A PROPOS DE LA DISTRIBUTION DES CAS DE MALADIE ENTRE LES ASSURES ET PAR RAPPORT A LA DUREE

par

MARIO ALBERTO COPPINI

Rome (Italie)

1. Il est un fait bien connu que la technique des assurances, qu'elles soient privées ou sociales, qui versent une indemnité en cas de maladie, se base premièrement sur la connaissance des fréquences et des durées moyennes de maladie, ou bien des coefficients de morbidité. Pour résoudre des problèmes particuliers, des instruments plus raffinés sont toutefois nécessaires, savoir:

a) la distribution des assurés d'après le nombre de cas dont ils ont été frappés au cours d'une année;
b) la distribution des cas d'après la durée en jours;
c) la distribution des assurés d'après les jours de maladie dont ils ont été frappès au cours d'une annèe.

Nous avons eu la possibilité d'elaborer les fiches de 44.829 travailleurs salariés (36.134 hommes et 8.695 femmes âgés de 15 à 65 ans) résidant à Rome et assurés obligatoirement auprès de l'Institut National d'assurance-maladie (Istituto Nazionale per l'assicurazione contro le malattie — INAM), et de relever les cas de maladie qui les ont frappés au cours de l'année 1960, en même temps que les durées correspondantes. [1]).

Puisqu'il s'agit, comme on le verra, de données très détaillées, il nous a paru utile d'étudier les caractéristiques principales des trois distributions en question et les relations existant entre elles: l'exposé de certains résultats obtenus dans ce sens constitue l' objet de la présente note.

2. La table 1 recueille en premier lieu les deux distributions d'après le sexe visées à la lettre a) du paragraphe précédent, dont on obtient les indices suivants:

---

[1]) Il s'agit plus précisément de cas qui ont pris fin en 1960, à l'exclusion de ceux qui n'ont pas donné lieu à annotations sur la fiche individuelle pour le motif que le pronostic formulé était inférieur à quatre jours.

https://doi.org/10.1017/S0515036100007613 Published online by Cambridge University Press

	Hommes	*Femmes*
moyenne (nombre moyen de cas rapporté à tous les assurés)	0,502	0,534
indice de répétition (nombre moyen de cas rapporté aux seuls assurés frappés d'un cas ou plus)	1,467	1,573
variance	0,7129	0,8600

Nous pouvons nous demander quelle est la distribution théorique des assurés dans l'hypothèse où, une moyenne donnée ayant été fixée, l'attribution des cas de maladie aux différents individus aurait lieu au hasard: abstraction faite de la variation de la probabilité de tomber malade au cours de l'année et de la durée de la maladie, il est facile de se rendre compte que pareille distribution est celle de Poisson, puisque l'hypothèse en question correspond au schéma sur la base duquel, pour chaque assuré, grâce à une série de tirages de boules blanches et noires contenues dans une urne (les boules noires étant dans une proportion assez réduite), on attribue un cas de maladie toutes les fois que sort une boule noire.

TABLE I

Distribution des assures d'après le nombre de cas de maladie

Nombre de cas	Nombre des assurés			
	Hommes		Femmes	
	Données observées	Données ajustées (a)	Données observées	Données ajustées (a)
0	23.771	23.765	5.741	5.729
1	8.439	8.398	1.890	1.902
2	2.624	2.725	662	677
3	905	860	253	245
4	293	267	94	91
5	64	83	38	33
6	31	25	12	12
7	5	7	4	4
8	2	4	1	2
Au total	36.134	36.134	8.695	8.695

(a) — Sur la base de la fonction (5).
Source: INAM — Travailleurs salariés de la province de Rome.
Période d'observation: Année 1960.

On n'ignore pas que, d'après cette loi, la fréquence de ceux qui sont frappés de r cas assume la forme

$$(1) \qquad f(r) = e^{-\lambda} \frac{\lambda^r}{r!} \qquad \left(\sum_0^\infty {}_r f(r) = 1 \right)$$

En indiquant par m et s^2 respectivement la moyenne et la variance des distributions examinées et par ρ le coefficient de répétition, [1]) si les distributions susdites suivent la loi de Poisson, nous devrions avoir

$$(2) \qquad \lambda = m = s^2$$

$$(3) \qquad \rho_p = \frac{\lambda}{1 - e^{-\lambda}} = \frac{m}{1 - e^{-m}}$$

ρ_p étant le coefficient de répétition dans le cas de la dite loi.

