REVUE DE STATISTIQUE APPLIQUÉE

DAVID SROUR

Sur quelques distributions de probabilité applicables aux problèmes de sécurité routière

Revue de statistique appliquée, tome 19, nº 2 (1971), p. 77-94
<http://www.numdam.org/item?id=RSA_1971__19_2_77_0>

SUR QUELQUES DISTRIBUTIONS DE PROBABILITÉ APPLICABLES AUX PROBLÈMES DE SÉCURITÉ ROUTIÈRE

David SROUR
SEMA-Prévision

INTRODUCTION

Nous commencerons par dire quelques mots des problèmes de Sécurité Routière tels qu'ils se posent régulièrement en pratique et de la façon dont ces problèmes sont actuellement résolus. Nous proposerons ensuite certaines formalisations probabilistes des phénomènes qui peuvent aider à fournir des moyens d'études plus élaborés.

EFFICACITE D'UNE MESURE DE PREVENTION

Une mesure de prévention étant décidée en un certain lieu L et pendant une certaine période T, il faut pour mesurer son effet étudier non seulement la réduction du nombre d'accidents, mais aussi la diminution éventuelle de la gravité des accidents (par exemple : nombres moyens de tués et de blessés graves par accident).

L'étude du nombre d'accidents ne pose pas de problèmes. On sait que ce nombre suit une loi de Poisson de moyenne $\lambda(L , T)$ que l'on peut estimer à l'aide des observations d'accidents pendant des périodes successives de longueur T, choisies avant l'entrée en vigueur de la réglementation. On effectue, alors, un test classique pour voir si la moyenne d'accidents a significativement diminué [Réf. 1, 2].

En fixant une condition sur la puissance du test, on peut aussi aboutir à un seuil de réduction significative, indispensable pour une étude coûts-bénéfices de la mesure en vigueur [Réf. 2].

Dans cette dernière étude, on tient, par contre, rarement compte de l'évolution éventuelle de la gravité de l'accident, cela parce qu'on ne possède pas explicitement la distribution du nombre de tués ou de blessés par accident. On se contente, en général, de vérifier grâce à un test d'homogénéité que la répartition des impliqués en tués, blessés et indemnes "n'a pas trop varié" ; on en déduit alors :

- d'abord que l'on peut se contenter du bilan "accidents" pour l'étude "coûts-bénéfices",

- ensuite que l'efficacité de la mesure se traduit par une réduction des victimes certes, mais une réduction seulement conséquente à la diminution d'accidents [Réf. 3].

On comprend sans peine les insuffisances d'une telle position :

- si le test d'homogénéité conduit à la conclusion "répartition inchangée", il n'y a pas forcément incomptabilité avec la conclusion "décroissance significative du nombre de tués ou de blessés par accident", ce dernier résultat éventuel étant obtenu de façon directe à l'aide d'un test explicite sur le nombre de tués ou de blessés par accident,

- si le test d'homogénéité conduit à la conclusion "changement dans la répartition des impliqués en tués, blessés et indemnes", il faut alors tenir explicitement compte de ce changement dans l'étude "coûts-bénéfices".

Or, le test effectué ne permet que des conclusions qualitatives, c'est-à-dire que l'on ne possédera que rarement les nouveaux pourcentages de répartition statistiquement valables.

Afin de lever cette indétermination les paragraphes 1, 2, 3 proposeront certaines distributions de probabilité, en particulier pour les nombres d'impliqués, de tués et de blessés par accident.

EFFICACITE D'UN AMENAGEMENT ROUTIER LOCALISE

Si une mesure globale de prévention (telle une limitation de la vitesse par exemple) est, en général, soumise à l'appréciation du public à qui il faut fournir explicitement des réductions des nombres de tués et de blessés par accident, un aménagement routier, qui est une mesure de prévention localisée, conserve un caractère plus technique. La notion de coût moyen d'un accident (coût direct auquel on adjoint éventuellement un coût social) se substitue alors à celle du couple "nombres moyens de tués ou de blessés".

Fondamentalement, le problème n'a cependant pas changé et les études de sécurité routière se contentent souvent, à tort ou à raison, de baser le bilan économique sur la décroissance "avant-après" du nombre d'accidents seulement [Réf. 4].

Le paragraphe 4 propose des distributions de probabilité permettant de faire intervenir explicitement les coûts des accidents.

Afin de faciliter la lecture du texte, on donne ci-dessous la liste des problèmes qui seront successivement étudiés dans les quatre paragraphes de cet article.

Problèmes étudiés	Paramètres dont dépendent les lois
1 - Loi du nombre d'impliqués I par accident (I = 1, 2, ...)	μ = paramètre spécifique p = proportion d'accidents impliquant un seul véhicule
2 - Loi du nombre de blessés B (ou de tués) par accident, ou pour n accidents (n connu) (B = 0, 1, 2, ...)	
2.a - Cas d'un accident	μ, p = comme ci-dessus π = probabilité d'être blessé (tué) dans un accident, supposée indépendante du nombre d'impliqués

2.b - Cas de n (connu) accidents	μ, p, π = comme ci-dessus n = nombre d'accidents
3 - Loi du nombre de blessés B (ou de tués) en fonction de la dangérosité (nombre d'accidents aléatoire) (B = 0, 1, 2, ...)	
3.a - Lieu de dangérosité connue	μ, p, π = comme ci-dessus λ = paramètre de dangérosité
3.b - Lieu de dangérosité quelconque = λ aléatoire, de loi γ (m , c)	μ, p, π = comme ci-dessus m, c = paramètre de la loi de λ
3.c - Lieu de dangérosité quelconque = λ aléatoire, de loi N^+ (m , σ) total	μ, p, π = comme ci-dessus m, σ = paramètres de la loi de λ
4 - Loi du coût total S des accidents en fonction de la dangérosité (nombre d'accidents aléatoire) ($0 \leq S \leq \infty$)	
4.a - Lieu de dangérosité connue	μ = paramètre de la loi du coût λ = paramètre de dangérosité
4.b - Lieu de dangérosité quelconque = λ aléatoire, de loi γ (m , c)	μ = comme ci-dessus m , c = paramètre de la loi de λ
4.c - Lieu de dangérosité quelconque = λ aléaoite, de loi N^+ (m , σ)	μ = comme ci-dessus m , σ = paramètre de la loi de λ

1 - NOMBRE D'IMPLIQUES PAR ACCIDENT

On peut admettre que le nombre d'impliqués par accident, soit I, suit une loi de Poisson de moyenne μ, tronquée de la valeur O ou des valeurs O ou 1 selon que l'accident implique un véhicule unique ou 2 véhicules au moins en mouvement.

