0,811	101	0,783	83	0,762	60	0,724
1/2	46	0,689	35	0,649	30	0,625	22	0,572
1/3	19	0,544	15	0,498	13	0,469	10	0,413
1/5	9	0,390	8	0,363	7	0,333	5	0,256
1/10	5	0,256	4	0,206	4	0,206	3	0,145

θ_1/θ_0	$\alpha = 0,05$ $\beta = 0,01$		$\alpha = 0,05$ $\beta = 0,05$		$\alpha = 0,05$ $\beta = 0,10$		$\alpha = 0,05$ $\beta = 0,25$	
	k	C/θ_0	k	C/θ_0	k	C/θ_0	k	C/θ_0
2/3	95	0,837	67	0,808	55	0,789	35	0,739
1/2	33	0,732	23	0,683	19	0,655	13	0,592
1/3	13	0,592	10	0,543	8	0,498	6	0,436
1/5	7	0,469	5	0,394	4	0,342	3	0,272
1/10	4	0,342	3	0,272	3	0,272	2	0,178

θ_1/θ_0	$\alpha = 0,10$ $\beta = 0,01$		$\alpha = 0,10$ $\beta = 0,05$		$\alpha = 0,10$ $\beta = 0,10$		$\alpha = 0,10$ $\beta = 0,25$	
	k	C/θ_0	k	C/θ_0	k	C/θ_0	k	C/θ_0
2/3	77	0,857	52	0,827	41	0,806	25	0,754
1/2	26	0,758	18	0,712	15	0,687	9	0,604
1/3	11	0,638	8	0,582	6	0,525	4	0,436
1/5	5	0,487	4	0,436	3	0,367	3	0,367
1/10	3	0,367	2	0,266	2	0,266	2	0,266

θ_1/θ_0	$\alpha = 0,25$ $\beta = 0,01$		$\alpha = 0,25$ $\beta = 0,05$		$\alpha = 0,25$ $\beta = 0,10$		$\alpha = 0,25$ $\beta = 0,25$	
	k	C/θ_0	k	C/θ_0	k	C/θ_0	k	C/θ_0
2/3	52	0,903	32	0,876	23	0,853	12	0,793
1/2	17	0,827	11	0,784	8	0,744	5	0,674
1/3	7	0,726	5	0,674	4	0,634	2	0,481
1/5	3	0,576	2	0,481	2	0,481	1	0,288
1/10	2	0,481	2	0,481	1	0,288	1	0,288

d'où :

$$\frac{\theta_o}{\theta_1}\ \chi^2_\alpha = \chi^2_{1-\beta}\ , \qquad \nu = 2k$$

k est donc théoriquement défini par :

$$\frac{\theta_o}{\theta_1} = \frac{\chi^2_{1-\beta}(\nu = 2\ k)}{\chi^2_\alpha\ (\nu = 2\ k)}$$

En fait, k devant être entier, on prendra pour k le plus petit entier vérifiant la condition :

$$\frac{\theta_o}{\theta_1} \geqslant \frac{\chi^2_{1-\beta}}{\chi^2_\alpha}$$

le test ainsi défini satisfera alors aux conditions d'acceptation :

$$P_a(\theta = \theta_o) = 1 - \alpha \qquad P_a(\theta = \theta_1) \leqslant \beta$$

Pour diverses valeurs de α, β et θ_o/θ_1, Epstein et Sobel[1] ont donné un tableau de k et C/θ_o, définissant la règle d'acceptation (tableau 2)

$$\hat{\theta}_k \geqslant C$$

Remarque

Une théorie plus complète de ce test est basée sur la théorie générale des tests d'hypothèses de Neyman et Pearson (tests basés sur le rapport des vraisemblances de l'ensemble observé dans les hypothèses $\theta = \theta_o$ et $\theta = \theta_1$.

II.1.9 - Estimation de la vie moyenne (loi exponentielle), test arrêté au bout d'un temps t_o

1/ Essai avec remplacement des défaillants

Si n éléments à l'origine de l'essai, la somme des durées de vie est :

$$T = n t_o \quad \text{d'où} \quad \hat{\theta} = \frac{n t_o}{k},$$

si k défaillances entre 0 et t_o

L'intervalle de confiance à $1 - \alpha$ pour θ est défini par :

$$\frac{\alpha}{2} = Pr[x \leqslant k | \theta = \theta_i]$$

$$\frac{\alpha}{2} = Pr[x \geqslant k | \theta = \theta_s]$$

θ_i θ_s

----

(1) Epstein and Sobel "Life testing" Journal of the American Statistical Association, Vol 48 n° 263, Sept. 1953 pp. 486-502

$$1 - \frac{\alpha}{2} = \Pr[x \leqslant k - 1 | \theta = \theta_s]$$

or on a dans ce cas (remplacement) cf. II.2.6.1

$$\Pr[x \leqslant k/\theta] = \sum_{x=o}^{k} e^{-nt_o/\theta} \frac{(nt_o/\theta)^x}{x!}$$

$$= \Pr\left[\chi^2 > \frac{2\,n\,t_o}{\theta},\quad \nu = 2(k+1)\right]$$

d'où les deux conditions :

$$\frac{\alpha}{2} = \Pr\left[\chi^2 > \frac{2\,n\,t_o}{\theta_i},\quad \nu = 2(k+1)\right]$$

$$1 - \frac{\alpha}{2} = \Pr\left[\chi^2 > \frac{2\,n\,t_o}{\theta_s},\quad \nu = 2\,k\right]$$

d'où :

$$\theta_i = \frac{2\,n\,t_o}{\chi^2_{1-\alpha/2}},\qquad \theta_s = \frac{2\,n\,t_o}{\chi^2_{\alpha/2}}$$

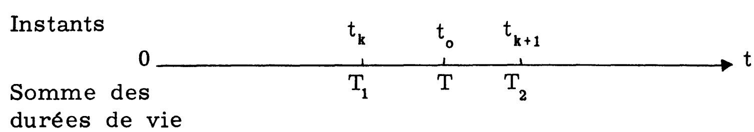
$$(\nu = 2\,k + 2)\qquad (\nu = 2\,k)$$

2/ Essai sans remplacement des défaillants

Dans le cas précédent toute l'information sur la vie moyenne était résumée par la connaissance de n, t_o et k.

Il n'en est pas de même dans le cas présent où l'on veut tenir compte des durées de vie individuelle des k éléments défaillants, qui sont précisément les instants t_1, t_2, ... t_k des défaillants(1).

Une solution approchée du problème peut être présentée comme suit:



La somme des durées de vie, observée à l'instant t_o est :

$$T = \sum_{i=o}^{k} t_i + (n - k)\, t_o$$

d'où $\hat{\theta}_k = T/k$.

Si l'essai avait été arrêté à l'instant t_k de la $k^{\text{ième}}$ défaillance, on aurait eu l'intervalle à $1 - \alpha$:

----

(1) Dans l'essai sans remplacement étudié au § II.1.5 on ne tenait compte que du nombre des défaillants et non de leurs instants d'arrivée.

$$\theta_{i1} = \frac{2\ T_1}{\chi^{2\ (2k)}_{1-\alpha/2}} < \theta < \frac{2\ T_1}{\chi^{2\ (2k)}_{\alpha/2}} = \theta_{s1} \qquad \text{(voir II.1.7)}$$

avec :

$$T_1 = \sum_{o}^{k} t_i + (n - k)\ t_k$$

De même si l'essai avait été poursuivi jusqu'à l'instant t_{k+1} de la $(k + i)^{\text{ème}}$ défaillance, on aurait eu l'intervalle

$$\theta_{i2} = \frac{2\ T_2}{\chi^{2\ (2k+2)}_{1-\alpha/2}} < \theta < \frac{2\ T_2}{\chi^{2\ (2k+2)}_{\alpha/2}} = \theta_{s2}$$

avec :

$$T_2 = \sum_{i=1}^{k+1} t_i + (n - k - 1)\ t_{k+1}$$

Pour l'essai effectif, arrêté à l'instant t_o, k défaillances étant observées avant t_o, on pourra prendre comme estimation approchée de l'intervalle de confiance à $1 - \alpha$

$$\theta_i = \frac{2\ T}{\chi^{2\ (2k+2)}_{1-\alpha/2}}, \qquad \theta_s = \frac{2\ T}{\chi^{2\ (2k)}_{\alpha/2}}$$

Compte tenu de $T_1 < T < T_2$, on a

$$\theta_i < \theta_{12}, \qquad \theta_s > \theta_{s1}$$

II.1.10 - Estimation de la durée de vie x_p qui a une probabilité p d'être dépassée. (Essais censurés et tronqués)

Si θ est la vie moyenne, on veut estimer une durée x_p telle que :

$$\Pr[X > x_p] = p = e^{-x_p/\theta}$$

d'où :

$$x_p = \theta\ \mathrm{Log}_e\, 1/p$$

qui sera estimée par :

$$\hat{x}_p = \hat{\theta}\ \mathrm{Log}_e 1/p$$

L'intervalle de confiance centré à $1 - \alpha$ pour x_p sera défini à partir des intervalles correspondants pour θ, suivant qu'il s'agit d'un essai censuré ou d'un essai tronqué, les limites de ces intervalles étant multipliées par $\mathrm{Log}_e 1/p$.

Inversement la probabilité p_t qu'un élément de la population testée ait une durée de vie supérieure à une durée donnée t est :

$$p_t = e^{-t/\theta}$$

Elle sera estimée par :

$$\hat{p}_t = e^{-t/\hat{\theta}}$$

avec un intervalle de confiance à $1 - \alpha$ défini à partir des intervalles de confiance calculés pour $\hat{\theta}$, selon qu'il s'agit d'un essai tronqué ou d'un essai censuré.

II.2 - PROCESSUS DE POISSON

II.2.1 - Définition

On dit q'une suite d'évènements E_1, E_2 ..... se réalisant aux temps t_1, t_2, ... obéit à un processus de Poisson, si les intervalles de temps $t_2 - t_1$, $t_3 - t_2$ ..... suivent indépendamment une même loi de probabilité et si la probabilité d'arrivée de l'un quelconque des évènements E_i pendant un intervalle de temps infiniment petit dt est λ dt (λ constante positive indépendante de t).

Soit $P_x(0, t)$ la probabilité pour que x évènements E aient lieu au cours de l'intervalle (0, t), ou plus généralement entre T et T + t on a :

$$P_x(0, t + dt) = P_x(0, t)\ P_o(t, t + dt) + P_{x-1}(0, t)\ P_1(t, t + dt)$$

$$= P_x(0, t)\ (1 - \lambda dt) + P_{x-1}(0, t)\ \lambda dt$$

(dt infiniment petit, tel que $P_{k>1}(t, t + dt)$ est négligeable).

On en tire :

$$\frac{P_x(0, t + dt) - P_x(0, t)}{dt} = -\lambda P_x(0, t) + \lambda P_{x-1}(0, t)$$

soit :

$$\frac{dF_x(0, t)}{dt} + \lambda P_x(0, t) = \lambda P_{x-1}(0, t),$$

système récurrent d'équations différentielles dont la solution, avec les conditions initiales :

$$P_o(0, 0) = 1$$

$$P_k(0, 0) = 0 \qquad \text{pour } k \geqslant 1$$

est :

$$P_x(0, t) = e^{-\lambda t}\ \frac{(\lambda t)^x}{x!}$$

En particulier, on a :

$$P_o(0, t) = e^{-\lambda t}$$

Le nombre d'arrivées d'évènements E dans l'intervalle (0, t) fixé, est une variable aléatoire X telle que :

$$Pr[X \leqslant x \mid (0, t)] = \sum_{x=0}^{x} e^{-\lambda t}\ \frac{(\lambda t)^x}{x!}$$

La variable X est donc distribuée suivant une loi de Poisson de paramètre λt.

$$E(X) = \lambda t$$

$$V(X) = \lambda t$$

Remarques

1/ On peut encore raisonner comme suit : quels que soient t' et t'', tout intervalle de temps fini $t'' - t'$ peut être partagé en n intervalles élémentaires infiniment petits et égaux à dt, soit :

$$dt = \frac{t'' - t'}{n}$$

La probabilité pour que x de ces n intervalles voient la réalisation d'un évènement E_i est donnée par la loi binomiale :

$$P_x(t', t'') = C_n^x(\lambda dt)^x (1 - \lambda dt)^{n-x},$$

dont la forme limite est, pour $dt \longrightarrow 0$:

$$P_x(t', t'') = e^{-n\lambda dt} \frac{(n \lambda dt)^x}{x!}$$

loi de Poisson de moyenne $n\lambda dt$, soit, pour $t' = 0$, $t'' = t$

$$P_x(0, t) = e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

2/ L'hypothèse faite sur les évènements E était que l'instant d'arrivée d'un évènement E_i était indépendante de l'instant d'arrivée du précédent : ceci peut être exprimé comme suit,

La probabilité R(t) pour qu'aucun évènement n'arrive entre T et T + t ne dépend que de t quel que soit T. Si l'on considère maintenant les instants successifs T + t et T + t + s = (T + t) + s, pour qu'aucune arrivée n'ait lieu entre T et T + t + s, il faut qu'il n'y ait aucune arrivée entre T et T + t, ni non plus aucune arrivée entre (T + t) et (T + t) + s.

Ces deux évènements étant supposés indépendants, on a donc quels que soient t et s

$$R(t + s) = R(t)\ R(s),$$

dont on sait que la solution générale est de la forme :

$$R(t) = e^{-\lambda t}, \text{ avec } \lambda > 0 \text{ pour } 0 < R(t) < 1$$

On retrouve ainsi directement le résultat précédent.

