

Ajustements à des séries de Poisson

Journal de la société statistique de Paris, tome 98 (1957), p. 139-144

http://www.numdam.org/item?id=JSFS_1957__98__139_0

© Société de statistique de Paris, 1957, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Ajustements à des séries de Poisson

M. LE PRÉSIDENT, M. le PROFESSEUR,
MESDAMES, MESSIEURS,

Pour mal défini qu'est le terme « statistique », elle recherche les lois du hasard et l'art d'utiliser le calcul des probabilités.

Le sujet est l'étude d'un problème biologique que l'on a essayé de résoudre par la loi de Poisson. C'est le problème du tirage d'une boule noire dans une urne où se trouvent un grand nombre de boules blanches. Le passé n'influant par sur l'avenir, l'éventualité la plus probable sur 1.000 épreuves peut être que la boule noire sorte une fois. Or elle peut très bien ne pas sortir, ou sortir plusieurs fois (2, 3, 4... fois).

L'étude suivante n'a nullement la prétention d'une théorie mathématique de la question, d'autant qu'à des mathématiciens purs il nous est impossible, sinon ridicule, de rivaliser.

La méthode est une méthode *graphique* et nous appliquerons brutalement des formules, peu connues d'ailleurs car presque exclusivement mises au point vers 1926 et employées par les Bell Telephone Laboratories aux États-Unis pour l'étude des fréquences d'appels et à des fins électriques je crois, tout comme le chercheur ou le technicien applique telle ou telle formule pour résoudre tel ou tel problème sans rechercher la démonstration de celle-ci, ce qui serait inutile. Seul le résultat compte.

Avant de continuer, je tiens tout spécialement à remercier Monsieur le Professeur Dufrenoy d'une part, pour avoir mis à ma disposition tout le matériel nécessaire pour mener à bien cette étude, et d'autre part, M. Lhoste qui a bien voulu me confier ses séries de chiffres dont il est l'auteur.

En 1955, nous avons étudié en Bretagne le mode d'action de divers insecticides dans le sol contre les larves d'« *Elatерidae* », petits coléoptères causant de gros dégâts aux tubercules de pommes de terre en particulier.

Trouver un champ bien infesté de ces « Taupins » et de dimensions convenables est pratiquement introuvable aujourd'hui car la Bretagne, d'arrière qu'elle fut dans le domaine agricole, est devenue dans certaines régions à l'avant-garde du progrès. Je m'explique : le cultivateur de nos jours traite et sait traiter; les méthodes deviennent plus rationnelles... et ce qui nous a valu de trouver nos terrains d'expérience fut simplement la connaissance d'un brave paysan que la routine a conduit à ne pas s'écarter des méthodes ancestrales. Peut-être est-il dans la vérité? Car nous autres, ne nous intoxiquons-nous pas

par des produits qui, en principe, sont destinés à conserver la race humaine, par là même où nous recherchons toutes les qualités d'une nourriture saine (végétale comme animale) en collaboration avec nos confrères de la Technologie alimentaire?

Donc nos champs ont été divisés les uns en « carrés latins », les autres dans un compromis assez particulier du « bloc » et du « carré ». En voici la distribution de l'un d'eux qui tente de se rapprocher des meilleures conditions pour que les résultats puissent être soumis à l'étude statistique.

V	T	K	A
T	K	A	V
K	A	V	T
A	V	T	K

FIG. 1.

Ce champ a été traité par différents produits à tester que l'on compare évidemment au terme de référence qu'est le « Témoin ».

A l'époque critique les comptages ont eu lieu. Dans chaque parcelle, 1.500 tubercules environ, et ce dans la partie centrale pour éviter l'influence des bords, sont observés de la manière suivante : nous relevons le nombre (ou fréquences) de tubercules ayant subi 0, 1, 2, 3, 4 15 et plus de 15 piqûres d'insectes (= Manifestations).

Sur l'ensemble des comptages, nous n'étudierons qu'une distribution pour ne pas nous écarter du sujet.

PRODUIT V

<u>m</u>	<u>h</u>	<u>m × h</u>	<u>%</u>	<u>% cumulatifs</u>
0	1.371	0	91,33	99,97
1	63	63	4,19	8,64
2	31	62	2,065	4,45
3	16	48	1,066	0,239
4	9	36	0,599	0,133
5	4	20	0,266	00,731
6	2	12	0,133	00,465
7	2	14	0,133	00,332
8	2	16	0,133	00,199
9	1	9	0,066	00,066
	<u>1.501</u>	<u>280</u>		

$$\bar{x} = \frac{280}{1.501} = 0,186 \text{ (}\bar{x} \text{ est la moyenne de la distribution).}$$

Le Bell System Technical Journal en octobre 1926 fit paraître les « Poissons's Exponential Binomial Limit » contenant 2 tables, l'une des termes individuels, la seconde des termes cumulés.

Expliquons sommairement l'une et l'autre :

Dans la table I (*Individual Terms of the Poisson Formula* $\frac{a^x e^{-a}}{x}$) pour chaque valeur des moyennes « a », allant de .001 à 100, des distributions, on a les termes individuels d'une distribution théorique.