Si l'on appelle p la probabilité qu'un accident implique un véhicule seulement, les probabilités attachées à I sont les suivantes :

$$\begin{cases} P(I = 0) = 0 \\ P(I = 1) = \dfrac{\mu p e^{-\mu}}{1 - e^{-\mu}} = \dfrac{\mu p}{e^{\mu} - 1} \\ P(I = x) = \dfrac{p e^{-\mu} \mu^x}{x!\,(1 - e^{-\mu})} + \dfrac{(1 - p)\, e^{-\mu}\, \mu^x}{x!\,(1 - e^{-\mu} - \mu e^{-\mu})} \qquad x = 2, 3, \ldots \end{cases}$$

soit après calcul et en posant $K = K(p, \mu) = \dfrac{(e^{\mu} - 1 - p\mu)}{(e^{\mu} - 1)\,(e^{\mu} - 1 - \mu)}$

$$
(1.1) \qquad P(I = x) = \begin{cases} 0 & \text{si } x = 0 \\ \dfrac{p\mu}{e^{\mu} - 1} & \text{si } x = 1 \\ \dfrac{K\mu}{x!} & \text{si } x = 2, 3, \dots \end{cases}
$$

On note de plus les caractéristiques de la distribution (1.1)

$$(1.2) \qquad E(I) = \mu \left[\frac{p}{e^{\mu} - 1} + K(e^{\mu} - 1)\right]$$

$$(1.3) \qquad E[I(I - 1)] = K\mu^2 e^{\mu}$$

$$(1.4) \qquad E(I^2) = \mu \left[\frac{p}{e^{\mu} - 1} + K(e^{\mu} - 1 + \mu e^{\mu})\right]$$

$$(1.5) \qquad V(I) = \mu\{(P + K')(1 - \mu(P + K')) + K\mu e\}$$

$$\text{avec } P = \frac{p}{e^{\mu} - 1} \; ; \; K' = K(e^{\mu} - 1)$$

2 - NOMBRE DE BLESSES PAR ACCIDENT (S)

La méthode utilisée ci-après serait exactement la même pour le nombre de tués par exemple.

a) Cas d'un seul accident

Si le nombre d'impliqués est connu, soit x par exemple, la variable B représentant le nombre de blessés par accident suit une loi binomiale $\mathcal{B}(x, \pi)$ où π, probabilité d'être blessé dans un accident, est supposée être indépendante du nombre d'impliqués

$$(2.1) \qquad P(B = y/I = x) = C_x^y \pi^y (1 - \pi)^{x-y}$$

On peut à l'aide de cela, déduire les probabilités $P(B = y)$.

Pour $y \geq 2$, ces probabilités sont données par :

$$P(B = y) = \sum_{x=y}^{\infty} P(B = y/I = x) \, P(I = x) \text{ avec } P(I = x) = K\mu^x/x!$$

puisqu'il faut qu'il y ait au moins deux impliqués par accident. D'où :

$$P(B = y) = K \sum_{x=y}^{\infty} C_x^y \pi^y (1 - \pi)^{x-y} \mu^x/x! \quad \text{et après calcul :}$$

$$(2.2) \qquad P(B = y) = (Ke^{\mu}) \, e^{-\mu\pi} (\mu\pi)^y/y! \qquad \text{pour } y = 2, 3, \dots$$

Pour les valeurs 0 et 1 de y, il faut aussi tenir compte du cas où le nombre d'impliqués est égal à 1. Pour y = 0 par exemple,

$$P(B = 0) = P(B = 0/I = 1)\ P(I = 1) + \sum_{x=2}^{\infty} P(B = 0/I = x)\ P(I = x)$$

$$= \frac{\mu p(1-\pi)}{(e^{\mu}-1)} + K \sum_{x=2}^{\infty} \frac{\mu\ (1-\pi)^x}{x!}$$

$$= \frac{\mu p(1-\pi)}{(e^{\mu}-1)} + K(e^{\mu(1-\pi)} - 1 - \mu(1-\pi)$$

La même procédure permet de calculer P(B = 1). D'où finalement :

$$(2.3)\quad P(B = y) = \begin{cases} \dfrac{\mu p(1-\pi)}{e^{\mu}-1} + K(e^{\mu(1-\pi)} - 1 - \mu(1-\pi)) & \text{pour } y = 0 \\ \dfrac{\mu p \pi}{e^{\mu}-1} + K\pi\mu\,(e^{\mu(1-\pi)} - 1) & \text{"}\quad y = 1 \\ Ke^{\mu(1-\pi)}\ (\mu\pi)^y/y! & \text{"}\quad y = 2, 3, \ldots \end{cases}$$

résultat qui constitue en quelque sorte une généralisation de l'étude de la somme à nombre de termes poissonien, de variables binomiales parentes de $\mathcal{B}(1, \pi)$ [Réf 5, p. 71, 76].

On peut facilement vérifier la relation $\sum_{y=0}^{\infty} P(B = y) = 1$.