II.2.2 - Intervalle de temps entre deux arrivées successives d'évènements E

Si un évènement est arrivé à l'instant t = 0 (arbitrairement pris comme origine) la probabilité f(t) dt pour que le suivant ait lieu entre t et t + dt est égale au produit de la probabilité de zéro arrivée entre 0 et t par la probabilité de l'arrivée d'un évènement E entre t et t + dt, soit :

$$f(t)\ dt = \lambda e^{-\lambda t}\ dt$$

La probabilité pour que l'intervalle de temps T entre deux arrivées successives soit inférieur à t est :

$$\Pr(T < t) = F(t) = \int_0^t f(t)\, dt = 1 - e^{-\lambda t}$$

compte tenu de $F(0) = 0$.

On retrouve ainsi la loi exponentielle.

La durée moyenne d'attente entre deux arrivées est :

$$E(T) = 1/\lambda = \theta$$

et la variance de cette durée d'attente est :

$$V(T) = 1/\lambda^2 = \theta^2$$

II.2.3 - <u>Loi du nombre d'intervalles de temps, entre deux arrivées successives, dont la durée est supérieure à une valeur donnée T</u>

Supposons que l'on considère n intervalles de temps entre arrivées successives, le nombre de ces intervalles de durée supérieure à T est une variable aléatoire $Y (y = 0, 1, \ldots, n)$.

La probabilité que $Y = y$ est donnée par la loi binomiale :

$$f(y) = \Pr[Y = y] = C_n^y (e^{-\lambda T})^y \left[1 - e^{-\lambda T}\right]^{n-y}$$

L'espérance mathématique et la variance de Y étant respectivement

$$E(Y) = n\, e^{-\lambda T}$$

$$V(Y) = n\, e^{-\lambda T}\left[1 - e^{-\lambda T}\right]$$

II.2.4 - <u>Loi de l'intervalle de temps entre la $n^{\text{ième}}$ et la $(n + k)^{\text{ième}}$ arrivée de l'évènement E_i</u> (généralisation de la Loi Gamma)

Considérons un évènement A défini par l'arrivée du k^{ieme} des évènements E_i définis précédemment $(n = 0)$.

La probabilité pour que A se produise entre t et t + dt est égale au produit de la probabilité de $(k - 1)$ arrivées entre 0 et t par la probabilité d'une arrivée entre t et t + dt, soit :

$$f(t)\, dt = P_k(t,\ t + dt) = e^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k - 1)!} \lambda dt, \quad \lambda > 0, \quad t > 0$$

$$= \frac{1}{\Gamma(k)} \lambda e^{-\lambda t} (\lambda t)^{k-1}\, dt,$$

(loi d'Erlang ou loi γ_k pour λt).

La forme de la courbe de distribution ne dépend que de k, λ étant un paramètre d'échelle pour les temps.

La probabilité pour que la durée d'attente de A soit inférieure à t est :

$$F(t) = \int_0^t f(t)\, dt = \frac{1}{\Gamma(k)} \int_0^{\lambda t} e^{-x} x^{k-1}\, dx, \qquad (\text{loi } \gamma_k)$$

soit d'après I.6.2 :

$$F(t) = 1 - \sum_{x=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

Le taux instantané d'arrivée de A, qui est dans ce cas une fonction de t, est défini par :

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{\lambda e^{-\lambda t} (\lambda t)^{k-1}}{\Gamma(k) - \int_0^t \lambda e^{-\lambda t} (\lambda t)^{k-1}\, dt} = \frac{\lambda (\lambda t)^{k-1}}{(k-1)! \sum_{x=0}^{k-1} \frac{(\lambda t)^x}{x!}}$$

Ce taux variable avec t, est croissant, constant (égal à λ) suivant que k est supérieur ou égal à l'unité.

Les valeurs caractéristiques de la loi d'Erlang sont :

Mode : $\frac{k-1}{\lambda}$

Médiane : définie par $\frac{1}{\Gamma(k)} \int_0^{\lambda t} e^{-x} x^{k-1}\, dx = \frac{1}{2}$

Moyenne : $E(t) = \frac{k}{\lambda}$

Variance : $V(t) = \frac{k}{\lambda^2}$

On notera que f(t) est aussi la loi de l'intervalle de temps entre la $n^{\text{ème}}$ et la $(n + k)^{\text{ème}}$ arrivée de l'évènement E. La probabilité pour que ce temps soit inférieur à t est encore :

$$F(t) = \int_0^t f(t)\, dt = \frac{1}{\Gamma(k)} \int_0^{\lambda t} e^{-x} x^{k-1}\, dx$$

Remarque

On notera aussi ce résultat important :

Si la loi du nombre X d'arrivées de l'évenement E entre 0 et t est une loi de Poisson de paramètre λt :

$$P_x(0, t) = e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

$$\Pr[X \leqslant c \,|\, (0, t)] = \sum_{x=0}^{c} e^{-\lambda t} \frac{(\lambda t)^x}{x!}$$

$$= \Pr[\chi^2 > 2\lambda t, \ \nu = 2(c+1)]$$

la loi de l'intervalle de temps T entre la $n^{\text{ème}}$ et la $(n + k)^{\text{ème}}$ arrivée de l'évènement E est une loi Gamma de paramètre k :

$$f(t)\, dt = \frac{1}{\Gamma(k)} \lambda\, e^{-\lambda t} (\lambda t)^{k-1}\, dt$$

$$\Pr[T < t] = \frac{1}{\Gamma(k)} \int_0^{\lambda t} e^{-x} x^{k-1}\, dt_x$$

qui pour k = 1 se réduit à la loi exponentielle :

$$f(t)\,dt = \lambda e^{-\lambda t}\,dt$$

$$Pr[T < t] = 1 - e^{-\lambda t}$$

II.2.5 - Test de l'existence d'un processus de Poisson

1/ Pour une durée donnée t, la probabilité de x arrivées entre T et T + t est d'après (II.2.1) :

$$P_x(0,\ t) = e^{-\lambda t}\frac{(\lambda t)^x}{x!} \quad \text{avec } E(X) = \lambda t$$

En présence d'une importante série d'observations, soit m_i le nombre de fois où l'on a observé x_i arrivées de E dans un intervalle de temps fixé t : on pourra utiliser un test de χ^2 pour vérifier que la distribution de x peut être représentée par une loi de Poisson de moyenne λt estimée par la moyenne des x_i observés :

$$\hat{\lambda} t = \frac{\sum m_i x_i}{\sum m_i} = \frac{1}{m}\sum m_i x_i$$

m = nombre total des intervalle de durée t observés.

2/ On pourra aussi, lorsque le nombre n des arrivées de E est grand, procéder graphiquement(1).

Soient t_1, t_2, ..., t_n les instants d'arrivée des évènements E_i, considérons les intervalles successifs :

$$u_1 = t_1,\ u_2 = t_2 - t_1,\ \ldots\ldots\ldots\ldots,\ u_n = t_n - t_{n-1},$$

que l'on pourra classer par ordre de grandeur, afin d'étudier la distribution de ces intervalles.

Si y_i est le nombre de ces intervalles dont la durée est supérieure à une valeur w_i, on aura, dans le cas d'un processus de Poisson :

$$E(Y_i) = n\ e^{-\lambda w_i} \quad (\text{cf. §.II.2.3})$$

soit :

$$\mathrm{Log}_e E(Y_i) = -\ \lambda w_i + \mathrm{Log}_e n$$

ou

$$\log E(Y_i) = -\ 0{,}4343\ \lambda w_i + \log n$$

Dans un graphique semi-logarithmique les points $(w_i,\ y_i)$ devront se placer au voisinage d'une droite issue du point (w = 0, y = n).

Si on trace la droite Δ issue du point (0, n) représentant au mieux le nuage des points-observations $(w_i,\ y_i)$, la pente de cette droite, calculée compte tenu des unités graphiques, permettra d'estimer le paramètre λ de la loi de Poisson.

On notera que la marge de dispersion possible des points observations autour de la droite Δ, marge caractérisée par :

----

(1) Blanc-Lapierre et Fortet. Théorie des fonctions aléatoires (Masson).

$$V(Y_i) = n\, e^{-\lambda w_i}\,(1 - e^{-\lambda w_i})$$

varie avec w_i ; nulle pour $w = 0$, elle est d'abord croissante, passe par un maximum pour $e^{-\lambda w} = 1/2$, soit $w_i = \frac{0,69}{\lambda}$, pour décroître ensuite.

II.2.6 - Plan d'échantillonnage pour acceptation d'un lot d'éléments appartenant à une population dont la durée de vie est distribuée suivant la loi exponentielle (Essai de durée T avec remplacement des défaillants, essai tronqué)

Un élément étant soumis à l'essai à l'instant $t = 0$ est immédiatement remplacé par un élément de la même population et ainsi de suite. La probabilité d'observer un nombre x de défaillances inférieur ou égal à c est, si la vie moyenne est égale à θ :

$$\Pr[x \leqslant c] = \sum_{x=0}^{c} e^{-T/\theta}\,\frac{(T/\theta)^x}{x!}$$

Si au lieu d'un seul élément, n éléments sont soumis à l'épreuve dans les mêmes conditions de remplacement, tout se passe comme dans le cas particulier ci-dessus, mais avec une durée nT

$$\Pr[x \leqslant c] = \sum_{x=0}^{c} e^{-nT/\theta}\,\frac{(nT/\theta)^x}{x!}$$

Ce plan d'échantillonnage défini comme au § II.1.4.1 par n (effectif de l'échantillon), c (critère d'acceptation), T (durée de l'épreuve) et correspondant aux risques $\alpha\ \theta_0$, $\beta\ \theta_1$, $\theta_1 < \theta_0$ devra satisfaire aux conditions :

$$1 - \alpha = \sum_{x=0}^{c} e^{-nT/\theta_0}\,\frac{(nT/\theta_0)^x}{x!}$$

$$\beta = \sum_{x=0}^{c} e^{-nT/\theta_1}\,\frac{(nT/\theta_1)^x}{x!}$$

La détermination d'un tel plan peut se faire immédiatement à l'aide de l'abaque Cavé(1) en y remplaçant p_2/p_1 par θ_0/θ_1 et np par nT/θ (fig.7).

En effet les deux conditions précédentes peuvent s'écrire :

$$\left.\begin{array}{l} 2\,nT/\theta_1 = \chi^2_{1-\beta} \\ 2\,nT/\theta_0 = \chi^2_{\alpha} \end{array}\right\} \quad \nu = 2(c+1)$$

$$\theta_0/\theta_1 = \frac{\chi^2_{1-\beta}}{\chi^2_\alpha} \qquad \nu = 2(c+1)$$

Pour α et β fixés, θ_0/θ_1 est donc une fonction de c dont on pourra construire la courbe représentative $C_{\alpha\beta}$ par points en calculant à l'aide d'une table de χ^2 les valeurs du rapport $\frac{\chi^2_{1-\beta}}{\chi^2_\alpha}$ pour les valeurs successives de c.

----

(1) Cavé R. "Le contrôle statistique des fabrications" Eyrolles (1961)

La courbe $C_{\alpha\beta}$ permettra ensuite de déterminer la valeur de c correspondant au rapport θ_0/θ_1.

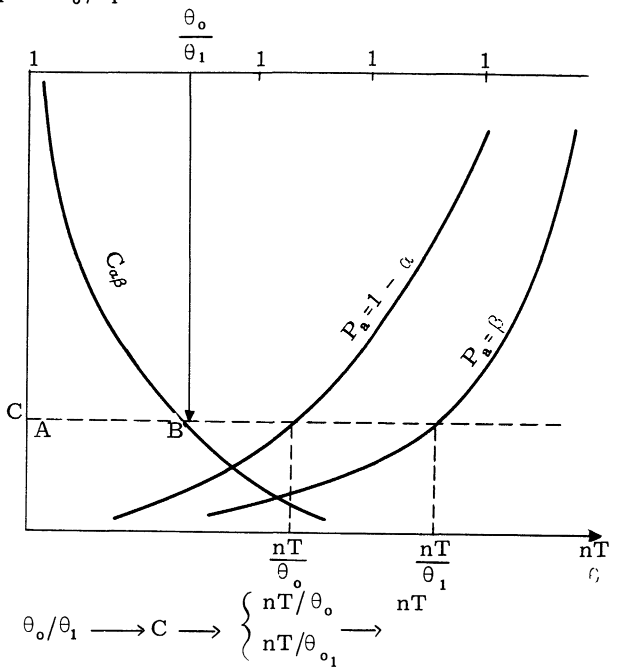


Fig. 7 - Abaque Cavé

Ayant choisi pour c, la valeur entière la plus voisine, les deux équations précédentes donneront pour le produit nT deux valeurs légèrement différentes :

$$(nT)_1 = \frac{1}{2}\,\theta_1\,\chi^2_{1-\beta}$$

$$(nT)_0 = \frac{1}{2}\,\theta_0\,\chi^2_{\alpha}$$

dont on pourra prendre la moyenne pour obtenir tous les plans nT répondant approximativement aux conditions initiales.

(L'abaque Cavé, où sont aussi tracées les courbes $c = f(np) = f(nT/\theta)$, pour diverses valeurs de α et β, permet de lire directement les valeurs de nT/θ_0 et nT/θ_1).

Pour une valeur quelconque de θ, la probabilité d'acceptation est

$$P_a = \sum_{x=0}^{c} e^{-nT/\theta}\,\frac{(nT/\theta)^x}{x!}$$

d'où $2\,nT/\theta = \chi^2_{1-P_a}$, $\quad \nu = 2(c + 1)$.