Par exemple :

x	$a = M$
0	.91
1	.05
2	.03
3	.01
Total .	100

Par contre dans la Table II, ce sont ces termes mais cumulés de bas en haut.

(*Cumulative Terms of the Poisson Formula* $P(c, a) = \sum_{x=c}^{\infty} \frac{a^x e^{-a}}{x}$)

x	$a = M$
0	100
1	.09
2	.04
3	.01

Les courbes de la figure (2) sont justement construites à l'aide des pourcentages cumulés. Prenons dans la table II les termes cumulés pour la moyenne $a = .08$, nous lisons :

c	$a = 0.08$
0	1.000 000 0
1	.076 883 7
2	.003 034 3
3	.000 080 4
4	.000 001 6
5	

et pour $a = 1.0$ nous lisons :

c	$a = 1.0$
0	1.000 000
1	.6321.....
2	.264.....
3	.08.....
4	.018.....
	etc...

Ces valeurs ont été marquées sur le graphique.

La première ligne est toujours égale à 1 par définition. Les courbes sont construites sur un système de coordonnées Probabilité-normale/log. Sur la

perpendiculaire à $a = .08$ marquons les valeurs ci-dessus. Les courbes passeront donc par les points ainsi déterminés pour les différentes valeurs de la moyenne. Il est à préciser qu'ayant ces graphiques, les tables deviennent à ce moment-là inutiles.

La Probabilité qu'un événement se manifeste au moins (c) fois dans un grand nombre d'essais, la moyenne des manifestations étant α , est :

$$P = 1 - \left(1 + \frac{\alpha}{1!} + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^{c-1}}{(c-1)!} \right) e^{-\alpha}.$$

Sur les graphiques, les points correspondants aux niveaux de pourcentages (P) sur chacune des courbes $C = 1, C = 2 \dots C = n$ doivent s'aligner selon la perpendiculaire élevée de l'abscisse correspondant à la moyenne de la distribution, si la distribution correspond à une Série de Poisson.

Dans la plupart des distributions observées, l'interdépendance (ou la contagion) des « essais » constituant les éléments de la Série, tend à faire incliner vers la droite la ligne passant par les points.

Pour la distribution du Témoin et pour le produit « V », on peut faire passer une droite par l'ensemble des points.

Pour un autre essai nous avons obtenu une cassure.

La méthode graphique permet une évaluation rapide et précise de la différence entre résultats des traitements.

Une autre méthode est apparue beaucoup plus élégante : celle des « log normale ».

Il suffit de cumuler cette fois-ci les pourcentages de *haut en bas* pour obtenir sur une échelle appropriée un ajustement très supérieur et nettement plus facile à lire. Cette échelle est loglog. (fig. 3).

Témoin « T »

m	f	m × f	%	m + 1	% tubercules ayant moins de (m + 1) attaques
0	217	0	46,1	1	16,1
1	166	166	12,3	2	28,4
2	157	314	11,8	3	
3	152	456	11,3	4	51,5
4	117	468	8	5	
5	127	635	9	6	
6	74	444	5	7	
7	74	518	5	8	81,5*
8	74	592	5	9	
9	56	504	3,9	10	
					100

(* On peut grouper, ce qui fait penser aux tables de contingence.)

Produit « K »

m	f	m × f	%	m + 1	% tubercules
0	1.480	0	98,6	1	98,6
1	16	16	1,06	2	99,66
2	2	4	0,13	3	99,79
3	1	3	0,06	4	99,85
4	1	4	0,06	5	99,91