En remarquant après quelques calculs que P (B = 1) s'écrit aussi $\mu\pi\left(Ke^{\mu(1-\pi)} - \dfrac{q}{e^{\mu} - 1 - \mu}\right)$ avec $q = 1 - p$, on trouve pour la distribution (2.3) les caractéristiques suivantes :

$$(2.4)\qquad E(B) = \pi\mu\left(Ke^{\mu} - \frac{q}{e^{\mu} - 1 - \mu}\right)$$

$$(2.5)\qquad E[B(B-1)] = K(\pi\mu)^2 e^{\mu}$$

$$(2.6)\qquad E(B^2) = K\pi p e^{\mu}\ (1 + \pi\mu) - \frac{q\pi\mu}{e^{\mu} - 1 - \mu}$$

$$(2.7)\qquad V(B) = \pi\mu\,[(K'' - Q)\ (1 - \pi\mu\ (K'' - Q)) + \pi\mu K'']$$

$$\text{où } K'' = Ke^{\mu}\ ;\ Q = \frac{q}{e^{\mu} - 1 - \mu}$$

Application : Test sur la réduction du nombre de blessés par accident

La détermination explicite des probabilités (2.3) attachées à B permet d'effectuer un test sur la réduction éventuelle du nombre de blessés par accident, consécutive à une certaine mesure de prévention. On supposera à cet effet que les paramètres p et μ égaux respectivement à p_0 et μ_0 dans la situation "avant" n'ont pas varié "après" (ce qui est souvent le cas en pratique) et on effectuera le test H_0 : $\pi = \pi_0$ contre H : $\pi \leq \pi_0$.

Soit donc n accidents et $y_1, y_2, \ldots y_n$ les nombres correspondants de blessés. On peut classer les y_i de la façon suivante :

$y_1 = y_2 = \dots y_{n_0} = 0$ c'est-à-dire qu'il y a n_0 accidents sans aucun blessé

$y_{n_0+1} = y_{n_0+2} = \dots y_{n_0+n_1}$ c'est-à-dire qu'il y a n_1 accidents à nombre de blessés par accident égal à l'unité

$y_{n_0+n_1+i} \geq 2 \quad \forall i = 1, 2, \dots, n_2$ avec $n_2 = n - (n_0 + n_1)$ = nombre d'accidents pour lesquels le nombre de blessés est ≥ 2.

En notant $y_{n_0+n_1+i} = z_i \geq 2 \; \forall i = 1, 2, \dots n_2$, le rapport de vraisemblance s'écrit

$$\frac{L_0}{L} = \left(\frac{P_{00}}{P_0}\right)^{n_0} \left(\frac{P_{10}}{P_1}\right)^{n_1} \prod_{i=1}^{n_2} \frac{P_{20}(z_i)}{P_2(z_i)} \tag{2.8}$$

avec (cf. 2.3)

$$P_0 = P_0(p_0, \mu_0, \pi) = \frac{\mu_0 p_0(1-\pi)}{e^{\mu_0} - 1} + K(p_0, \mu_0)\,(e^{\mu_0(1-\pi)} - 1 - \mu_0(1-\pi))$$

$$\text{et } P_{00} = P_0(p_0, \mu_0, \pi_0)$$

(2.9)

$$P_1 = P_1(p_0, \mu_0, \pi) = \frac{\mu_0 p_0 \pi}{e^{\mu_0} - 1} + K(p_0, \mu_0)\mu_0\pi(e^{\mu_0(1-\pi)} - 1)$$

$$\text{et } P_{10} = P_1(p_0, \mu_0, \pi_0)$$

$$P_2(z_i) = P_2(p_0, \mu_0, \pi, z_i) = K(p_0, \mu_0)e^{\mu_0}e^{-\mu_0\pi}(\mu_0\pi)^{z_i}/z_i!$$

$$\text{et } P_{20}(z_i) = P_2(p_0, \mu_0, \pi_0, z_i)$$

On peut alors aisément montrer que la région critique W est donnée par $\sum_{i=1}^{n_2} z_i < k$ où k est une constante à déterminer, une fois l'erreur de première espèce α fixée. Bien entendu, il faut préalablement déterminer la loi de $\sum_{i=1}^{n_2} z_i$, somme de n_2 variables indépendantes, de fonction génératrice :

$$g_{z_i}(t) = \sum_{z_i=2}^{\infty} Ke^{\mu(1-\pi)}(\mu\pi t)^{z_i}/z_i! = Ke^{\mu(1-\pi)}(e^{\mu\pi t} - 1 - \mu\pi t) \tag{2.10}$$

L'expression (2.11) de $g_{\sum_1^{n_2} z_i}(t)$:

$$g_{\Sigma z_i}(t) = [Ke^{\mu(1-\pi)}]^{n_2}[e^{\mu\pi t} - 1 - \mu\pi t]^{n_2} \text{ ou encore} \tag{2.11}$$

$$g_{\Sigma z_i}(t) = \left[\frac{K\mu^2 t^2 e^{\mu(1-\pi)}}{2}\right]^{n_2} \left(\sum_{j=0}^{\infty} \frac{(\mu\pi t)^j}{(j+2)^{[j]}}\right)^{n_2} \tag{2.11'}$$

permet de déduire les probabilités attachées à $\sum_1^{n_2} z_i$ dans l'hypothèse H_0 et par conséquent le seuil critique k.

Bien entendu, le test sur la réduction du nombre de tués par accident serait essentiellement le même, le paramètre π de blessure devant simplement être remplacé par un paramètre ψ de décès.

b) Cas de n accidents ($n \geq 2$)

En appelant B_1, B_2, ..., B_n les nombres de blessés relatifs aux n accidents et en admettant l'hypothèse d'indépendance des B_i, on a :

(2.12) $$g_{\sum_{i=1}^{n} B_i}(t) = [g_B(t)]^n$$

où $g_X(t)$ désigne la fonction génératrice de la variable aléatoire X.

D'où $g_{\Sigma B_i}(t) = \left[P_0 + tP_1 + \sum_{y=2}^{\infty} P_2(t)\, t^y\right]$ où P_0, P_1 et $P_2(y)$ ont des significations évidentes obtenues en se reportant à (2.3).