Le point d'ordonnée c d'une quelconque des courbes P_a donnera la valeur de nT/θ donc de θ qui correspond à la probabilité d'acceptation P_a, ce qui permettra de construire par points la courbe d'efficacité du plan.

Le manuel H. 108, déjà cité en II.1.4 donne les plans (n, r = c + 1 = critère de rejet) correspondant à diverses combinaisons des parametres α, β, θ_o, θ_1 et T (voir tableau III).

ESSAIS TRONQUES AVEC REMPLACEMENT

Tableau III

Valeurs de r pour α, β, θ_1/θ_o et T/θ_o donnés

θ_1/θ_o	r	T/θ_o				r	T/θ_o			
		1/3	1/5	1/10	1/20		1/3	1/5	1/10	1/20
		n	n	n	n		n	n	n	n
		α = 0,01		β = 0,01			α = 0,01		β = 0,05	
2/3	136	331	551	1103	2207	101	237	395	790	1581
1/2	46	95	158	317	634	35	68	113	227	454
1/3	19	31	51	103	206	15	22	37	74	149
1/5	9	10	17	35	70	8	8	14	29	58
1/10	5	4	6	12	25	4	3	4	8	16
		α = 0,01		β = 0,10			α = 0,05		β = 0,01	
2/3	83	189	316	632	1265	95	238	397	795	1591
1/2	30	56	93	187	374	33	72	120	241	483
1/3	13	18	30	60	121	13	23	38	76	153
1/5	7	7	11	23	46	7	9	16	32	65
1/10	4	2	4	8	16	4	4	6	13	27
		α = 0,05		β = 0,05			α = 0,05		β = 0,10	
2/3	67	162	270	541	1082	55	130	216	433	867
1/2	23	47	78	157	314	19	37	62	124	256
1/3	10	16	27	54	108	8	11	19	39	93
1/5	5	6	10	19	39	4	4	7	13	34
1/10	3	3	4	8	16	3	3	4	8	10
		α = 0,10		β = 0,01			α = 0,10		β = 0,05	
2/3	77	197	329	659	1319	52	128	214	429	859
1/2	26	59	98	197	394	18	38	64	128	256
1/3	11	21	35	70	140	8	13	23	46	93
1/5	5	7	12	24	48	4	5	8	17	34
1/10	3	3	5	11	22	2	2	3	5	10
		α = 0,10		β = 0,10			α = 0,25		β = 0,25	
2/3	41	99	165	330	660	12	28	47	95	190
1/2	15	30	51	102	205	5	10	16	33	67
1/3	6	9	15	31	63	2	2	4	9	19
1/5	3	4	6	11	22	1	1	2	3	6
1/10	2	2	2	5	10	1	1	1	2	5

On notera que dans ces plans on a généralement choisi pour r, non la valeur entière la plus voisine de la valeur théorique, mais le plus petit entier satisfaisant à la condition :

$$\frac{\chi^2_\alpha}{\chi^2_{1-\beta}} (\nu = 2\,r) \geqslant \frac{\theta_1}{\theta_o}$$

et pour n l'entier immédiatement inférieur ou égal à $\frac{1}{2\,T}\,\theta_o\,\chi^2_{\alpha}$, $\nu = 2r$.

Mais il faut noter aussi que lorsque θ_o/θ_1 devient important, le critère d'acceptation c diminue et ceci d'autant plus que α et β augmentent (cf. fig. 8). Il en résulte que l'erreur due à la nécessité de choisir pour c une valeur entière - erreur de plus en plus grande en valeur relative - conduit à des valeurs de nT/θ_o et nT/θ_1 donc de nT donnant des plans dont les risques réels peuvent s'écarter notablement des valeurs théoriques α et β.

Les plans (n, r) donnés par le tableau III avec, pour n et r, des valeurs de quelques unités, correspondent à des risques α' et β' pouvant être assez différents des risques α et β figurant dans les cases du tableau.

Ainsi, par exemple, dans la case correspondant aux risques $\alpha = \beta = 0,10$, on trouve pour $\theta_1/\theta_o = 1/10$ et $T/\theta_o = 1/3$, le plan (r = 2, n = 2).

Les risques correspondant à ce plan sont en réalité définis par :

$$\left.\begin{aligned} \chi^2_{1-\beta'} &= 2\,n\frac{T}{\theta_1} = 13,33 \\ \chi^2_{\alpha'} &= 2\,n\frac{T}{\theta_o} = 1,33 \end{aligned}\right\} \quad \nu = 4$$

soit $\alpha' \# 0,15$, $\beta' \# 0,01$ au lieu de 0,10 et 0,10.

Le plan (r = 2, n = 2) est seulement le plan "le moins éloigné" dans l'ensemble des conditions théoriques envisagées.

(Cette remarque s'applique également aux plans donnés dans le tableau I) .

Remarque

Si on se fixe seulement l'un des risques, par exemple θ_o, α, on a :

$$\chi^2_\alpha = \frac{2\,nT}{\theta_o} \qquad \nu = 2(c+1)$$

Exemple : $n = 10 \qquad c = 4 \qquad \theta_o = 1000 \qquad \alpha = 0,10$

$$T = \frac{\theta_o\,\chi^2_{0,10}}{20} \qquad \nu = 10$$

$$T = 243 \text{ h.}$$

On acceptera le lot si au cours d'un essai de 243 heures, on a observé au plus 4 défaillances sur 10 éléments essayés.

La brochure H. 108 donne T/θ_o en fonction de $r = c + 1$ (...... 100), de $n = kr$ (k = 2 ...... 10 et 20) pour α = 0,01, 0,05, 0,10, 0,25 et 0,50.

On pourrait ensuite calculer un des éléments θ_1, β à l'aide de

$$\frac{2\,nT}{\theta_1} = \chi^2_{1-\beta} \qquad \nu = 2(c+1)$$

soit, pour l'exemple précédent si $\theta_1 = 200$

$$20 \times 1,215 = 24,3 = \chi^2_{1-\beta} \qquad (\nu = 10)$$

$$1 - \beta \sim 0,992$$

$$\beta \sim 0,008$$

LOI EXPONENTIELLE
RÉSUMÉ DES TECHNIQUES D'ESSAIS

I - ESTIMATION DE LA VIE MOYENNE

I.1 - Essais censurés (arrêtés à la k^e défaillance) (II.1.7)

I.1.1 - Essai sans remplacement des défaillants

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n-k)\, t_k\right]$$

Intervalle de confiance à $1-\alpha$:

$$\theta_i = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{1-\alpha/2}}, \qquad \theta_s = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{\alpha/2}}, \qquad (\nu = 2\,k)$$

I.1.2 - Essai avec remplacement des défaillants

$$\hat{\theta}_k = \frac{n\,t_k}{k}$$

$$\theta_i = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{1-\alpha/2}}, \qquad \theta_s = \frac{2\,k\,\hat{\theta}_k}{\chi^2_{\alpha/2}}, \qquad (\nu = 2\,k)$$

I.2 - Essais tronqués (arrêtés à l'instant t_o) (II.1.9)

I.2.1 - Essai sans remplacement des défaillants

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n-k)\, t_o\right]$$

Intervalle de confiance à $1-\alpha$ (approximation)

$$\theta_i = \frac{2\,k\,\hat{\theta}_k}{\chi^{2(2k+2)}_{1-\alpha/2}}, \qquad \theta_s = \frac{2\,k\,\hat{\theta}_k}{\chi^{2(2k)}_{\alpha/2}}$$

I.2.2 - Essai avec remplacement des défaillants

$$\hat{\theta}_k = \frac{n\,t_o}{k}$$

$$\theta_i = \frac{2\,k\,\hat{\theta}_k}{\chi^{2(2k+2)}_{1-\alpha/2}}, \qquad \theta_s = \frac{2\,k\,\hat{\theta}_k}{\chi^{2(2k)}_{\alpha/2}}$$

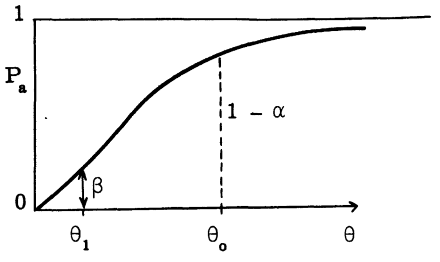
(k = rang de la dernière défaillance observée).

I.2.3 - Essai de durée t_o sans remplacement des défaillants (estimation basée uniquement sur le nombre k des défaillants) (II.1.5)

$$\theta_i = t_o\left[\mathrm{Log}_e\left(1 + \frac{k+1}{n-k}\,F_{1-\alpha/2}\right)\right]^{-1}, \qquad \begin{cases} \nu_1 = 2\,k+2 \\ \nu_2 = 2\,n-2\,k \end{cases}$$

$$\theta_s = t_o \left[\mathrm{Log}_e \left(1 + \frac{k}{n-k+1} F_{\alpha/2}\right)\right]^{-1} \qquad \begin{cases} \nu_1 = 2\,k \\ \nu_2 = 2\,n - 2\,k + 2 \end{cases}$$

II - PLANS D'ECHANTILLONNAGE POUR ACCEPTATION



n : effectif de l'échantillon
P_a : probabilité d'acceptation
α : risque du producteur
β : risque du client
c : critère d'acceptation
r = c + 1 : critère de refus

II.1 - Essai tronqué (durée fixée T, critère d'acceptation c, critère de refus r = c + 1)

II.1.1 - Essai avec remplacement des défaillants (§ II.2.6)

1/ Abaque Cavé : α, β = 0,001 – 0,01 – 0,05 – 0,10 – 0,20 – 0,30
c = 0,1 .... (1) 24

2/ Tableau III : Valeurs de r et n pour :

T/θ_o = 1/3, 1/5, 1/10, 1/20 ; θ_1/θ_o = 2/3, 1/2, 1/3,1/5, 1/10
α, β = 0,01 – 0,05 – 0,10 et α = β = 0,25

r est le plus petit entier tel que :

$$\frac{\theta_1}{\theta_o} \leqslant \frac{\chi^2_\alpha}{\chi^2_{1-\beta}} \qquad (\nu = 2\ r)$$

n est l'entier immédiatement inférieur ou égal à :

$$\frac{1}{2\,T}\,\theta_o\,\chi^2_\alpha \qquad (\nu = 2\ r)$$

II.1.2 - Essai sans remplacement des défaillants (§ II.1.4.1)

Tableau I : valeurs de r et n pour

T/θ_o = 1/3, 1/5, 1/10, 1/20 ; θ_1/θ_o = 2/3, 1/2, 1/3, 1,5, 1/10
α, β = 0,01 – 0,05 – 0,10 et α = β = 0,25

$$F_{1-\beta} = \frac{n-c}{c+1}\left(e^{T/\theta_1} - 1\right) \qquad \begin{cases} \nu_1 = 2(c+1) \\ \nu_2 = 2(n-c) \end{cases}$$

$$F_\alpha = \frac{n-c}{c+1}\left(e^{T/\theta_o} - 1\right)$$

II.2 - Essai censuré (basé sur la vie moyenne $\hat{\theta}_k$ estimée à partir des k premières défaillances, t_k = instant de la k^e défaillance) (II.1.8)

II.2.1 - Essai sans remplacement

$$\hat{\theta}_k = \frac{1}{k}\left[\sum_{i=1}^{k} t_i + (n-k)\, t_k\right]$$

II.2.2 - Essai avec remplacement

$$\hat{\theta}_k = \frac{n\, t_k}{k}$$

Dans les deux cas, la condition d'acceptation est $\hat{\theta}_k \geqslant C$, les valeurs de C étant déduites du tableau 2.

Tableau 2 : Valeurs de k et C/θ_o pour

θ_1/θ_o = 2/3, 1/2, 1/3, 1/5, 1/10
α, β = 0,01 - 0,05 - 0,10 - 0,25

$$\left[C = \theta_o \frac{\chi^2_\alpha}{2\,k}, \quad \nu = 2\,k; \right.$$

$$\left. k = \text{plus petit entier tel que} \frac{\theta_o}{\theta_1} \geqslant \frac{\chi^2_{1-\beta}}{\chi^2_\alpha}, \quad \nu = 2\,k \right]$$

II.2.7 - Interprétation de la loi Gamma dans le cas de défaillance par usure

On peut considérer k comme un paramètre caractérisant la résistance à l'usure dans l'état initial et prenant successivement les valeurs k, k - 1, ... 1, ... 0 sous l'influence de causes élémentaires successives, la défaillance due à la seule usure se produisant lors de l'action de la $k^{\text{ème}}$ cause élémentaire.

Ceci revient à dire que l'organe considéré passe de l'état initial E_k (pour t = 0) à l'état final E_0 (défaillance) par transitions successives E_{k-1}, E_{k-2}, ... E_1, E_0.

Faisons l'hypothèse que la probabilité de passage de l'état E_i à l'état E_{i-1} pendant un petit intervalle de temps dt est μdt, μ étant une constante et l'intervalle de temps dt étant suffisamment petit pour que la probabilité de passage de E_i à E_{i-r} soit nulle pour $r > 1$.

Une étude directe de ce problème conduirait à des calculs analogues à ceux du § III.1.