On a vu que (2) n'est pas vérifié: en outre, en calculant ρ_p, on obtient les valeurs 1,272 pour les hommes et 1,291 pour les femmes, valeurs toutes deux inférieures aux valeurs observées: on peut donc conclure que *la répétition plus élevée dépend du fait que celui qui est tombé malade a une probabilité plus forte d'être frappé à nouveau* [2]).

D'autre part, au cas où l'alternative opposée se vérifierait, on aurait évidemment une valeur de ρ inférieure à celle fournie par (3) On peut donc affirmer comme *indice de propension aux maladies* (*subséquentes*) la quantité

$$(4) \qquad R = \frac{\rho - \rho_p}{\rho_p}$$

qui assume des valeurs positives ou negatives par rapport aux deux alternatives énoncées.

---

[1]) On observera que le rapport entre la moyenne des cas et le coefficient de répétition donne la probabilité de tomber malade au moins une fois.

[2]) Il faut observer que la probabilité plus élevée de tomber malade pour une personne ayant déjà été frappée de maladie, peut dépendre du fait que la première maladie rend l'organisme plus exposé à de nouveaux événements morbides, ou bien encore de l'existence au sein du groupe observé d'une hétérogénéité *a priori* des individus considérés par rapport à la possibilité de tomber malade. La discrimination entre ces deux possibilités n'est pas facile à faire ni sur le plan théorique ni sur celui de la récolte des données: nous avons donc préféré nous abstenir de toute considération à cet égard.

Pour les deux distributions examinées ici, R assume la valeur de $+ 0{,}153$ pour les hommes et de $+ 0{,}218$ pour les femmes.

Pour des phénomènes analogues où l'on avait également relevé une valeur de la variance supérieure à la moyenne, GREENWOOD et YULE [1]) ont proposé, sur la base de considérations que nous ne reproduisons pas ici, la distribution de fréquences suivante

$$(5) \quad f(r) = \binom{p+r-1}{r} \left(\frac{a}{a+1}\right)^p \left(\frac{1}{a+1}\right)^r, \qquad \left(\sum_0^\infty {}_r f(r) = 1\right)$$

connue comme la distribution binomiale négative, pour laquelle on a

---

[1]) Cfr. M. GREENWOOD et G. U. YULE, *An inquiry into the nature of frequency-distributions of multiple happenings, etc.*, dans „Journal of the Royal Statistical Society", 1920, vol. 83.

Il peut être utile d'observer que la distribution binomiale négative (formule (5) du texte) résulte d'un mélange de distributions de Poisson (formule (1) du texte) qu'on obtient en donnant au paramètre λ ($= m$) toutes les valeurs de 0 à ∞ et en attribuant à chacune de ces distributions de Poisson un poids exprimé par la fonction

$$g(\lambda) = \frac{a^p}{\Gamma(p)} e^{-a\lambda} \lambda^{p-1}$$

c'est-à-dire qu'on obtient

$$\int_0^\infty g(\lambda)\, e^{-\lambda} \frac{\lambda^r}{r!}\, d\lambda = \binom{p+r-1}{r} \left(\frac{a}{a+1}\right)^p \left(\frac{1}{a+1}\right)^r.$$

La bonne adhérence des données observées à la fonction (5) peut donc être interprétée dans le sens que de nombreux sous-groupes coexistent au sein du groupe d'assurés observé, chacun de ces sous-groupes ayant une valeur propre de λ ($= m$), c'est-à-dire son propre nombre moyen de cas de maladie auxquels on peut s'attendre pendant la période annuelle prise en considération.