On trouve finalement

(2.13) $$g_{\Sigma B_i}(t) = [Q_0 + tQ_1 + Q_2 e^{\pi\mu t}]^n$$

avec $Q_0 = P_0 - Ke^{\mu(1-\pi)}$; $Q_1 = P_1 - \pi\mu Ke^{\mu(1-\pi)} = \dfrac{-\pi\mu q}{e^{\mu} - 1 - \mu}$; $Q_2 = Ke^{\mu(1-\pi)}$

On peut aussi écrire (2.13) sous la forme

(2.14) $$g_{\Sigma B_i}(t) = \left[\sum_{j=0}^{\infty} R_j t^j\right]^n$$

avec $R_0 = Q_0 + Q_2 = P_0$; $R_1 = Q_1 + \pi\mu Q_2 = P_1$; $R_j = Q_2 \pi^j \mu^j / j!$ $j = 2, 3, \ldots$

La relation (2.13) ou (2.14) permet de déterminer plus explicitement les probabilités attachées à $\sum_{i=1}^{n} B_i$.

3 - NOMBRE DE BLESSES QUAND LE NOMBRE N D'ACCIDENTS EST ALEATOIRE

a) Lieu de dangérosité connue λ

Si l'on s'intéresse à un certain lieu (section de route par exemple) dont la "dangérosité" (ou indice caractérisant le danger) est λ, la fonction générative de $S = \sum_{i=1}^{n} B_i$ dans l'hypothèse où N est un aléa poissonien, est donnée par :

(3.1) $$g_S(t) = e^{\lambda\left[\left(\sum_{j=0}^{\infty} R_j t^j\right)^n - 1\right]}$$

On sait en effet que si $S = B_1 + B_2 + \ldots + B_N$ où les B_i sont des variables indépendantes de même loi et N aléatoire, on a alors :

(3.2) $$g_S(t) = g_N(g_B(t)) \qquad \text{[Réf. 6]}$$

On sait de plus que la propriété bien connue des fonctions génératrices :

(3.3) $$g_X^{(h)}(1) = E(X^{[h]})$$

(où $g_X^{(h)}(1)$ représente la valeur de la dérivée d'ordre h au point 1) permet d'aboutir aux relations classiques :

(3.4) $$E(S) = E(N) \,.\, E(B)$$

(3.5) $$V(S) = E(N) \,.\, V(B) + V(N) \,.\, E^2(B) \qquad \text{[Réf. 6]}$$

Dans le cas particulier où $E(N) = V(N) = \lambda$ (loi de Poisson), (3.5) s'écrit d'ailleurs encore :

(3.5)' $$V(S) = E(N) \,.\, E(B^2)$$

On trouve alors facilement : (cf. 2.4 et 2.6)

(3.6) $$E(S) = \lambda\pi\mu\left[Ke^{\mu} - \frac{q}{e^{\mu} - 1 - \mu}\right]$$

(3.7) $$V(S) = \lambda\pi\mu\left[Ke^{\mu}(1 + \pi\mu) - \frac{q}{e^{\mu} - 1 - \mu}\right]$$

Il arrive aussi souvent que l'on s'intéresse au nombre total de blessés survenu en une section de route quelconque (c'est-à-dire dont la dangérosité λ est inconnue).

b) Cas où λ suit la loi $\gamma(m, c)$

Il est courant d'admettre que sur l'ensemble des sections de route , qui est maintenant aléatoire, est distribué selon la loi $\gamma(m, c)$: Réf. 7, 8

(3.8) $$f_{\lambda}(u) = \frac{e^{-cu}c^{m}u^{m-1}}{(m-1)!}$$

Le résultat est la loi binomiale négative pour N :

(3.9) $$P(N = n) = \frac{(m + n - 1)!}{(m-1)!\, n!}\,\frac{c}{(1 + c)^{n+m}} \qquad n = 0, 1, 2, \ldots \text{ [Réf. 6, 7]}$$

loi dont la fonction génératrice est :

(3.10) $$g_N(t) = \left[1 + \frac{1}{c}(1 - t)\right]^{-m}$$

et dont les caractéristiques sont :

(3.11) $$E(N) = \frac{m}{c},$$ (3.12) $$V(N) = \frac{m}{c}\left(1 + \frac{1}{c}\right) \qquad \text{[Réf. 6]}$$

Les relations (3.1), (3.6) et (3.7) deviennent alors :

(3.13) $$g_S(t) = \left[1 + \frac{1}{c}\left(1 - \left(\sum_{j=0}^{\infty} R_j t^j\right)^n\right)\right]^{-m}$$

(3.14) $$E(S) = \frac{m\pi\mu}{c}\left(Ke^{\mu} - \frac{q}{e^{\mu} - 1 - \mu}\right)$$

(3.15) $$V(S) = \frac{m\pi\mu}{c}\left[(K'' - Q)\left(1 + \frac{\pi\mu}{c}(K'' - Q)\right) + \pi\mu K''\right]$$

où K'' et Q sont les mêmes que dans (2.7).

c) Cas où λ suit la loi $N^+(m, \sigma)$

Enfin, dans le cas courant en pratique où λ est distribué selon une loi normale, $N^+(m, \sigma)$, restreinte à l'intervalle $[0, \infty]$, on peut démontrer (cf. plus loin la relation (4.28)) :

$$(3.16) \qquad g_N(t) = \frac{F\left(\frac{m}{\sigma} + \sigma(t-1)\right)}{F\left(\frac{m}{\sigma}\right)} \exp \left\{ (t-1) \left[m + (t-1) \frac{\sigma^2}{2} \right] \right\} \quad \text{où}$$

$$(3.17) \qquad F(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-x^2/2} dx$$

est la fonction de répartition de la variable $N(0, 1)$, variable normale centrée réduite.

De même, on donne plus loin $E(N)$ et $V(N)$, ce qui permet de déduire, dans le cadre de cette hypothèse, les formules homologues de (3.13), (3.14) et (3.15). Ces formules, très complexes, ne sont pas explicitement fournies ici.

4 - ETUDE DU COUT TOTAL DES ACCIDENTS

a) Lieu de dangérosité connue λ

Il est courant que le coût d'un accident suive une distribution exponentielle [Réf. 9.10]. On considère alors la variable aléatoire "coût global" $S_N = C_1 + C_2 + \ldots + C_N$ où N, nombre d'accidents, suit une loi $P(\lambda)$ et C_i, coût d'un accident, une loi $\gamma\left(1, \frac{1}{\mu}\right)$ de densité $\frac{1}{\mu} e^{-x/\mu}$ $\forall$ i = 1, 2, ..., N. On admet S = 0 si N = 0.