En désignant par $P_i(t)$ la probabilité pour que l'organe soit dans l'état E_i à l'instant t, on aurait :

$$P_i(t + dt) = P_i(t)\,[1 - \mu dt] + P_{i+1}(t)\,\mu\, dt$$

soit :

$$P_i'(t) - \mu P_i(t) = \mu P_{i+1}(t)$$

dont la solution générale, compte tenu des conditions initiales :

$$P_k(0) = 1 \quad \text{pour } i = k$$

$$P_i(0) = 0 \quad \text{pour } i = 1,\ 2,\ \ldots\ldots\ k-1,$$

est :

$$P_i(t) = \frac{(\mu t)^{k-1}}{(k-i)!} e^{-\mu t}$$

soit en particulier :

$$P_1(t) = \frac{(\mu t)^{k-1}}{(k-1)!} e^{-\mu t}$$

La probabilité pour que la défaillance ait lieu entre t et t + dt, c'est-à-dire pour que l'organe considéré étant dans l'état E_1 à l'instant t passe à l'état E_0 entre t et t + dt sera donnée par la loi Gamma :

$$f(t)\,dt = P_1(t)\,\mu dt = \frac{1}{\Gamma(k)} \mu(\mu t)^{k-1} e^{-\mu t}\,dt$$

d'où, en posant $\mu t = X$

$$F(t) = \int_0^t f(t)\,dt = \frac{1}{\Gamma(k)} \int_0^{\mu t} e^{-x} x^{k-1}\,dx$$

$$= 1 - \sum_{x=0}^{k-1} e^{-\mu t}\frac{(\mu t)^x}{x!},\ (k \text{ entier})$$

le taux instantané de défaillance étant

$$h(t) = \frac{f(t)}{1-F(t)} = \frac{\mu(\mu t)^{k-1}}{(k-1)!\ \sum_{x=0}^{k-1} \frac{(\mu t)^x}{x.}}$$

D'une manière générale, on pourrait envisager une loi de défaillance définie par :

$$f(t)\,dt = \frac{1}{\Gamma(\alpha)} \mu^{\alpha}\, t^{\alpha-1} e^{-\mu t}\,dt \qquad \alpha > 0$$

caractérisée par un taux croissant, constant (et égal à μ) ou décroissant, suivant que α est supérieur, égal ou inférieur à l'unité.

Pour $\alpha = 1$, cette loi Gamma se réduit à la loi exponentielle

II.2.8 - Loi mixte de défaillance (usure et défaillance aléatoire)

Considérons le cas où, en plus de l'usure représentée par le schéma précédent, l'organe est soumis à des accidents aléatoires pouvant le faire passer brusquement d'un quelconque état intermédiaire E_i ($i = k$, $k-1$, ... 2, 1) à l'état E_0 (défaillance).

Soit λdt la probabilité de passage d'un état E_i à l'état E_0 pendant un petit intervalle de temps dt, λ étant une constante indépendante de l'état E_i et de t (loi exponentielle).

La probabilité R(t) de survie à l'instant t est égale au produit de la probabilité $e^{-\lambda t}$ de non défaillance aléatoire avant l'instant t par la probabilité de non défaillance par usure :

$$R(t) = e^{-\lambda t}\left[1 - \int_0^t \frac{1}{\Gamma(\alpha)}\,\mu(\mu t)^{\alpha-1}\,e^{-\mu t}\,dt)\right]$$

d'où :

$$f(t) = e^{-\lambda t}\left[\frac{\mu(\mu t)^{\alpha-1}\,e^{-\mu t}}{\Gamma(\alpha)} + \lambda\left(1 - \frac{1}{\Gamma(\alpha)}\int_0^t \mu(\mu t)^{\alpha-1}\,e^{-\mu t}\,dt\right)\right]$$

(probabilité globale de défaillance entre t et t + dt, que l'on peut encore écrire pour α entier (d'après I.6.2)

$$f(t) = e^{-\lambda t}\left[\frac{\mu(\mu t)^{\alpha-1}\,e^{-\mu t}}{\Gamma(\alpha)} + \lambda\sum_{x=0}^{\alpha-1} e^{-\mu t}\,\frac{(\mu t)^x}{x!}\right]$$

Le taux instantané de défaillance, fonction de t étant :

$$h(t) = \frac{\mu(\mu t)^{\alpha-1}\,e^{-\mu t}}{\Gamma(\alpha) - \int_0^t \mu(\mu t)^{\alpha-1}\,e^{-\mu t}\,dt} + \lambda$$

$$= \lambda + \frac{\mu(\mu t)^{\alpha-1}}{\Gamma(\alpha)\sum_{x=0}^{\alpha-1}\frac{(\mu t)^x}{x!}}$$

(pour α entier).

Cas particuliers

1/ $\mu = 0$, on a

$$f(t) = \lambda e^{-\lambda t}$$

(défaillance par le seul effet d'accident aléatoire)

2/ $\lambda = 0$, on a

$$f(t) = \frac{\mu(\mu t)^{\alpha-1}\,e^{-\mu t}}{\Gamma(\alpha)}$$

(défaillance par le seul effet de l'usure)

3/ λ et $\mu > 0$, $\alpha \nearrow \infty$

f(t) tend vers la distribution exponentielle $f(t) = \lambda e^{-\lambda t}$

4/ $\lambda = 0$, α grand mais fini, on a :

$$f(t)\,dt = \frac{\mu e^{-\mu t}\,(\mu t)^{\alpha-1}}{\Gamma(\alpha)}\,dt$$

soit pour $x = \mu t$, la distribution Gamma :

$$f(x)\,dx = \frac{e^{-x}\,x^{\alpha-1}}{\Gamma(\alpha)}\,dx,$$

qui tend vers la distribution normale de moyenne α et de variance α, soit pour la variable t :

$$E(t) = \alpha/\mu \qquad V(t) = \alpha/\mu^2$$

Remarque

D'une manière plus générale, si la défaillance peut être due à l'action de deux causes indépendantes, la probabilité de survie à l'instant t sera :

$$R(t) = R_1(t).\ R_2(t)$$

d'où :

$$F = 1 - R = 1 - (1 - F_1)(1 - F_2) = F_1 + F_2 - F_1F_2$$
$$f = f_1 + f_2 - f_1F_2 - f_2F_1 = f_1(1 - F_2) + f_2(1 - F_1)$$

et :

$$h = \frac{f}{1 - F} = \frac{f_1}{1 - F_1} + \frac{f_2}{1 - F_2} = h_1 + h_2$$

II.3 - LOI DE WEIBULL

Une généralisation du mécanisme employé pour définir la loi exponentielle consiste à admettre comme hypothèse de départ que le taux instantané de défaillance est une fonction de t, soit :

$$\frac{f(t)}{1 - F(t)} = h(t)$$

d'où, ainsi qu'on l'a vu au début :

$$f(t) = h(t) \exp.\left[- \int_0^t h(t)\, dt\right]$$

Alors que la loi exponentielle supposait que les défaillances survenaient de manière entièrement aléatoire, le taux de défaillance étant constant, l'hypothèse actuelle tient compte de la variation du risque avec la durée.

Des formes très diverses de loi de survie peuvent évidemment être obtenues suivant l'hypothèse faite sur la fonction h(t) : fonction puissance, fonction exponentielle....

Une loi particulièrement importante par ses applications est la loi de Weibull qui correspond à l'hypothèse :

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}, \qquad \beta > 0 \qquad \eta > 0$$

cas particulier très simple d'une fonction puissance de t.

Il en résulte :

$$f(t)\, dt = Pr(t < T < t + dt) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} dt$$
$$F(t) = Pr(T < t) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}} \qquad \text{pour } t > 0$$
$$R(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

Suivant que β est supérieur, égal ou inférieur à 1 la loi de Weibull peut correspondre à un taux instantané de défaillance croissant, constant ou décroissant.

On notera que : η est un simple paramètre de changement d'échelle des temps, souvent désigné sous le nom de vie caractéristique ou caractéristique de vie, qui n'intervient pas dans la forme de distribution, laquelle ne dépend que de β.

Une forme plus générale de la loi de Weibull impliquant un changement d'origine dans l'échelle des temps est(1)

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

$$f(t)\,dt = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}} dt$$

II.3.1 - Propriétés

Les valeurs caractéristiques de la loi de t sont respectivement

Mode : $t = \gamma + \eta(1 - 1/\beta)^{1/\beta}$

Médiane : $t = \gamma + \eta(\mathrm{Log}_e 2)^{1/\beta}$

Moyenne : $E(T) = \gamma + \eta\,\Gamma(1 + 1/\beta) = m$

Variance : $V(T) = \eta^2[\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)] = \sigma^2$

On remarquera que pour $\gamma = 0$, le rapport m/σ est indépendant de η.

II.3.2 - Etude graphique [14], [12]

De la relation

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$

On déduit :

$$\mathrm{Log}_e.\ \mathrm{Log}_e \frac{1}{1 - F(t)} = \beta\ \mathrm{Log}_e\, t - \beta\ \mathrm{Log}_e\, \eta$$

relation linéaire entre $\mathrm{Log}_e t$ et $\mathrm{Log}_e\, \mathrm{Log}_e \frac{1}{1 - F(t)}$

Si pour une valeur de t, F(t) est estimé par la fréquence cumulée F_i des observations inférieures à t, on pourra :

----

(1) Certains auteurs les présentent aussi sous l'une des formes équivalentes :

$$F(t) = 1 - \exp\left[-\frac{(t-\gamma)^{\beta}}{\alpha}\right]$$

ou

$$F(t) = 1 - \exp\left[-\lambda(t-\gamma)^{\beta}\right]$$

les relations entre les paramètres α, λ, η étant

$$\eta^{\beta} = \alpha = 1/\lambda$$

- soit calculer $Y_i = \mathrm{Log}_e \mathrm{Log}_e \dfrac{1}{1 - F_i}$, $X_i = \mathrm{Log}_e t_i$ et vérifier, si la variable t est distribuée suivant la loi de Weibull, que les points (X_i, Y_i) sont au voisinage d'une droite dont l'équation permettre d'estimer β et η.

- soit plus simplement, utiliser un graphique à échelles fonctionnelles, permettant de placer immédiatement les points (t_i, F_i), [10] et [11] (fig. 9).

Le graphique donne immédiatement η et β dans le cas d'une loi à deux paramètres et permet, s'il y a lieu de déterminer le paramètre γ dans le cas d'une loi de Weibull du type $F(t) = 1 - \exp\left[-\left(\dfrac{t-\gamma}{\eta}\right)^{\beta}\right]$ (1).

Cas particuliers

1/ Pour $\gamma = 0$ et $\beta = 1$ on obtient la loi exponentielle

$$f(t) = \frac{1}{\eta} e^{-t/\eta} \qquad h(t) = \frac{1}{\eta} = \text{Cte}$$

$$F(t) = 1 - e^{-t/\eta}$$

$$E(t) = \eta \qquad V(t) = \eta^2 .$$

II.4 - LOIS DES VALEURS EXTREMES

La loi exponentielle et la loi de Weibull semblent constituer jusqu'à maintenant les deux principaux outils statistiques des études de fiabilité associés à l'idée d'une loi de durée de vie caractérisée par son taux instantané de défaillance.

D'autres lois peuvent évidemment être envisagées suivant l'hypothèse faite sur la façon dont le taux de défaillance varie en fonction de t.

Une autre approche du problème peut être envisagée en admettant que la défaillance d'un organe est provoquée par le ou les défauts les plus graves de cet organe, ce qui conduit à envisager l'utilisation des lois asymptotiques des valeurs extrêmes pouvant représenter la distribution des défauts suivant leur gravité.

Ces modèles ont été surtout étudiés en vue de leur application aux études de fatigue des matériaux : Freudenthal, Gumbel [8], [9]. Dans ce domaine, où la vérification de l'adéquation d'un modèle à la réalité et l'estimation statistique de ses paramètres deviennent pratiquement impossibles lorsqu'il s'agit de structures et non d'éprouvettes, les travaux modernes s'orientent vers la recherche de modèles tenant compte non seulement de l'étude statistique des durées de vie mais aussi des caractères physiques associés à la réalisation de la défaillance.