Les valeurs de la fonction $g(\lambda)$ ont résulté être les suivants

λ ($= m$)	Hommes	Femmes	λ ($= m$)	Hommes	Femmes
0 — 0,1	1.597	1.513	1,4 — 1,6	195	281
0,1 — 0,2	1.445	1.285	1,6 — 1,8	126	203
0,2 — 0,3	1.248	1.090	1,8 — 2,0	80	145
0,3 — 0,4	1.048	925	2,0 — 2,5	97	211
0,4 — 0,5	872	785	2,5 — 3,0	31	92
0,5 — 0,6	710	666	3,0 — 3,5	11	41
0,6 — 0,7	589	566	3,5 — 4,0	3	18
0,7 — 0,8	479	479	4,0 — 4,5	1	8
0,8 — 0,9	388	408	4,5 — 5,0	1	3
0,9 — 1,0	317	345	5,0 — 5,5		2
1,0 — 1,2	460	542	5,5 — 6,0		1
1,2 — 1,4	302	391		10.000	10.000

(6) $$m = \frac{p}{a}$$

(7) $$s^2 = \frac{p}{a} + \frac{p}{a^2}$$

En déterminant a et p sur la base des (6) et (7) et des valeurs observées de m et de s^2, on a obtenu

	Hommes	*Femmes*
a	2,377	1,641
p	1,193	0,877

Les distributions théoriques ainsi déterminées, qui sont également exposées dans la Table 1, constituent un excellent ajustement de celles déduites de l'expérience.

On peut observer enfin que pour la distribution (5) on a, en rappelant (6) et (7),

(8) $$\rho = \frac{m}{1 - \left(\frac{s^2}{m}\right)^{-p}}$$

3. Les distributions des cas observés par rapport à la durée sont reproduites en synthèse dans la Table 2 et en détail dans les tableaux I et II annexés à la présente note: dans ces derniers documents, les cas sont distingués d'après les deux critères suivants:

a) *l'ordre* dans lequel ils ont eu lieu vis-à-vis de l'assuré;

b) leur *multiplicité* toujours vis-à-vis de l'assuré (c'est-à-dire le fait de se référer à des assurés frappés d'un seul cas, de deux seuls cas, etc.).

Nous nous sommes demandés si ces classifications excercent ou non une influence sur la distribution des cas par durée. Après avoir effectué l'analyse de la variance, les résultats suivants ont été obtenus.

Distributions d'apres l'ordre

	Hommes	
	Variances	Degrés de liberté
Variance entre les moyennes	608,62	4
Variance résiduelle	648,05	18.126
Variance globale	648,04	18.130

rapport entre les variances = 1,06 F 5% = 5,63

	Femmes Variances	Degrés de liberté
Variance entre les moyennes	386,66	4
Variance résiduelle	468,53	4.642
Variance globale	468,46	4.646

rapport entre les variances = 1,21 F 5% = 5,63.

Distributions d'apres la multiplicite

	Hommes Variances	Degrés de liberté
Variance entre les moyennes	3.110,34	4
Variance résiduelle	647,50	18.126
Variance globale	648,04	18.130

rapport entre les variances = 4,80 F 5% = 2,37

	Femmes Variances	Degrés de liberté
Variance entre les moyennes	372,94	4
Variance résiduelle	468,54	4.642
Variance globale	468,46	4.646

rapport entre les variances = 1,26 F 5% = 5,63

Comme on peut le constater, une seule classification résulte être significative, et précisément celle d'après la multiplicité pour le sexe masculin.

Les deux distributions globales ont les moyennes et variances suivantes:

	Hommes	*Femmes*
durée moyenne (en jours)	17,45	14,61
variance (en jours)	648,11	468,29

Au moyen de ces données, on peut se proposer de vérifier quelles sont les moyennes de chaque distribution des tableaux I et II qui diffèrent d'une manière significative des durées moyennes globales, afin d'approfondir les résultats obtenus par l'analyse des variances. Dans ce but, ayant indiqué par A et σ^2 respectivement la moyenne

https://doi.org/10.1017/S0515036100007613 Published online by Cambridge University Press

et la variance des distributions globales, par N le nombre global des observations et par n le nombre des observations contenues dans une distribution générique partielle, nous avons calculé les quantités $2\sigma\sqrt{\frac{N-n}{(N-1)\,n}}$ et nous avons vérifié si les différences entre la moyenne globale et les moyennes de chaque distribution sont supérieures en valeur absolue aux intervalles de confiance susdits. Les résultats sont les suivants:

Distributions d'après l'ordre

	Hommes			Femmes		
	Moyennes	Différence avec la moyenne générale	Intervalle de confiance	Moyennes	Différence avec la moyenne générale	Intervalle de confiance
premiers cas	17,40	— 0,05	0,26	14,66	+ 0,05	0,48
deuxièmes cas	17,72	+ 0,27	0,72	14,88	+ 0,27	1,17
troisièmes cas	17,27	— 0,18	1,36	14,66	+ 0,05	2,06
quatrièmes cas	18,27	+ 0,82	2,53	11,48	— 3,13	3,49
cinquièmes cas et cas appartenant à des ordres plus élevés	13,95	— 3,50	4,15	14,57	— 0,04	4,86

Distributions d'après la multiplicité

	Hommes			Femmes		
	Moyennes	Différence avec la moyenne générale	Intervalle de confiance	Moyennes	Différence avec la moyenne générale	Intervalle de confiance
cas uniques	17,52	+ 0,07	0,41	14,80	+ 0,19	0,77
cas doubles	17,70	+ 0,25	0,59	15,12	+ 0,51	1,01
cas triples	17,84	+ 0,39	0,90	14,29	— 0,32	1,44
cas quadruples	17,16	— 0,29	1,44	13,25	— 1,36	2,14
cas quintuples et cas d'une multiplicité plus élevée	12,89	— 4,56	2,12	13,61	— 1,00	2,43

Comme on peut le constater, seule la moyenne de la distribution relative au sexe masculin visant les cas quintuples et d'une multiplicité plus élevée résulte être significative.

On peut ainsi conclure que, *sauf cas exceptionnels, ni l'ordre ni la multiplicité influent sur la distribution par durées.*

Pour les cas doubles on a en outre élaboré au moyen du matériel original les tables à double entrée des assurés classifiés par rapport aux durées des deux cas, et on a obtenu les coefficients de corrélation suivants:

pour les hommes $+0{,}139$
pour les femmes $+0{,}121$

Les valeurs assez modestes de ces coefficients permettent d'affirmer que, *tout au moins pour les cas doubles, la durée du second cas résulte être presque indépendante de celle du premier cas.*

TABLE 2

Distribution des cas de maladie d'après la durée

Durée (en jours)	Hommes				Femmes			
	Données observées	Données ajustées (a)			Données observées	Données ajustées (a)		
		distribution globale	distributions partielles			distribution globale	distributions partielles	
1 — 5	3.237	3.372	3.372		3.889	3.822	3.822	
6 — 10	2.484	2.349	2.349		2.311	2.380	2.378	2
11 — 15	1.250	1.226	1.225	1	1.192	1.167	1.130	37
16 — 20	712	608	595	13	657	623	492	131
21 — 30	861	567	402	165	824	722	288	434
31 — 40	451	423	81	342	383	448	47	401
41 — 60	432	686	18	668	364	466	8	458
61 — 90	298	486	1	485	187	242		242
91 — 120	127	167		167	103	74		74
121 — 150	62	64		64	45	28		28
Au délà de 150 jours	86	52		52	45	28		28
	10.000	10.000	8.043	1.957	10.000	10.000	8.165	1.835

(a) Sur la base de la fonction (9) et de ses composantes. Source: INAM — Travailleurs salariés de la province de Rome. Période d'observation: Année 1960.

Il a été procédé enfin à l'ajustement de la distribution de la Table 2 avec la fonction suivante, proposée dans un de mes travaux antérieurs [1])

$$(9)\quad \varphi(u) = k_1 \varphi_1(u) + k_2 \varphi_2(u) = k_1 \frac{1}{a_1^{p_1+1}\,\Gamma(p_1+1)} u^{p_1} e^{-\frac{u}{a_1}} + $$

$$+ k_2 \frac{a_2^{p_2-1}}{\Gamma(p_2-1)} u^{-p_2} e^{-\frac{a_2}{u}}; \left(\int_0^\infty \varphi(u)\,du = 1;\, k_1 + k_2 = 1\right)$$

qui décompose la distribution globale en deux distributions partielles, fournies par le premier et le deuxième terme de la fonction (9), respectivement pour les maladies à élimination (de l'état morbide) rapide et pour les maladies à élimination lente.