Les probabilités attachées à S_N s'écrivent alors :

$$P(S_N = x) = \sum_{n=1}^{\infty} P(S_n = x) \; P(N = n) \qquad \text{pour } x > 0.$$

En tenant compte du fait que S_n a pour densité $f_{S_n}(x) = \frac{e^{-x/\mu} x^{n-1}}{(n-1)! \mu^n}$, on trouve la densité de S_N

$$(4.1) \qquad f_{S_N}(x) = \begin{cases} e^{-\lambda} & \text{pour } x = 0 \\ e^{-\lambda - x/\mu} \displaystyle\sum_{n=1}^{\infty} \frac{(\lambda/\mu)^n x^{n-1}}{n! \; (n-1)!} & \text{pour } x > 0 \end{cases}$$

Cette densité s'écrit encore :

$$(4.2) \qquad f_{S_N}(x) = \begin{cases} e^{-\lambda} & \text{pour } x = 0 \\ \dfrac{\lambda}{\mu} e^{-\lambda - x/\mu} \displaystyle\sum_{k=0}^{\infty} \frac{(\lambda x/\mu)^k}{k! \; (k+1)!} = \sqrt{\frac{\lambda}{\mu x}} \, e^{-\lambda - x/\mu} \, I_1\left(2\sqrt{\frac{\lambda x}{\mu}}\right) & \text{où} \end{cases}$$

$$(4.3) \qquad I_1(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+1}}{k! \; (k+1)!}$$

est la fonction de Bessel modifiée d'ordre 1.

Il est évidemment facile de vérifier que l'on a :

$$\int_0^\infty f_{S_N}(x)\ dx = 1 - e^{-\lambda} \tag{4.4}$$

L'application des formules (3.4) et (3.5) permet d'obtenir les caractéristiques de cette loi

$$E(S_N) = \lambda\mu \tag{4.5}$$

$$E(S_N^2) = \lambda\mu^2\ (2 + \lambda) \tag{4.6}$$

$$V(S_N) = 2\lambda\mu^2 \tag{4.7}$$

Dans le cadre de ce modèle, on trouve pour le coefficient de variation du coût global C.V. (S_N) la valeur

$$\text{C.V.}\ (S_N) = \sqrt{\frac{2}{\lambda}} \tag{4.8}$$

La relation C.V. (S_N) = $\sqrt{2}$ C.V. (N) montre que le coefficient de variation du coût global est plus élevé que celui du nombre d'accidents, ce qui est tout à fait normal [Réf. 11]. Mais est-ce une raison suffisante, comme plusieurs auteurs le pensent, pour se dispenser d'une étude de la variation des coûts d'accidents après un aménagement de sécurité et de se restreindre simplement à une étude du nombre d'accidents ? Certainement pas : on sait en effet combien le coût d'une mauvaise décision, basée seulement sur une étude des fréquences d'accidents, est élevé pour la collectivité.

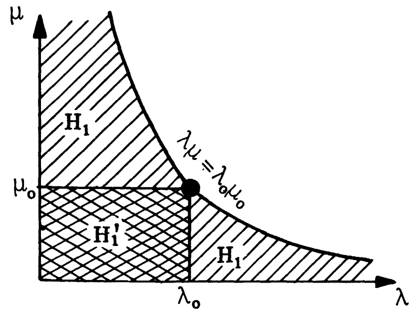
Application : Test sur la réduction du coût global des accidents

La distribution donnée en 4.2 permet, dans la mesure où elle cadre avec la réalité, de poser clairement le problème du test de sécurité : à l'aide des observations coûts d'accidents en nombre aléatoire, c'est-à-dire d'une observation "coût global d'accidents" tester la situation

"avant" $H_0 \begin{cases} \lambda = \lambda_0 \\ \mu = \mu_0 \end{cases}$ contre la situation favorable "après" la plus générale

$H_1 \begin{cases} \lambda \text{ et } \mu \\ \text{tels que } \lambda\mu < \lambda_0\mu_0 \end{cases}$

Formulé ainsi, ce test apparaît bien plus précis que celui qui ne tiendrait pas compte de l'évolution du coût des accidents ($\lambda = \lambda_0$ contre $\lambda < \lambda_0$ avec bruit constant μ_0 dans les deux hypothèses). Il est en outre moins sévère que le test qui n'admettrait comme situation favorable qu'une situation pour laquelle on aurait $\lambda < \lambda_0$ et $\mu < \mu_0$ simultanément, dans l'hypothèse H_1. Ce dernier test serait d'ailleurs effectué d'abord sur λ à l'aide de la distribution poissonienne du nombre d'accidents, et serait suivi dans le cas où l'on rejette $\lambda = \lambda_0$ d'un test sur μ à l'aide de la loi du coût d'un accident. A titre d'illustration on donne ci-joint les deux domaines définissant l'hypothèse alternative : le domaine hachuré H_1 que nous proposons, limité par les deux axes et l'hyperbole $\lambda\mu = \lambda_0\mu_0$ et le rectangle doublement hachuré H_1', souvent choisi en pratique. L'hypothèse H_0 est représentée par le point (λ_0, μ_0).



Le modèle du paragraphe 4. a) peut être étendu au cas où le paramètre λ de Poisson est aléatoire. Conformément à ce qui précède, l'étude en a été faite dans le cadre des lois $\gamma(m, c)$ (cf. 3.8) et $N^+(m, \sigma)$ pour λ.

b) Lieu de dangérosité λ quelconque : λ suivant la loi $\gamma(m, c)$

La démarche du calcul est pareille à 4a) sauf que N suit la loi binomiale négative rappelée en (3.9).