Laissant de côté cet aspect de la question, nous examinerons rapidement dans ce qui suit la loi généralement connue sous le nom de loi

----

(1) On recherchera par approximations successives, s'il existe une valeur de γ, telle que la variable $t' = t - \gamma$ satisfasse à la loi :

$$F(t') = 1 - \exp.\left[-\left(\frac{t'}{\eta}\right)^{\beta}\right]$$

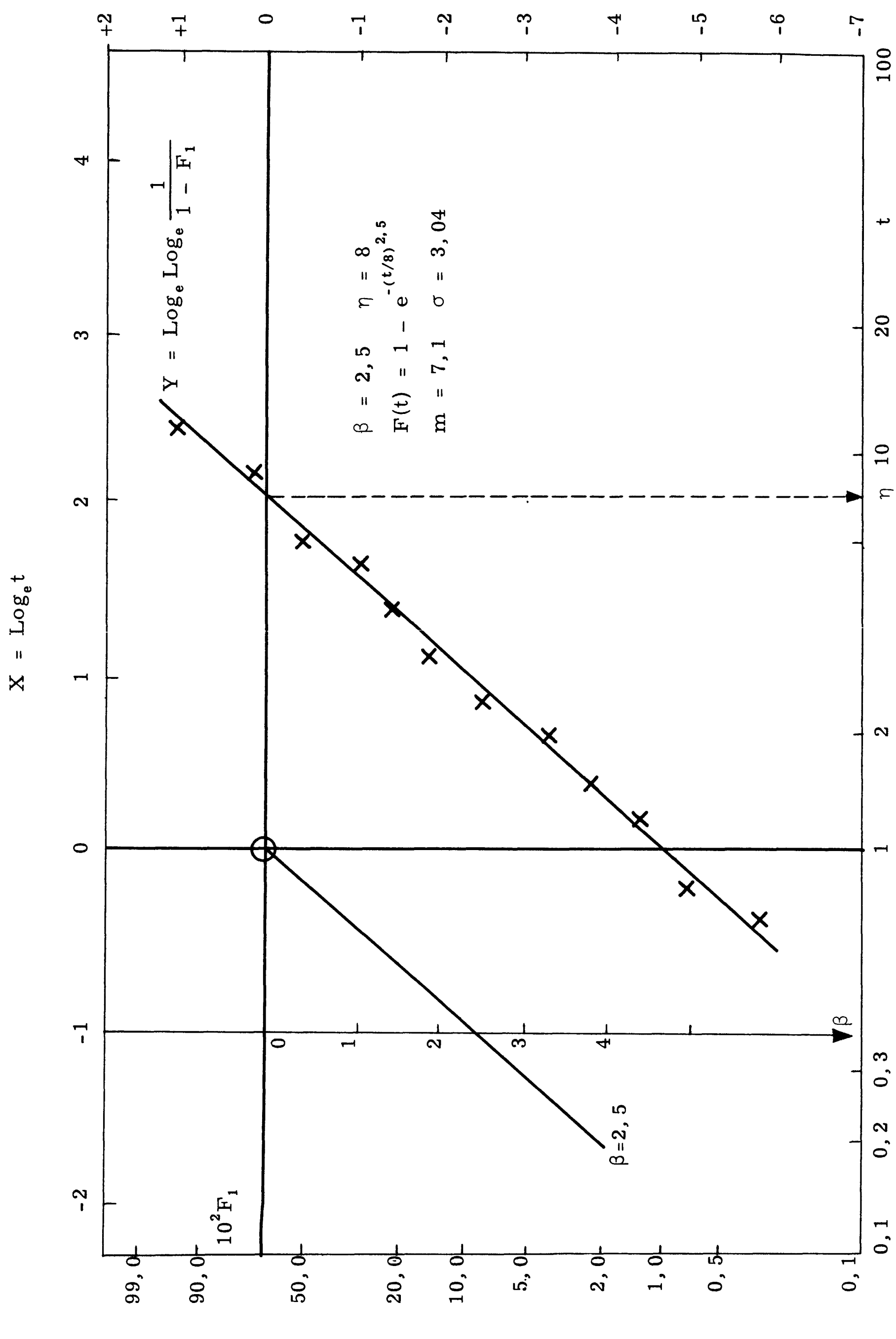


Fig. 9 - Graphique de Weibull

Tableau 4

Loi de Weibull :

$$f(x)\,dx = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right] dx \qquad m = E(X) = \eta\,\Gamma\left(1+\frac{1}{\beta}\right)$$

$$R(x) = 1 - F(x) = \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right] \qquad \sigma^2 = V(X) = \eta^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]$$

β	$m/\eta = \Gamma(1+1/\beta)$	σ/η	β	$m/\eta = \Gamma(1+1/\beta)$	σ/η	β	$m/\eta = \Gamma(1+1/\beta)$	β
0	∞	∞	1,3	0,9235	0,716	2,6	0,8882	0,37
0,1	10!	$\sqrt{20! - (10!)^2}$	1,4	0,9114	0,659	2,7	0,8893	0,36
0,2	120	1901	1,5	0,9028	0,613	2,8	0,8905	0,34
0,3	9,2605	47	1,6	0,8966	0,594	2,9	0,8917	0,33
0,4	3,3234	10,43	1,7	0,8922	0,530	3,0	0,8938	0,32
0,5	2,0000	4,472	1,8	0,8893	0,512	3,1	0,8943	0,315
0,6	1,5046	2,645	1,9	0,8874	0,486	3,2	0,8957	0,31
0,7	1,2658	1,851	2,0	0,8862	0,463	3,3	0,8970	0,30
0,8	1,1330	1,428	2,1	0,8857	0,44	3,4	0,8984	0,29
0,9	1,0522	1,171	2,2	0,8856	0,42	3,5	0,8998	0,28
1,0	1,0000	1,000	2,3	0,8859	0,41	3,6	0,9011	0,27
1,1	0,0649	0,878	2,4	0,8865	0,39	3,8	0,9038	0,26
1,2	0,9407	0,785	2,5	0,8873	0,38	4,0	0,9064	0,25

des valeurs extrêmes et les lois asymptotiques de la plus petite valeur dont la précédente n'est d'ailleurs qu'un cas particulier.

II.4.1 - Loi des valeurs extrêmes

Dans la distribution de Weibull, le taux instantané de défaillance est une fonction puissance de t.

Si l'on suppose maintenant que le taux instantané de défaillance, h(t), croît lui-même à taux constant au cours du temps (hypothèse de Gompertz), on aura :

$$\frac{1}{h(t)} \frac{dh(t)}{dt} = \lambda = \text{Cte}, \quad \lambda > 0$$

d'où :

$$h(t) = k\, e^{\lambda t}, \quad k > 0$$

k étant la valeur initiale du taux de défaillance.

On en déduit :

$$f(t) = h(t) \exp\left[-\int_0^t h(t)\, dt\right]$$
$$= k\, e^{\lambda t} \exp\left[-\frac{k}{\lambda}(e^{\lambda t} - 1)\right]$$

Si l'on pose $k/\lambda = \alpha$, on aura :

$$\Pr(t < T < t + dt) = f(t)\, dt = \alpha \quad e^{\lambda t}\, e^{-\alpha(e^{\lambda t}-1)}\, dt$$

$$\Pr(T < t) = F(t) = 1 - e^{-\alpha(e^{\lambda t}-1)}$$

compte tenu de la condition initiale $F(0) = 0$, et [$h(t) = \alpha\, \lambda\, e^{\lambda t}$ définissant, pour $\alpha > 0$, un taux instantané croissant.

Comme précédemment, λ est un paramètre d'échelle qui n'intervient pas dans la forme de la courbe de distribution, laquelle ne dépend que de α.

II.4.1.1 - Exemple de génération de la loi des valeurs extrêmes(1)

Supposons :

1/ que la défaillance d'un ensemble composé d'un très grand nombre N d'éléments E se produit dès la première défaillance de l'un de ses éléments.

2/ qu'à l'instant initial chacun de ces éléments soit caractérisé par un paramètre a_i pouvant prendre toute valeur entre zéro et A, la distribution de ces éléments a_i étant une distribution exponentielle tronquée de paramètre γ et que dès que l'un de ces éléments atteint la valeur A, la défaillance se produit.

----

(1) Lloyd and Lipow [10].

3/ que l'évolution des a_i (i = 1, 2, ... N) au cours du temps vers la valeur A se fait suivant une loi linéaire ; soit :

$$\theta_i = k(A - a_i),$$

le temps nécessaire pour que a_i devienne égal à A.

On a alors :

$$\Pr(a_i \geqslant a) = \frac{e^{-\gamma a} - e^{-\gamma A}}{1 - e^{-\gamma A}}, \qquad 0 < a < A$$

D'autre part, si t_i est l'instant de la défaillance de l'élément E_i, on aura :

$$G(t) = \Pr(\theta_i \leqslant t) = \Pr\left[a_i \geqslant A - \frac{t}{k}\right]$$

$$= \frac{e^{\frac{\gamma t}{k}} - 1}{e^{\gamma A} - 1} \qquad 0 \leqslant t \leqslant kA$$

Soit T l'instant de la défaillance, c'est-à-dire la plus petite valeur des θ_i, on aura :

$$F(t) = \Pr(T \leqslant t) = 1 - [1 - G(t)]^N$$

Soit pour $N \nearrow \infty$

$$F(t) \nearrow 1 - e^{-NG(t)}$$

c'est-à-dire :

$$F(t) \nearrow 1 - \exp\left[-\frac{N}{e^{\gamma A} - 1}\left(e^{\frac{\gamma t}{k}} - 1\right)\right]$$

qui est de la forme :

$$F(t) = 1 - e^{-\alpha(e^{\lambda t} - 1)}$$

si on pose :

$$\frac{N}{e^{\gamma A} - 1} = \alpha \qquad \gamma/k = \lambda$$

Ce modèle a été utilisé par Epstein [4] pour étudier un phénomène de corrosion, la défaillance se produisant lors de la première perforation de la paroi considérée d'épaisseur A.

La loi des valeurs extrêmes a aussi été utilisée dans des études de fatigue des matériaux par divers auteurs par exemple par Gumbal [8] et [9], et Freudenthal sous des formes un peu différentes étudiées ci-après.

II.4.2 - Généralisation

La distribution envisagée ci-dessus sous le nom de "loi des valeurs extrêmes" n'est en réalité qu'un cas particulier des lois asymptotiques de distribution de la plus petite (ou de la plus grande) valeur, de n observations indépendantes d'un échantillon, d'une variable aléatoire X distribuée elle-même suivant une certaine loi F(x).

Soit $X_1 = x_1$, la plus petite valeur de cette série de n observations de X.

X_1 est aussi une variable aléatoire dont la fonction de répartition est :

$$\Phi_n(x) = \Pr[X_1 < x] = 1 - [1 - F(x)]^n$$

Fisher et Tippett [3] ont montré l'existence de trois formes limites de $\Phi_n(x)$ pour $n \nearrow \infty$: deux d'entre elles correspondent respectivement à la loi des valeurs extrêmes étudiée ci-dessus et à la loi de Weibull.

Elles ont été étudiées par Gumbel [4] et [5], qui les a appliquées à des études de fatigue des matériaux et mises en tables [41] avec des notations différentes de celles utilisées ci-dessus.

II.4.2.1 - Loi limite du type I

Sous ce nom, Gumbel envisage la loi limite

$$\Pr[X_1 < x] = \Phi(x) = 1 - \exp\left[- e^{\frac{x-u}{\beta}}\right] \quad - \infty < x < \infty^{(1)}, \quad \beta > 0$$

Si l'on pose :

$$y = \frac{x - u}{\beta} \qquad \text{(variable réduite)}$$

on aura :

$$\Phi(x) = 1 - \exp[- e^y]$$

La densité de probabilité de X_1, étant :

$$\varphi(x) = \frac{1}{\beta} \exp(y - e^y)$$

On aura aussi :

$$R(x) = \Pr[X_1 > x] = \exp[- e^y]$$

soit :

$$\mathrm{Log}_e [- \mathrm{Log}_e R(x)] = y$$

Les valeurs caractéristiques de cette distribution sont :

Mode : u

Médiane : $u - 0,36651\ \beta$, $- 0,36651 = \mathrm{Log}_e(\mathrm{Log}_e 2)$

Moyenne : $E(X) = u - 0,57722\ \beta$, 0,57722 : constante d'Euler

Variance : $V(X) = \left(\frac{\pi\beta}{\sqrt{6}}\right)^2$

----

(1) Intervalle non borné à gauche pourvu que la probabilité initiale $1 - F(x)$ converge vers zéro au moins aussi rapidement que e^{-x} lorsque x croît.

Une table de la fonction $y = \mathrm{Log}_e\,[-\mathrm{Log}_e\,R(x)]$ a été publiée par le National Bureau of Standards [41]. Elle permet de construire une échelle graphique graduée en valeurs de R(x), les longueurs étant proportionnelles à y. Si la loi est vérifiée par les observations étudiées, les points (x, R(x)) devront être dispersés autour de la droite d'équation :

$$y = \frac{x - u}{\beta}$$

Les deux paramètres u et β peuvent d'ailleurs être estimés à l'aide de la moyenne et de l'écart-type empiriques ($\bar{x}$ et s) par :

$$\hat{u} = \bar{x} + 0{,}5772\ \beta \qquad \beta = \frac{s\sqrt{6}}{\pi},$$

On notera que moyennant le changement de variables et de paramètres défini, (Gumbel [9]), par :

$$e^{\lambda t} - 1 = e^{x/\beta}, \qquad \alpha = e^{-u/\beta}$$

la loi $\Phi(x)$ est la loi des valeurs extrêmes précédemment indiquée (II.4.1)

II.4.2.2 - Loi limite du type III

Si on pose :

$$y = \mathrm{Log}_e \left(\frac{x - \varepsilon}{V - \varepsilon}\right)^{\beta'} \qquad \text{avec } \beta' > 0$$

on obtient la loi limite

$$\Psi(x) = \Pr[X_1 < x] = 1 - \exp\left[-\left(\frac{x - \varepsilon}{V - \varepsilon}\right)^{\beta'}\right]$$

pour $\beta' > 0 \quad x \geqslant \varepsilon \geqslant 0, \quad V > \varepsilon$

La densité de probabilité est :

$$\psi(x) = \frac{\beta'}{V - \varepsilon}\left(\frac{x - \varepsilon}{V - \varepsilon}\right)^{\beta' - 1} \exp\left[-\left(\frac{x - \varepsilon}{V - \varepsilon}\right)^{\beta'}\right]$$

La loi de Weibull (pour $\gamma = 0$) est un cas particulier de cette loi correspondant à $\varepsilon = 0$, $V = \eta$.

Les valeurs caractéristiques de cette distribution sont :

Mode : $\varepsilon + (V - \varepsilon)\,(1 - 1/\beta')^{1/\beta'}$ (si $\beta' > 1$)

Médiane : $\varepsilon + (V - \varepsilon)\,[\mathrm{Log}_e 2]^{1/\beta'}$

Moyenne : $E(X) = \varepsilon + (V - \varepsilon)\,\Gamma(1 + 1/\beta')$

Variance : $V(X) = (V - \varepsilon)^2\,[\Gamma(1 + 2/\beta') - \Gamma^2(1 + 1/\beta')]$

On remarquera que :

1/ $\Phi(x = u) = \Psi(x = V) = 1 - 1/e$

2/ Si on pose $x - \varepsilon = e^z$, $V - \varepsilon = e^u$, on a

$$y = \mathrm{Log}_e\left(\frac{x - \varepsilon}{V - \varepsilon}\right)^{\beta'} = \beta'(z - u)$$

Il en résulte que si X_1 suit une loi du type III, la variable

$$Z = \mathrm{Log}_e(X - \varepsilon)$$

suit une loi du type I.

$$\Psi(z) = 1 - \exp\left[- e^{-(z-u)\beta'}\right]$$

permettant l'emploi du papier à échelles fonctionnelles envisagé ci-dessus.