Les valeurs des paramètres obtenus sont les suivantes:

	Hommes	*Femmes*
k_1	0,804	0,816
a_1	5,60	5,00
p_1	0,50	0,50
k_2	0,196	0,184
a_2	248,91	139,68
p_2	6,00	5,00

En indiquant par A_1, A_2, σ_1^2, σ_2^2, respectivement les moyennes et les variances de la fonction $\varphi_1(u)$ et de la fonction $\varphi_2(u)$ on obtient en outre:

	Hommes	*Femmes*
A_1	8,40	7,50
A_2	62,23	46,56
σ_1^2	47,04	37,50
σ_2^2	1.290,86	1.083,92

4. Etablissons maintenant la relation intervenant entre la distribution des assurés d'après le nombre de jours de maladie que nous indiquerons ci-après par $\psi(u)$ et celles des assurés d'après le nombre des cas et des cas d'après la durée.

---

[1]) Cfr. M. A. Coppini, *Reduction factors and sickness distribution by duration*, dans „Transactions of the First International Conference of Social Security Actuaries and Statisticians", Bruxelles, novembre 1956 — ISSA. Nous renvoyons le lecteur à ce travail pour la méthode utilisée dans la détermination des paramètres.

En attendant, il est évident qu'en indiquant par $\psi_1(u)$, $\psi_2(u)$, ..., $\psi_r(u)$ ... les distributions analogues des assurés qui ont été frappés respectivement de 1, 2, ... r ... cas, on a

$$(10)\qquad \psi(u) = \bar{f}(1)\,\psi_1(u) + \bar{f}(2)\,\psi_2(u) + \ldots + \bar{f}(r)\,\psi_r(u) + \ldots$$

$$\left(\int_0^\infty \psi(u)\,du = 1;\ \int_0^\infty \psi_r(u)\,du = 1\right)$$

où on a posé

$$(11)\qquad \varphi_{21}(u) * \varphi_{22}(u) = \bar{f}(r) = \frac{f(r)}{1 - f(0)}$$

Indiquons par $\varphi_{11}(u)$ la distribution des cas de maladie uniques, par $\varphi_{21}(u)$ et $\varphi_{22}(u)$ respectivement la distribution du premier et du second des cas doubles, par $\varphi_{r1}(u)\ \varphi_{r2}(u) \ldots \varphi_{rr}(u)$ respectivement la distribution du premier, du deuxième, ... du $r^{\text{ème}}$ cas des cas r^{uples}, etc. [1])

C'est évidemment

$$(12)\qquad \psi_1(u) = \varphi_{11}(u).$$

D'autre part, supposé que le nombre de jours de maladie qui frappe un assuré lequel a subi deux cas de maladie puisse être considéré comme la somme de deux variables casuels indépendants (qu'on se rappelle la légère corrélation verifiée au paragraphe 3), on obtient comme on le sait

$$(13)\qquad \psi_2(u) = \varphi_{21}(u) * \varphi_{22}(u)$$

en ayant posé

$$(14)\qquad \varphi_{21}(u) * \varphi_{22}(u(= \int_0^\infty \varphi_{21}(u - t)\,\varphi_{22}(t)\,dt$$

En raisonnant d'une manière analogue pour les cas triples, ... r^{uples} etc., on obtient en définitive

$$(15)\qquad \psi(u) = \bar{f}(1)\,\varphi_{11}(u) + \bar{f}(2)\,[\varphi_{21}(u) * \varphi_{22}(u)] + \ldots$$
$$\ldots + f(r)\,[\varphi_{r1}(u) * \varphi_{r2}(u) * \ldots * \varphi_{rr}(u)] + \ldots$$

[1]) On se rappellera le sens de cas uniques, doubles, triples, ... défini au paragraphe précédent.