On trouve alors facilement :

$$(4.9) \quad f_{S_N}(x) = \begin{cases} \left(\dfrac{c}{1+c}\right)^m & \text{pour } x = 0 \\[2ex] \dfrac{e^{-x/\mu}\, c^m}{\mu(m-1)!\,(1+c)^{m+1}} \displaystyle\sum_{k=0}^{\infty} \dfrac{\left(\dfrac{x}{\mu(1+c)}\right)^k (m+k)!}{k!\,(k+1)!} & \text{pour } x > 0 \end{cases}$$

Compte tenu de la relation

$$(4.10) \qquad E^2(C_i) = V(C_i) = \mu^2 \qquad \forall i = 1, 2, \ldots$$

et des relations (3.11), (3.12), l'application des formules (3.4), (3.5) conduit ici à :

$$(4.11) \qquad E(S_N) = \frac{m\mu}{c}$$

$$(4.12) \qquad V(S_N) = m\,\frac{\mu^2}{c^2}\,(2c+1)$$

c) Lieu de dangérosité λ quelconque : λ suivant la loi $N^+(m, \sigma)$

De même qu'en 4b), commençons par trouver la loi marginale de N quand λ est distribué selon $N^+(m, \sigma)$. La densité de cette dernière distribution s'écrit :

$$f_\lambda(x) = \frac{k}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x-m)^2\right]$$

où la constante k est donnée par $\int_0^\infty f_\lambda(x)\, dx = 1$. On trouve après calcul

$$k = \frac{2}{1 + \theta\left(\frac{m}{\sigma\sqrt{2}}\right)} = \frac{1}{F\left(\frac{m}{\sigma}\right)} \tag{4.13}$$

où $\theta(t)$ est la fonction-erreur tabulée, souvent notée erf(t) :

$$\theta(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-y^2} dy \tag{4.14}$$

et F(t), fonction de répartition de N(0 , 1) est donnée par (3.17) [Réf. 12].

Si le paramètre λ de Poisson suit la loi ainsi déterminée, la probabilité d'observer n accidents pendant une certaine période, en un lieu quelconque est donnée par :

$$P(N = n) = \int_0^\infty P(N = n/\lambda = u)\, f_\lambda(u)\, du, \text{ soit}$$

$$P(N = n) = \frac{J_n}{n\,!\ \sigma\sqrt{2\pi}\, F(m/\sigma)} \text{ en notant } J_n \text{ l'intégrale } \int_0^\infty u^n e^{-\left(u + \frac{(u-m)^2}{2\sigma^2}\right)} du \tag{4.15}$$

En remarquant la relation de récurrence obtenue après intégration par parties de J_{n-2} :

$$J_n = (m - \sigma^2)\, J_{n-1} + \sigma^2(n - 1)\, J_{n-2} \text{ [Réf. 13] ,} \tag{4.16}$$

on peut de proche en proche, calculer J_n grâce aux valeurs initiales obtenues après un calcul qui ne présente guère de difficultés majeures

$$J_0 = \sigma\sqrt{2\pi}\ e^{-m+\sigma^2/2} F\left(\frac{m}{\sigma} - \sigma\right) \tag{4.17}$$

$$J_1 = \sigma\sqrt{2\pi}\ e^{-m+\sigma^2/2} (m - \sigma^2)\, F\left(\frac{m}{\sigma} - \sigma\right) + \sigma^2 e^{-m^2/2\sigma^2} \text{[Réf. 7 , 13]} \tag{4.18}$$

Cela n'est d'ailleurs pas nécessaire si l'on s'intéresse au coût global des accidents, car les sommations peuvent s'effectuer sans nécessiter le calcul explicite de J_n.

On a en effet :

$$P(S_N = 0) = P(N = 0) = \frac{e^{-m+\sigma^2/2} F\left(\frac{m}{\sigma} - \sigma\right)}{F\left(\frac{m}{\sigma}\right)} \quad \text{d'après (4.15) et (4.17)}$$

$$P(S_N = x) = \sum_{n=1}^{\infty} P(S_n = x)\ P(N = n)$$

En remplaçant dans cette dernière expression de $P(S_N = x)$, $P(N = n)$ par (4.15) et $P(S_n = x)$ par $\frac{e^{-x/\mu} x^{n-1} dx}{(n-1)\,!\ \mu^n}$, on trouve après calcul

$$f_{S_N}(x) = \frac{e^{-x/\mu}}{\sigma\sqrt{2\pi x\mu}\ F\left(\frac{m}{\sigma}\right)} \int_0^\infty I_1\left(2\sqrt{\frac{xu}{\mu}}\right) \sqrt{u}\, e^{-\left(u + \frac{(u-m)^2}{2\sigma^2}\right)} du \tag{4.19}$$

pour x > 0 où $I_1\left(2\sqrt{\frac{xu}{\mu}}\right)$ est donnée par une relation analogue à (4.3).

En développant $I_1\left(2\sqrt{\frac{xu}{\mu}}\right)$, on peut aussi écrire (4.19) sous la forme

(4.19)' $$f_{S_N}(x) = \frac{e^{-x/\mu}}{\sigma_\mu \sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \sum \frac{(x/\mu)\ J_{k+1}}{k!\ (k+1)!}$$

Vu la complexité des calculs, il est plus que jamais utile de vérifier que la somme des probabilités est bien unitaire, c'est-à-dire que l'on a

$$\int_0^\infty f_{S_N}(x)\,dx = 1 - \frac{e^{-m+\sigma^2/2} F(m/\sigma - \sigma)}{F\left(\frac{m}{\sigma}\right)}$$

Dans la suite, on montre donc que l'on a :

(4.20) $$\sigma\sqrt{2\pi}\left[F\left(\frac{m}{\sigma}\right) - e^{-m+\sigma^2/2} F\left(\frac{m}{\sigma} - \sigma\right)\right] = \int_0^\infty\int_0^\infty e^{-x/\mu} \sqrt{\frac{u}{x\mu}}\, I_1\left(2\sqrt{\frac{xu}{\mu}}\right) e^{-\left(u + \frac{(u-m)^2}{2\sigma^2}\right)} du\, dx$$

Pour cela, on commence par intégrer le second membre par rapport à x : l'intégration se fait par parties en tenant compte de la relation