3/ Si $\varepsilon = 0$, considérons la transformation :

$$y = \beta'(\mathrm{Log}_e x - \mathrm{Log}_e V) = 2,3026\ \beta'(\log_{10} x - \log_{10} V)$$

Dans ce cas les points correspondant aux logarithmes des plus petites valeurs observées devront se répartir autour de la droite :

$$\log_{10} x = \log_{10} V + \frac{y}{2,3026\ \beta'}$$

Les deux paramètres V et β' pourront encore être estimés à partir de la moyenne empirique $m(\log_{10} x)$ et de l'écart-type empirique $s(\log_{10} x)$ par :

$$\log_{10} \hat{V} = m(\log_{10} x) + \frac{0,5772}{2,3026\ \beta'}$$

$$\frac{1}{2,3026\ \beta'} = \frac{\sqrt{6}}{\pi}\, s(\log_{10} x)$$

III. ANNEXES

On trouvera en annexe :

1/ Un tableau récapitulatif des lois étudiées et de leurs principales caractéristiques.

2/ des graphiques représentant f(t), R(t) et h(t) pour les principales distributions observées (compte tenu seulement du paramètre de forme pour les distributions Gamma, de Weibull et des valeurs extrêmes) ; le paramètre d'échelle étant pris égal à l'unité, la comparaison de ces lois ne peut évidemment porter que sur la forme et non sur les valeurs correspondant à une valeur donnée de t,

3/ une bibliographie d'ouvrages et d'articles sur ces questions avec des indications sur quelques graphiques à échelles fonctionnelles existant dans le commerce,

4/ une liste des principales tables statistiques permettant les calculs numériques exigés par les distributions étudiées.

ANNEXE I

I - LOI BINOMIALE

$f(x) = C_n^x p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots n$ Mode $(n+1)p - 1 < x < (n+1)p$

$F(x) = \sum_{x=0}^{x} C_n^x p^x (1-p)^{n-x}$

Médiane $F(x) = 1/2$
Moyenne $E(X) = np$
Variance $V(X) = np(1-p)$

I. 1 - Loi binomiale négative

1/ N = nombre des épreuves pour obtenir la $c^{ième}$ arrivée

$f(n) = C_{n-1}^{c-1} p^c q^{n-c}, \quad n = c, c+1, \ldots$ Mode : $\frac{c-1}{p} < n < \frac{c-q}{p}$

$F(n) = \sum_{j=c}^{n} C_{j-1}^{c-1} p^c q^{j-c} = \sum_{x=c}^{n} C_n^x p^x q^{n-x}$

Médiane $F(n) = 1/2$
Moyenne $E(N) = c/p$
Variance $V(N) = qc/p^2$

2/ Loi de $Z = N - c$ (Nombre des non-arrivées avant la $c^{ième}$ arrivée)

$f(z) = C_{z+c-1}^{c-1} p^c q^z, \quad z = 0, 1, \ldots.$ Mode $\frac{qc-1}{p} < z < \frac{qc-q}{p}$

$F(z) = \sum_{i=o}^{z} C_{c+i-1}^{c-1} p^c q^i$

Médiane $F(z) = 1/2$
Moyenne $E(Z) = qc/p$
Variance $V(Z) = qc/p^2$

3/ Loi géométrique (de N ou de Z) ; pour $c = 1$

II - LOI DE POISSON

$f(x) = e^{-m} \frac{m^x}{x!}, \quad x = 0, 1, \ldots..$

$F(x) = \sum_{x=0}^{x} e^{-m} \frac{m^x}{x!}$

Mode $m - 1 < x < m$
Médiane $F(x) = 1/2$
Moyenne $E(X) = m$
Variance $V(X) = m$

III - LOI NORMALE

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad -\infty < x < \infty$

$F(x) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$

avec $u = \frac{x-m}{\sigma}$

Mode m
Médiane m
Moyenne $E(X) = m$
Variance $V(X) = \sigma^2$

IV - LOI DE χ^2

$$f(\chi^2) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-\frac{\chi^2}{2}} (\chi^2)^{\nu/2-1}$$

$$F(\chi^2) = \int_0^{\chi^2} f(\chi^2)\, d(\chi^2)$$

Mode $\nu - 2$
Médiane $F(\chi^2) = 1/2$
Moyenne $E(\chi^2) = \nu$
Variance $V(\chi^2) = 2^{\nu}$

V - LOI LOG-NORMALE DE X

(Loi normale de Y = $\log_{10}(X - a)$

$$f(x) = \frac{1}{x-a}\,\frac{0,4343}{\sigma\sqrt{2\pi}}\, e^{-\frac{(y-m)^2}{2\sigma^2}}, \quad x > a$$

(m et σ : moyenne et écart-type de la variable $Y = \log_{10}(X - a)$)

Mode $a + e^{m'-\sigma'^2}$
Médiane $a + e^{m'}$
Moyenne $E(X) = a + e^{m'+\frac{\sigma'^2}{2}}$
Variance $V(X) = e^{2m'+\sigma'^2}(e^{\sigma'^2} - 1)$
avec

$m' = 2,3026\, m, \quad \sigma' = 2,3026\, \sigma$

VI - LOI EXPONENTIELLE

$f(t) = \lambda e^{-\lambda t}, \ \lambda > 0, \ t > 0$

$F(t) = 1 - e^{-\lambda t}$

$$h(t) = \frac{f(t)}{1 - F(t)} = \lambda$$

Médiane $\frac{1}{\lambda} \operatorname{Log}_e 2$

Moyenne $E(T) = 1/\lambda$

Variance $V(T) = 1/\lambda^2$

VII - PROCESSUS DE POISSON : X, nombre d'arrivées dans un intervalle de durée t

$$Pr[X = x \mid (0, t)] = e^{-\lambda t}\frac{(\lambda t)^x}{x!}$$

$$Pr[X < x \mid (0, t)] = \sum_{x=0}^{x} e^{-\lambda t}\frac{(\lambda t)^x}{x!}$$

$$Pr[E \mid (0, dt)] = \lambda dt$$

$(\lambda > 0)$

Mode $\lambda t - 1 \leqslant x \leqslant \lambda t$
Médiane $\sum_{x=0}^{x} e^{-\lambda t}\frac{(\lambda t)^x}{x!} = 1/2$
Moyenne $E(X) = \lambda t$
Variance $V(X) = \lambda t$

VII. 1 - Loi du nombre Y d'intervalles de temps, entre deux arrivées successives, dont la durée est supérieure à une valeur donnée T, pour un nombre total n d'intervalles entre deux arrivées successives.

$$Pr[Y = y] = C_n^y (e^{-\lambda T})\,[1 - e^{-\lambda T}]^{n-y}$$

$(\lambda > 0)$

$$Pr[Y \leqslant y] = \sum_{x=0}^{y} Pr[Y = y]$$

Mode $(n + 1)e^{-\lambda T} - 1 < y < (n + 1)e^{-\lambda T}$
Médiane $\sum_{y=0}^{y} Pr[Y = y] = 1/2$
Moyenne $E(Y) = n\, e^{-\lambda T}$
Variance $V(Y) = n\, e^{-\lambda T}[1 - e^{-\lambda T}]$

VIII- LOI GAMMA

$$f(t) = \frac{\lambda}{\Gamma(k)} e^{-\lambda t} (\lambda t)^{k-1}, \quad \lambda > 0, \ t > 0$$

Mode $\frac{k-1}{\lambda}$

$$F(t) = \frac{1}{\Gamma(k)} \int_0^{\lambda t} e^{-x} x^{k-1} dx$$

$$= 1 - \sum_{x=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^x}{x!}, \text{ (k entier)}$$

Médiane $F(t) = 1/2$
Moyenne $E(T) = k/\lambda$
Variance $V(T) = k/\lambda^2$

$$h(t) = \frac{f(t)}{1 - F(t)}$$

$$= \frac{\lambda(\lambda t)^{k-1}}{(k-1)! \sum_{x=0}^{k-1} \frac{(\lambda t)^x}{x!}}, \text{ k entier}$$

IX - LOI DE WEIBULL

$$f(t) = \frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

$$h(t) = \frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1}$$

$(\beta > 0,\ \eta > 0,\ t \geqslant \gamma \geqslant 0)$

Mode $\gamma + \eta(1 - 1/\beta)^{1/\beta}$
Médiane $\gamma + \eta(\mathrm{Log}_e 2)^{1/\beta}$
Moyenne $E(T) = \gamma + \eta\Gamma(1 + 1/\beta)$
Variance $V(T) = \eta^2[\Gamma(1 + 2/\beta) -$
$- \Gamma^2(1 + 1/\beta)]$

X - LOIS DES VALEURS EXTREMES

$$f(t) = \alpha\lambda e^{\lambda t} e^{-\alpha(e^{\lambda t}-1)}$$

$$F(t) = 1 - e^{-\alpha(e^{\lambda t}-1)}$$

$$h(t) = \alpha\lambda e^{\lambda t}$$

$(\alpha > 0,\ \lambda > 0)$

Mode $\frac{1}{\lambda} \mathrm{Log}_e \frac{1}{\alpha}$
Médiane $\frac{1}{\lambda} \mathrm{Log}_e [1 + \frac{1}{\alpha} \mathrm{Log}_e 2]$

Loi type I de GUMBEL

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

$$F(x) = 1 - e^{-e^{\frac{x-u}{\beta}}}$$

$$h(x) = 1/\beta . e^{\frac{x-u}{\beta}}$$

$(\beta > 0,\ -\infty < x < \infty)$

Mode u
Médiane $u - 0,36651\ \beta$
Moyenne $E(X) = u - 0,57722\ \beta$
Variance $V(X) = \left(\frac{\pi\beta}{\sqrt{6}}\right)^2$

Loi type III de GUMBEL

$$f(x) = \frac{\beta}{V-\varepsilon} \left(\frac{x-\varepsilon}{V-\varepsilon}\right)^{\beta-1} \exp\left[-\left(\frac{x-\varepsilon}{V-\varepsilon}\right)^{\beta}\right]$$

$$F(x) = 1 - \exp -\left[\left(\frac{u-\varepsilon}{V-\varepsilon}\right)^{\beta}\right]$$

$$h(x) = \frac{\beta}{V-\varepsilon} \left(\frac{x-\varepsilon}{V-\varepsilon}\right)^{\beta-1}$$

$(\beta > 0,\ x \geqslant \varepsilon \geqslant 0,\ V > \varepsilon)$

Mode $\varepsilon + (V - \varepsilon)\ (1 - 1/\beta)^{1/\beta}$
Médiane $\varepsilon + (V - \varepsilon)\ [\mathrm{Log}_e 2]^{1/\beta}$
Moyenne $E(X) = \varepsilon + (V - \varepsilon).$
$\Gamma(1 + 1/\beta)$
Variance $V(X) = (V - \varepsilon)^2 [\Gamma(1 + 2/\beta)$
$- \Gamma^2\left(1 + \frac{1}{\beta}\right)]$

ANNEXE III

[1] CALOT - Cours du Calcul des probabilités, DUNOD, 1963.

[2] EUGENE - La fatigue des structures. Thèse.. Publications SUD-AVIATION 1962.

[3] FISHER and TIPPETT - Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings Cambridge Philosophical Society 24 Part. 2, 1928 pp. 180-190.

[4] EPSTEIN - The exponentiel distribution and its role in life testing Industrial Quality Control Déc. 1958.

[5] WEIBULL - A statistical representation of fatigue failure in solids. Trans. Royal Institute of Technology n° 27, Stockholm, 1949.

[6] WEIBULL - A statistical distribution fonction of wide applicability. Journal of Applied Mechanics Vol 18 n° 1 Mars 1951, pp. 293-297.

[7] Alan PLAIT - The Weibull distribution with tables.Industrial Quality Control Nov. 1962.

[8] GUMBEL - Statistics of extrems. Columbia University Press, 1958.

[9] GUMBEL - Etude statistique de la fatigue des matériaux. Revue de Statistique Appliquée, 1957 n° 4.

[10] LLOYD and LIPOW - Reliability : management, methods and mathematics. Prentice Hall, 1964.

[11] MORICE - Les graphiques à échelles fonctionnelles du statisticien. Revue de Statistique Appliquée, 1964 n° 3.

[12] "Fiabilité" - Revue publiée par le Centre National d'Etudes des Télécommunications n° 1 à 5.

Les papiers à échelles fonctionnelles, semi-logarithmique, logarithmique-logarithmique, log-normale se trouvent couramment dans le commerce avec des formats divers et des échelles logarithmiques à divers modules, par exemple à la Compagnie Française des Diagrammes, avenue d'Inkermann à Neuilly.

Un papier spécial pour étude graphique de la loi binomiale (échelles racine carrée : "Binomial Probability Paper"), est édité par CODEX Book C°, Norwood, Massachusets, U.S.A.

La Société "Technical and Engineering Aids for Management", 104 Belrose Avenue, Lowell, Massachusets (U.S.A.) édite des papiers à échelles fonctionnelles pour étude graphique de la loi de Weibull en 9 types différents : échelles des fréquences cumulées de 0,0001, ou 0,01, ou 1 à 99,9 %, échelles logarithmiques à 3,5 ou 7 modules.