Si on introduit encore l'hypothèse d'après laquelle toutes les fonctions $\varphi_{rs}(u)$ peuvent être remplacées, à des fins pratiques, par la fonction $\varphi(u)$, de même qu'il a été aussi vérifié au paragraphe 3, nous avons enfin

$$\begin{aligned}(16)\qquad \psi(u) = \bar{f}(1)\,\varphi(u) + \bar{f}(2)\,[\varphi(u) * \varphi(u)] + \ldots + \\ + \bar{f}(r)\,[\varphi(u) * \varphi(u) * \ldots * \varphi(u)] + \ldots \qquad r \text{ facteurs}\end{aligned}$$

La moyenne $\bar{A}$ et la variance $\bar{\sigma}^2$ de la distribution composée $\psi(u)$ s'obtiennent très facilement: en supposant connues les moyennes $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_r, \ldots$ et les variances $\bar{\sigma}_1^2, \bar{\sigma}_2^2, \ldots, \bar{\sigma}_r^2 \ldots$ des distributions composantes $\psi_1(u), \psi_2(u), \ldots, \psi_r(u) \ldots$, on obtient en effet

$$(17)\qquad \bar{A} = \Sigma_r\, f(r)\, \bar{A}_r$$

$$(18)\qquad \bar{\sigma}^2 = \Sigma\, \bar{f}(r)\, \bar{\sigma}_r^2 + \Sigma_{rs}\, \bar{f}(r)\, \bar{f}(s)\, (\bar{A}_r - \bar{A}_s)^2,$$

où le second membre de la somme s'entend étendu à toutes les combinaisons possibles des indices r et s. [1]) En admettant l'hypothèse d'indépendance sur laquelle (13) est basé, ainsi que l'hypothèse d'égalité que nous a mené à (16), nous obtenons en particulier

$$(19)\qquad \bar{A}_r = r\bar{A}$$

$$(20)\qquad \bar{\sigma}_r^2 = r\sigma^2$$

Si on se souvient que

$$(21)\qquad \rho = \Sigma_r\, f(r)\, r$$

et on pose

$$(22)\qquad \bar{s}^2 = \Sigma_{rs}\, f(r)\, \bar{f}(s)\, (r - s)^2,$$

en observant que la quantité correspondant au second membre de l'égalité est la moitié du carré de la différence moyenne quadratique avec répétition relative à la distribution $\bar{f}(r)$ et correspond donc [2]) à la variance de pareille distribution, on a en définitive

---

[1]) La formule (17) est immédiate; quant à la (18), il suffit d'observer que tous les moments depuis l'origine d'une distribution composée satisfont à une relation analogue à la (17), et que la variance est la différence entre le moment second et le carré du moment premier.

[2]) Cfr. C. Gini, *Memorie di metodologia statistica*, Vol. 1 , p. 228, Ed. Giuffré, Milano.

(23) $$\bar{A} = A\rho,$$

(24) $$\bar{\sigma}^2 = \sigma^2\rho + A^2\,\bar{s}^2.$$

Sur la base de la distribution des assurés par jours de maladie, qui est contenue dans la Table 3, et à la distribution des assurés par nombre des cas de la Table 1 (qui fournit pour $\bar{s}^2$ la valeur 0,6703 pour les hommes et la valeur 0,8980 pour les femmes), on a procédé par ailleurs à vérifier les relations (23) et (24). Sur la base des données analytiques du Tableau III qui contient les différentes distributions par nombre de jours des assurés frappés par un seul cas, par deux seuls cas, etc., on a également pu vérifier les relations (19) et (20) uniquement pour les quatre premières distributions. Les résultats sont les suivants:

	Moyennes		Variances	
	observées	calculées	observées	calculées
		Hommes		
distribution totale	25,45	25,60	1.159,93	1.154,88
distribution des assurés				
frappés de 1 cas	17,52	17,45	723,38	648,11
frappés de 2 cas	35,18	34,90	1.355,47	1.296,22
frappés de 3 cas	52,78	52,35	1.825,48	1.944,33
frappés de 4 cas	67,02	69,80	2.259,78	2.592,44
		Femmes		
distribution totale	22,96	22,98	907,98	928,30
distribution des assurés				
frappés de 1 cas	14,80	14,61	532,31	468,29
frappés de 2 cas	30,94	29,22	1.003,07	936,58
frappés de 3 cas	42,04	43,83	1.124,44	1.404,87
frappés de 4 cas	51,91	58,44	1.466,55	1.873,16

Comme on peut le voir, *ces données confirment également assez bien les hypothèses posées*, en tenant compte pour les ordres de multiplicité élevés du nombre limité d'observations.