(4.21) $$\frac{\delta I_0}{\delta x}\left(2\sqrt{\frac{xu}{\mu}}\right) = \sqrt{\frac{u}{\mu x}}\, I_1\left(2\sqrt{\frac{xu}{\mu}}\right) \quad \text{[Réf. 14 , 15]}$$

où

(4.22) $$I(t) = \sum_{k=0}^{\infty} \frac{\left(\frac{t}{2}\right)^{2k}}{(k!)^2}$$

Cela donne

$$K(u) = \int_0^\infty e^{-x/\mu} \sqrt{\frac{u}{\mu x}}\, I_1\left(2\sqrt{\frac{xu}{\mu}}\right) dx = \left[e^{-x/\mu} I_0\left(2\sqrt{\frac{xu}{\mu}}\right)\right]_0^\infty + \int_0^\infty \frac{e^{-x/\mu}}{\mu} I_0\left(2\sqrt{\frac{xu}{\mu}}\right) dx$$

soit puisque $I_0(0) = 1$:

$$K(u) = -1 + \frac{1}{\mu}\int_0^\infty \frac{e^{-x/\mu}}{\mu} I_o\left(2\sqrt{\frac{xu}{\mu}}\right) dx$$

En remplaçant $I_0\left(2\sqrt{\frac{xu}{\mu}}\right)$ par sa valeur, on a :

$$K(u) = -1 + \frac{1}{\mu}\sum_{k=0}^{\infty} \frac{u^k}{(k!)^2} \int_0^\infty e^{-x/\mu} \left(\frac{x}{\mu}\right)^k dx$$, et enfin :

(4.23) $$K(u) = e^u - 1$$

Le deuxième membre s'écrit alors

$$\int_0^\infty K(u)\, e^{\frac{-(u+(u-m)^2)}{2\sigma^2}}\, du = \int_0^\infty e^{\frac{-(u-m)^2}{2\sigma^2}}\, du - J_0$$

En calculant la première intégrale (poser $\frac{u-m}{\sigma} = t$) et en remplaçant J_0 par sa valeur donnée par (4.17), on trouve finalement :

$$\text{(4.20)} \qquad \int_0^\infty K(u)\, e^{\frac{-(u+(u-m)^2)}{2\sigma^2}}\, du = \sigma\sqrt{2\pi}\left[F\left(\frac{m}{\sigma}\right) - e^{-m+\sigma^2/2}\, F\left(\frac{m}{\sigma} - \sigma\right)\right]$$

Pour trouver les caractéristiques de S_N, on peut commencer par trouver celles de N.

La relation (4.15) fournit :

$$E(N) = \frac{1}{\sigma\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \int_0^\infty \sum_{n=1}^{\infty} \frac{u^{n-1}}{(n-1)!}\, u\, e^{-\left(u+\frac{(u-m)^2}{2\sigma^2}\right)} du$$

$$= \frac{1}{\sigma\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \int_0^\infty u\, e^{-\frac{(u-m)^2}{2\sigma^2}} du$$

L'intégrale se calcule aisément en posant $t = \frac{u-m}{\sigma}$ et l'on trouve :

$$\text{(4.24)} \qquad E(N) = m + \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}$$

Exactement de la même façon, on peut écrire

$$E[N(N-1)] = \frac{1}{\sigma\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \int_0^\infty u^2\, e^{-\frac{(u-m)^2}{2\sigma^2}} du$$

soit en modifiant convenablement l'intégrande

$$E[N(N-1)] \quad \frac{\sigma}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \int_0^\infty e^{-\frac{(u-m)^2}{2\sigma^2}} \left[\frac{(u-m)\,u}{\sigma^2} + \frac{um}{\sigma^2}\right] du$$

Cela fait apparaître une somme d'intégrales dont la première s'intègre facilement par parties et dont la seconde vient d'être calculée. On trouve finalement :

$$\text{(4.25)} \qquad E[N(N-1)] = m^2 + \sigma^2 + \frac{m\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}, \text{ ce qui fournit :}$$

$$\text{(4.26)} \qquad E(N^2) = \sigma^2 + (m+1)\left[m + \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}\right] \text{ et}$$

$$V(N) = \sigma^2 + m + \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)} \left[1 - m - \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}\right] \tag{4.27}$$

On peut d'ailleurs retrouver les formules (4.24) et (4.27) à l'aide de la fonction génératrice de N, dont on peut démontrer qu'elle a l'expression suivante :

$$g_N(t) = \frac{F\left(\frac{m}{\sigma} + \sigma(t-1)\right)}{F\left(\frac{m}{\sigma}\right)} \exp\{(t-1)\,[m + (t-1)\,\sigma^2/2]\} \tag{4.28}$$

Les relations (3.4) et (3.5) permettent maintenant d'écrire :

$$E(S_N) = \mu\left(m + \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}\right) \tag{4.29}$$

$$V(S_N) = \mu^2\left[\sigma^2 + 2m + \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}\left(2 - m - \frac{\sigma\, e^{-m^2/2\sigma^2}}{\sqrt{2\pi}\, F\left(\frac{m}{\sigma}\right)}\right)\right] \tag{4.30}$$

Dans certains cas enfin [Réf. 10], il s'est avéré que le coût d'un accident était relativement bien représenté par un "mélange de distributions exponentielles" [Réf. 16] de densité

$$u(x) = \alpha\, \theta_1\, e^{-\theta_1 x} + (1-\alpha)\, \theta_2\, e^{-\theta_2 x} \quad \text{pour } x \geq 0 \quad (\theta_1 \text{ et } \theta_2 > 0) \tag{4.31}$$

et il faudrait alors revoir les formalisations précédentes, compte tenu de cette distribution.

Nous donnons pour notre part en annexe la densité de la somme de variables parentes suivant la distribution (4.31) (coût total quand le nombre d'accidents est connu).

CONCLUSION

Bien entendu, l'ensemble des formalisations présentées ci-dessus ne sont que des propositions. Certaines des distributions de probabilité dont il a été question ont été testées sur des observations pratiques relatives aux accidents de la route, mais il reste que l'on peut encore améliorer les choses :

- en prenant pour le nombre d'impliqués des lois de Poisson tronquées aussi supérieurement, puisqu'il est rare que ce nombre dépasse un certain seuil considéré comme une limite pratique,

- en séparant éventuellement les accidents en trois types comme il est commun de le faire en pratique : accidents à un véhicule, à deux véhicules et à plus de deux véhicules impliqués.