Cette Société édite aussi des papiers spéciaux pour la loi des valeurs extrêmes, la loi de χ^2 (pour $\nu = 1, 2, 3, 4, 6, 8$ ou 10 degrés de liberté, ou loi Gamma de paramètre $\nu/2$).

En Allemagne, la maison "Schäfers Feinpapier", 4 Bergstrasse, Plauen (Vogtl) édite aussi les papiers ci-après :

- Binomial papier n° 584 format 21×30, $0 < x < 250$, $0 < y < 300$;

- Rosin-Rommler-Bennett papier n° 604 et 605 formats 30 × 40 et 42 × 60 (préparés pour l'étude graphique de la loi $R = e^{-(d/d')^n}$, ils peuvent être utilisés pour la loi de Weibull $R(t) = e^{-(t/\eta)^\gamma}$.

La Compagnie Française des Diagrammes vient d'éditer en deux types différents des papiers à échelles fonctionnelles pour la loi de Weibull : format 21 × 27, fréquences cumulées de 0,001 à 0,999, échelle logarithmique à trois modules. L'un comporte une échelle circulaire des pentes pour β avec une échelle de correspondance de β vers $\Gamma(1 + 1/\beta)$, l'autre une échelle de tangentes pour β.

ANNEXE IV

IV.1 - TABLES GENERALES

[21] Tables Statistiques - Centre de Formation aux Applications Industrielles de la Statistique (Université de Paris).

Loi binomiale : f et F pour n = 5, 10, 15, 20, 30, 40, 50, p = 1 (1) .... 10 %

Loi de Poisson : f et F pour m = 0,1 (0,1) ...... 1 (0,5) 10 (1) 18

Loi de χ^2 : ν = 1 (1) 30

α = 0,99, 0,975, 0,95, 0,90, 0,10, 0,05, 0,025, 0,01, 0,001

Loi normale ; f(u) pour u = 0 (0,1) 4
F(u) pour u = 0 (0,01) 3

Loi de F : 1 – α= 0,95 et 0,99.

[22] HALD - Statistical tables and formules. John Wiley, New-York

Loi normale : f(u) pour u = 0 (0,01) 4,99
F(u) pour u = 0 (0,01) 4,99

Loi de χ^2 : ν = 1 (1) 50

α et 1 – α = 0,05, 0,1, 0,5, 1, 2,5, 5, 10 (10) ..... 50 (en %)

Loi binomiale : limites de confiance à 0,05 et 0,90 à partir de l'estimation

$$f = \frac{x}{n - x} \text{ pour}$$

x et n – x = 0 (1) .... 20 (2) ..... 30 (5) .... 50, 60, 80, 100, 200, 500.

Loi de F : α = 0,50 – 0,70 – 0,90 – 0,05 – 0,975 – 0,99 – 0,999 – 0,9995

[23] FISHER and YATES - Statistical Tables for biological, agricultural and médical research. Hafner, New-York.

Loi normale : f(u) et F(u) pour u = 0 (0,01) 2,99
u pour F(u) = 0 (0,01) 0,99

Loi de χ^2 : ν = 1 (1) 30

α et 1 – α = 0,99, 0,98, 0,05, 0,90, 0,80, 0,70, 0,50 et 1 – α = 0,001

Loi de F : α = 0,80 – 0,90 – 0,95 – 0,99.

[24] PEARSON and HARTLEY - Biometrika Tables for Statisticians Vol I Cambridge University Press, 1954.

Tables très étendues dans lesquelles on trouve en particulier des tables détaillées de la loi normale, de la loi de Poisson [f(x)], de la loi de χ^2 et de la loi de F.

[25] M. BOLL - Tables numériques universelles. Dunod.

Recueil de tables dans lequel on trouve en particulier les tables suivantes :

Loi normale : $f(x)$ et $F(x) - 0,500$ pour $x = 0$ (0,001) .... 2 (0,01) ..... 6.

Loi de Poisson : $f(x)$ pour $m = 0,1$ (0,1) 15

Fonction $\Gamma(x)$ pour $x = 1$ (0,001) 2

Fonction $z = \log \log x$ et $y = \mathrm{Log}_e \mathrm{Log}_e x$ pour
$x = 1,001$ (0,001) ...... 1,180 (0,002)..... 1,300 (0,005) 1,500 (0,01) 2

avec des intervalles croissants jusqu'à $x = 100\,000$ et tables des fonctions inverses $x = 10^{(10^x)}$ et $x = e^{(e^y)}$.

Table pour le calcul de $\log e^x$ pour $\pm x = 0$ (0,00001) 9,99999

Table de e^x pour $\pm x = 0$ (0,01) ..... 7 (1) ..... 50 (10) .... 100 (50) .... 500.

IV.2 - TABLES PARTICULIERES

1/ Loi binomiale

[26] Tables of the binomial probability distribution
National Bureau of Standards, Applied Mathematical Series 6, Washington

$$C_n^x p^x q^{n-x} \text{ et } \sum_{s=r}^{n} C_n^s p^s q^{n-s} \quad \text{pour} \begin{cases} n = 2\ (1)\ \dots\dots\dots\dots\ 49 \\ p = 0,01\ (0,01)\ \dots\dots\ 0,50 \\ x = 0\ (1)\ \dots\dots\dots\dots\ n-1 \\ r = 0\ (1)\ \dots\dots\dots\dots\ n \end{cases}$$

H.G. ROMIG - 50-100 Binomial Tables. John Wiley, New-York.

$$C_n^x p^x q^{n-x} \text{ et } \sum_{s=o}^{r} C_n^s p^s q^{n-s} \quad \text{pour} \begin{cases} n = 50\ (5)\ \dots\dots\dots\ 100 \\ p = 0,01\ (0,01)\ \dots\ 0,50 \\ x = 0\ (1)\ \dots\dots\dots\ n-1 \\ r = 0\ (1)\ \dots\dots\dots\ n \end{cases}$$

2/ Loi de Poisson

MOLINA. - Poisson's exponential Binomial limits. Van Nostrand, New-York.

$$e^{-a}\frac{a^x}{x!} \text{ et } \sum_{x=o}^{\infty} e^{-a}\frac{a^x}{x!} \quad \text{pour} \quad \begin{aligned} a = {} & 0,001\ (0,001)\ \dots\ 0,01 \\ & 0,01\ (0,01)\ \dots\dots\ 0,40 \\ & 0,40\ (0,01)\ \dots\dots\ 15 \\ & 15\ (1,0)\ \dots\dots\dots\ 100 \end{aligned}$$

3/ Loi normale

[28] Tables of normal probability functions. National Bureau of Standards Applied Mathematics Series 23, Washington

$f(x)$ et $F(x)$, $F(-x)$ pour $x = 0$ (0,001), ... 1 (0,001) 7,80

[29] J. BERKSON dans Journal American Statistical Association, Juin 1955
a) Valeurs de x pour $F = 0$ (0,001) 1
b) Valeurs de F pour $x = 0$ (0,001) 2,50
Voir aussi Tables numériques générales de M. Boll pour tables de $f(x)$ et $F(x)$ [25].

4/ Loi de χ^2

[30] THOMPSON - Tables of the χ^2 distribution. Biométrika XXXII, 1941-1942.

Valeurs de χ^2 pour α et $1 - \alpha$ = 0,005, 0,01, 0,025, 0,05, 0,10, 0,25, 0,50 et ν = 1 (1) 30 (10) 100

[31] KENDALL - Distribution function of χ^2 for one degree of freedom. Biometrika tables for Statisticians, Vol I.

Valeurs de 1 - F pour χ^2 = 0 (0,01) .... 1 (0,1) ...... 10.

5/ Distribution exponentielle

[32] Tables of the exponentiel function e^x. National Bureau of Standards. Applied Mathematics Series 14, Washington.

Valeurs de e^x pour x = - 2,5 (0,0001) 2,5 (0,001) 5
5 (0,01) 10 (1) 100

[33] Tables of the descending exponentiel function e^{-x}. National Bureau of Standards. A.M.S. 46

Valeurs de e^{-x} pour x = 2,5 (0,001) 10

[34] HOLTAPPEL - Tables of e^x and e^{-x}. Scientific Computing Service. Londres.

Valeurs de e^x et e^{-x} pour x = 0 (0,001) 0,999

[35] Table of natural logarithms of decimal numbers. National Bureau of Standards. A.M.S. 31

Valeurs de Log_e N pour N = 0,001 5,0

[36] Table of natural logarithms. National Bureau of Standards. A.M.S. (4 Vol.)

Log_e N pour N = 10 (1) 100 000
N = 0,00001 (0,0001) 10

Voir aussi Tables numériques générales de M. Boll pour tables de e^x, e^{-x}, $\log_{10} e^x$, $\log_{10} e^{-x}$ [25].

6/ Distribution Gamma

[37] J. BROWNLEE - Tables of log $\Gamma(x)$. Tracts for Computers n° IX. Cambridge University Press.

Valeurs de log $\Gamma(x)$ pour x = 1 (0,01) 50

[38] K. PEARSON - Tables of the incomplete Gamma function

Valeurs de $I(u, p) = \frac{1}{\Gamma(p + 1)} \int_0^u e^{-x} x^p \, dx$, $\left(u = \frac{x}{\sqrt{p + 1}}\right)$

pour p = 0,1 (0,1) 5 (0,2) 50
et p = - 1 (0,05) .... 0, u par intervalles de 0,1.

[38] bis L. HARTER - New tables of the incomplete Gamma function ratio and percentage points of the Chi-Square and Beta distribution (U.S. Government Printing Office).

Valeurs de $I(u, p) = \frac{1}{\Gamma(p + 1)} \int_0^{u\sqrt{p+1}} e^{-\nu} v^p \, dv$ pour u par intervalles de 0,01 et p = - 0,5 (0,5) ..... 74 (1) ...... 164.

Centiles de $\chi^2(\nu)$ pour ν = 1 (1) 150 (2) 300

Voir aussi Tables numériques générales de M. Boll pour tables de $\Gamma(x)$ pour x = 1 (0,001) ..... 2 [25].

[38] K. PEARSON - Tables of the incomplete Beta Function. Valeurs de ter $I_x(p, q) = \frac{B_{x(p,q)}}{B_{(p\,q)}}$, pour p, q = 0,5 (0,5) ... 10 (1) ... 50 et x par intervalles de 0,1.

7/ Distribution de Weibull

[39] A. PLAIT - The Weibull Distribution. Industrial Quality Control. Nov. 1962

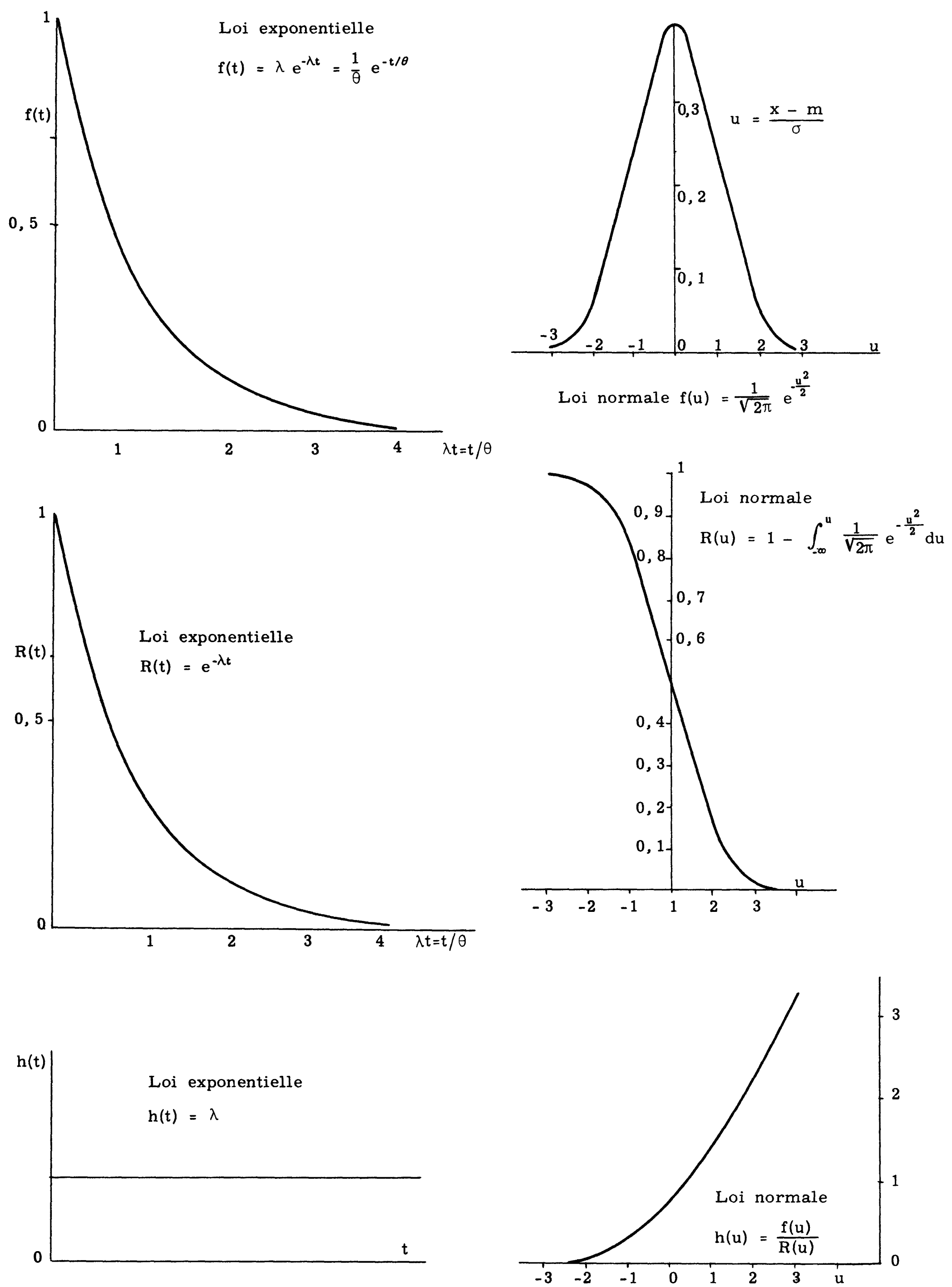
Tables de $f(x) = \beta\ x^{\beta-1}\ e^{-x^{\beta}}$ et $F(x) = \int_0^x f(x)\ dx$