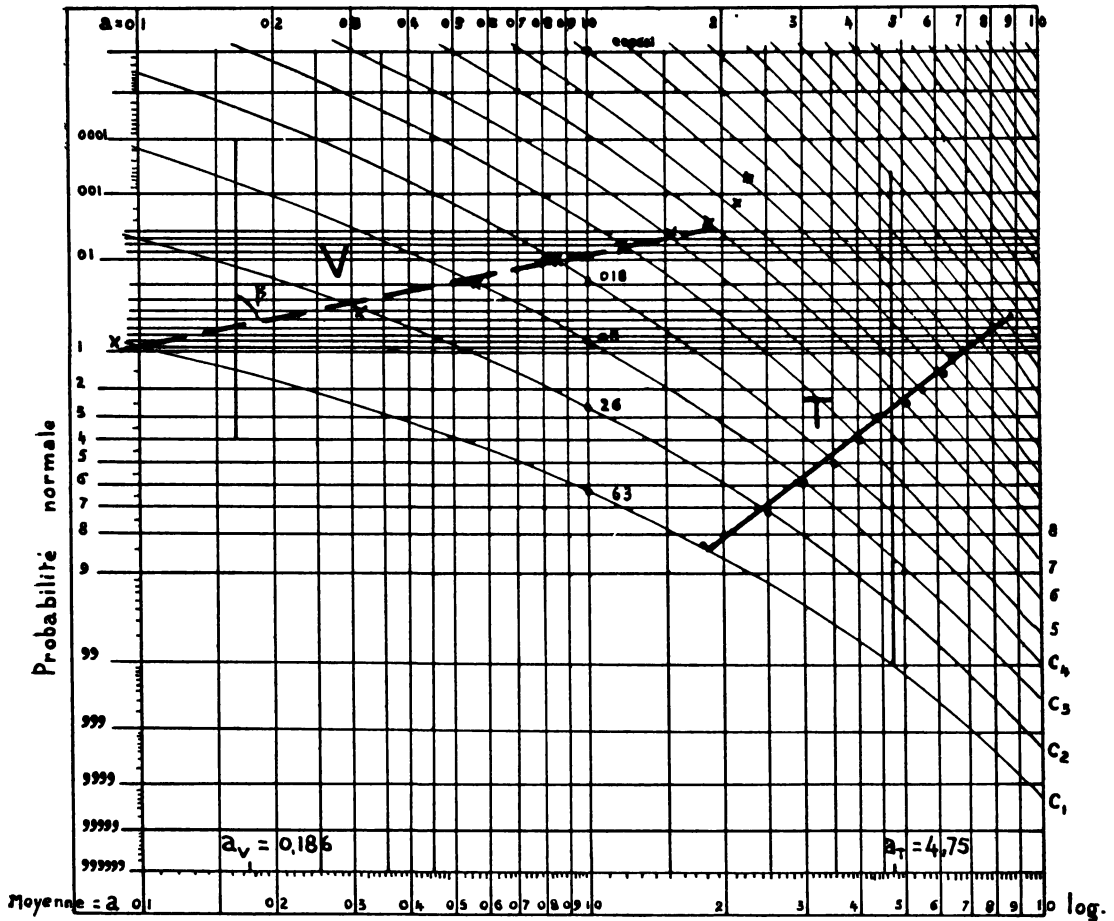


FIG. 2.

Probability curves showing Poisson's exponential summation

$$P = 1 - \left[1 + \frac{a}{1!} + \frac{a^2}{2!} + \dots + \frac{a^{c-1}}{(c-1)!} \right] e^{-a}$$

for the probability P that an event occur at least c times in a large group of trials for which the average number of occurrences is a. A scale proportional to the normal probability integral is used for P, a logarithmic scale is a.

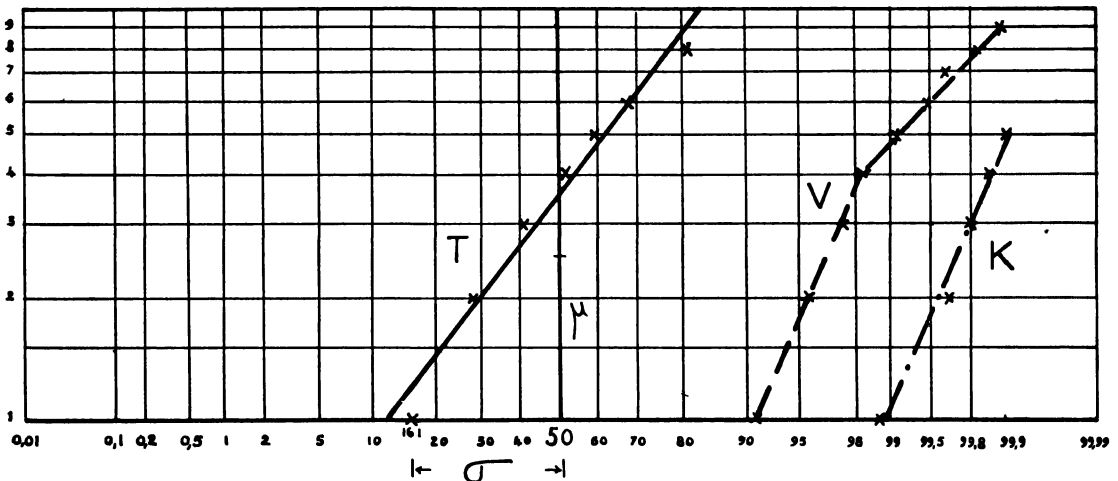


FIG. 3.

% tubercles ayant moins de (m + 1) attaques.

Produit « V »

0	1.371	0	91,33	1	91,33
1	63	63	4,19	2	95,52
2	31	62	2,065	3	97,58
3	16	48	1,0665	4	98,64
4	9	36'	0,599	5	99,23
5	4	20	0,2664	6	99,49
6	2	12	0,1332	7	99,62
7	2	14	0,1332	8	99,75
8	2	16	0,1332	9	99,88
9	1	9	0,06662	10	99,94

CONCLUSIONS

L'une et l'autre de ces méthodes ont l'avantage de la rapidité due aux moyens simples mis en œuvre. Un point seulement est délicat : comment estimer sur le graphique du Bell System une interprétation efficace pour les produits à action peu différente. Une publication ultérieure est à l'étude. Peut-être l'estimation et la comparaison de la pente β des droites nous apporteraient-elles des renseignements suffisants. Quant à l'explication des points d'infléchissement la question est aussi à l'étude car il ne nous a pas encore été possible d'obtenir la documentation nécessaire et complète, n'ayant travaillé qu'avec quelques rares ouvrages difficiles à se procurer.

Jean GÉRARD.

* * *