On notera cependant que les expressions deviennent rapidement complexes et il ne faut évidemment pas perdre de vue l'aspect utilisation pratique des formalisations proposées.

Un dernier problème enfin mérite d'être noté : celui de la comparaison dans le temps ou dans l'espace des "taux d'accidents". Etant censés mesurer le risque proprement dit, ces derniers sont représentés par le rapport du nombre d'accidents au nombre de kilomètres parcourus et constituent des indicateurs importants permettant de mesurer l'efficacité de mesures de prévention. Ils n'ont malheureusement subi jusqu'à présent qu'un traitement pour le moins empirique [Réf. 17 et 18] mais il semble que l'on puisse assez rapidement formaliser les problèmes de taux d'accidents de façon relativement opérationnelle.

ANNEXE

DISTRIBUTION DE LA SOMME DE VARIABLES PARENTES AYANT POUR DENSITE UN "MELANGE EXPONENTIEL"

On notera d'abord que si X est une variable aléatoire de densité

$$u(x) = \alpha\ \theta_1\ e^{-\theta_1 x} + (1-\alpha)\ \theta_2\ e^{-\theta_2 x}, \tag{1}$$

sa fonction génératrice est donnée par :

$$g_X(t) = \int_0^{\infty} e^{tx}\ u(x)\ dx = \alpha\left(1-\frac{t}{\theta_1}\right)^{-1} + (1-\alpha)\left(1-\frac{t}{\theta_2}\right)^{-1} \tag{2}$$

D'où $g_{\sum_1^n x_i}(t) = \left[\alpha\left(1-\frac{t}{\theta_1}\right)^{-1} + (1-\alpha)\left(1-\frac{t}{\theta_2}\right)^{-1}\right]^n$, c'est-à-dire en développant

$$g_{\sum_1^n x_i}(t) = \sum_{j=0}^{n} C_n^j\ \alpha^j\left(1-\frac{t}{\theta_1}\right)^{-j}(1-\alpha)^{n-j}\left(1-\frac{t}{\theta_2}\right)^{-n+j} \tag{3}$$

Or le terme courant de cette somme soit $\left(1-\frac{t}{\theta}\right)^{-j}\left(1-\frac{t}{\theta}\right)^{-n+j}$ représente la fonction génératrice de la somme de deux variables indépendantes suivant respectivement les lois $\gamma(j, \theta_1)$ et $\gamma(n-j, \theta_2)$ (cf. 3.8).

La variable de génératrice $\left(1-\frac{t}{\theta_1}\right)^{-j}\left(1-\frac{t}{\theta_2}\right)^{-n+j}$, a donc pour densité :

$$v_j(z, \theta_1, \theta_2) = \int_0^z \frac{\theta_1^j\ e^{-\theta_1 y}\ y^{j-1}}{(j-1)!}\ \frac{\theta_2^{n-j}\ e^{-\theta_2(z-y)}\ (z-y)^{n-j-1}}{(n-j-1)!}\,dy \tag{4}$$

et la densité de $\sum_{1}^{n} X_i$ est donc donnée par :

$$(5) \quad \boxed{w_n(x, \theta_1, \theta_2) = \frac{(1-\alpha)^n \theta_2^n e^{-\theta_2 x} x^{n-1}}{(n-1)!} + \sum_{j=1}^{n-1} C_n^j \alpha^j (1-\alpha)^{n-j} v(x, \theta_1, \theta_2) + \frac{\alpha^n \theta_1^n}{(n-1)!} e^{-\theta_1 x} x^{n-1}}$$

REFERENCES BIBLIOGRAPHIQUES

[1] MICHAEL - Two simple techniques for determining the significance of accident reducing measures. Public Roads 30, p. 238/239 (1959).

[2] DIETZ - Significance test for accident reductions based on classical statistics and economic consequences. Transportation Science 1, p. 206/217 (1967).

[3] O.N.S.E.R.* - Expérience de limitation de la vitesse sur 1600 km de routes nationales (rapport provisoire 1969).

[4] O.N.S.E.R.* - Efficacité des aménagements de sécurité (Cahiers d'études).

[5] CALOT - Exercices de calcul des probabilités (Dunod).

[6] MORAN P. - Introduction to probability theory. Clarendon Press, Oxford (1968).

[7] GREENWOOD et YULE - An inquiry into the nature of frequency distributions representative of multiple happenings J.R.S.S. 83, p. 255/279.

[8] O.N.S.E.R. - Classement des sections de routes nationales selon leur dangérosité.

[9] BILLINGSLEY et JORGENSON - Analyses of direct costs and frequencies of Illinois Motor-Vehicle accidents, 1958. Public Roads, Août 1963.

[10] WEBER - Notes on accident frequency models (unpublished).

[11] GARWOOD - Use of chi-squared distribution for comparing accident frequencies. Symposium on the use of satistical methods in the analysis of Road Accidents, April 69 (Crowthorne, Berkshire).

[12] BERLJAND, NAZAROV et PRESSMAN - i erfc-distribution or mixed-Gauss-Poisson distribution. Doklady Akademii Nauk S.S.S.R., 147, 1 005-1 007.

[13] SROUR D. - Méthodes mathématiques et statistiques appliquées à l'étude de la prédisposition aux accidents. Thèse de 3ème cycle, Université de Paris, (1968).

----

* Organisme National de Sécurité Routière.

[14] WHITTAKER et WATSON - A course of modern analysis. Cambridge University Press (1902).

[15] SCHWARTZ L. - Méthodes mathématiques pour les sciences physiques (Hermann).

[16] RIDER P. - The method of moments applied to a mixture of two exponential distributions.

[17] MORIN D.A. - Application of statistical concepts to accident data. Highway research board bulletin.

[18] NORDEN, ORLANSKY et JACOBS - Application of statistical quality control techniques to analysis of Highway-accident data. Highway research bulletin.