pour x = 0,10 (0,10) 2,5
β = 0,10 (0,10) 3 (1) 10

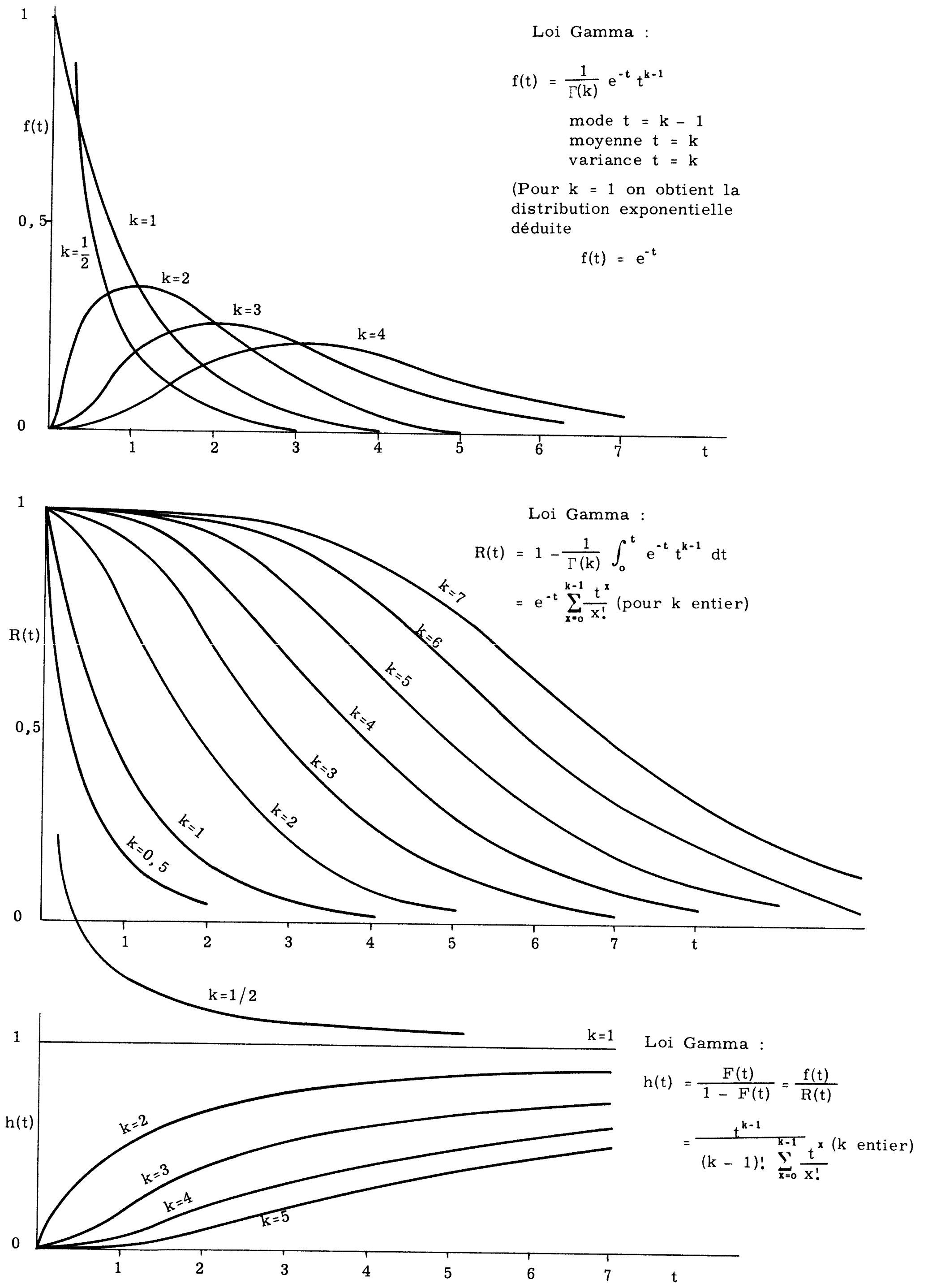
[40] Voir aussi Tables numériques générales de M. Boll pour tables de $\mathrm{Log}_e\ \mathrm{Log}_e\ x$ et de $\Gamma(x)$

[41] Distribution des Valeurs extrêmes

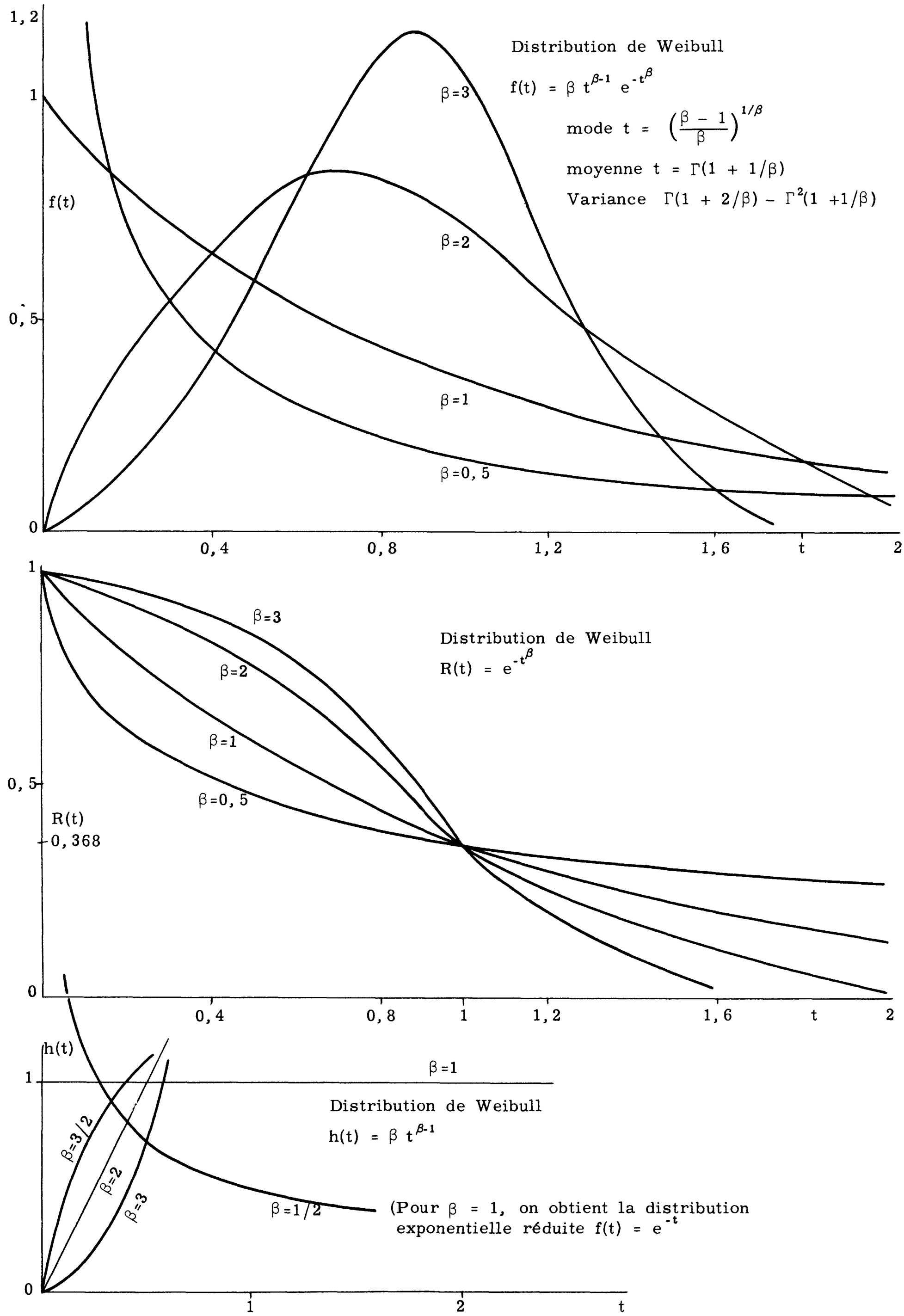
GUMBEL - Probability tables for the analysis of extrême values data. National Bureau of Standards. A.M.S. n° 22.


Loi exponentielle
f(t) = λ e^-λt = 1/θ e^-t/θ
f(t)
1
0,5
0
1 2 3 4 λt=t/θ
u = (x − m)/σ
0,3
0,2
0,1
-3 -2 -1 0 1 2 3 u
Loi normale f(u) = 1/√2π e^(-u²/2)
Loi exponentielle
R(t) = e^-λt
R(t)
1
0,5
0
1 2 3 4 λt=t/θ
Loi normale
R(u) = 1 − ∫_-∞^u 1/√2π e^(-u²/2) du
1
0,9
0,8
0,7
0,6
0,4
0,3
0,2
0,1
-3 -2 -1 1 2 3 u
Loi exponentielle
h(t) = λ
h(t)
0
t
Loi normale
h(u) = f(u)/R(u)
3
2
1
0
-3 -2 -1 0 1 2 3 u

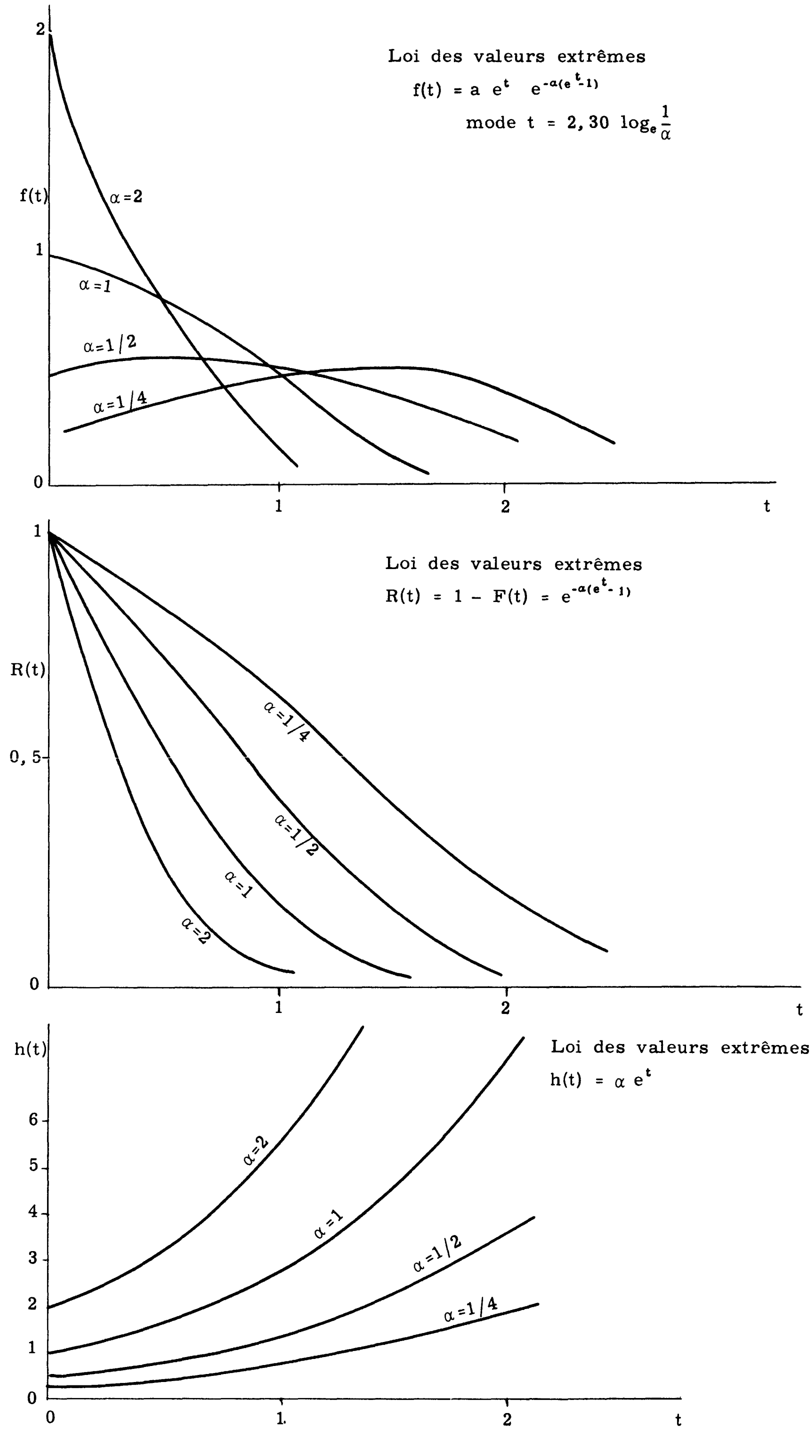
Graphique I



Graphique II



Graphique III



Graphique IV