5. Nous nous sommes proposé en dernier lieu la question de savoir si, la fonction $f(r)$ étant connue de même que la fonction $\varphi(u)$, en ses composantes $\varphi_1(u)$ et $\varphi_2(u)$, il serait possible de déterminer la fonction $\psi(u)$ grâce à une relation plus simple que (16), même si cette relation est d'une nature seulement expérimentale.

https://doi.org/10.1017/S0515036100007613 Published online by Cambridge University Press

TABLE 3

Distribution des assurés par nombre de jours de maladie

Durée (en jours)	Hommes				Femmes			
	Données observées	Données ajustées (a)			Données observées	Données ajustées (a)		
		distribution globale	distributions partielles			distribution globale	distributions partielles	
1 — 5	2.413	2.302	2.302		2.783	2.445	2.445	
6 — 10	2.071	1.623	1.604	19	1.929	1.590	1.522	68
11 — 15	1.243	1.030	836	194	1.158	1.129	723	406
16 — 20	798	848	406	442	761	895	315	580
21 — 30	1.047	1.365	274	1.091	1.087	1.518	184	1.334
31 — 40	650	922	55	867	663	990	30	960
41 — 60	698	977	12	965	731	715	5	710
61 — 90	487	511	1	510	437	433		433
91 — 120	239	230		230	244	146		146
121 — 150	137	86		86	71	60		60
Au délà de 150 jours	217	106		106	135	79		79
	10.000	10.000	5.490	4.510	10.000	10.000	5.224	4.776

(a) — Sur la base de la fonction (9) et de ses composantes. Source: INAM — Travailleurs salariés de la province de Rome. Période d'observation: Année 1960.

A cette fin, en rappelant (9) on peut écrire (16):

$$(25)\quad \psi(u) = \bar{f}(1)\, k_1\, \varphi_1(u) + \bar{f}(1)\, k_2\, \varphi_2(u) + \bar{f}(2)\, [\varphi(u) * \varphi(u)] + \ldots + \bar{f}(r)\, [\varphi(u) * \varphi(u) * \ldots * \varphi(u)] + \ldots$$

Admettons que la distribution constituée par tous les termes du second membre de (25) qui suivent le premier puisse être représentée par une fonction du type $\varphi_2(u)$, avec des valeurs opportunes des paramètres a_2 et p_2, c'est-à-dire:

$$(26)\qquad \psi(u) = \bar{f}(1)\, k_1\, \varphi_1(u) + (1 - \bar{f}(1)\, k_1)\, \varphi_2(u).$$

Or $\bar{A}$ et $\bar{\sigma}^2$ peuvent être déterminés sur la base des relations (23) et (24) en supposant connues la fonction $\bar{f}(r)$ et la fonction $\varphi(u)$; A_1 et σ_1^2 sont également connus sur la base de la connaissance de la fonction $\varphi_1(u)$. D'autre part, A_2 et σ_2^2 peuvent être déterminés, en tenant compte de (17) et de (18), au moyen des relations [1])

$$\bar{A} = \bar{f}(1)\, k_1\, A_1 + (1 - \bar{f}(1)\, k_1)\, A_2 \tag{27}$$

$$\bar{\sigma}^2 = \bar{f}(1)\, k_1 \sigma_1^2 + (1 - \bar{f}(1)\, k_1)\, \sigma_2^2 + \bar{f}(1)\, k_1\, (1 - \bar{f}(1)\, k_1)\, (A_1 - A_2)^2 \tag{28}$$

et les paramètres a_2 et p_2 au moyen des formules

$$A_2 = \frac{a_2}{p_2 - 2} \tag{29}$$

$$\sigma^2 = \frac{a_2}{(p_2 - 2)\sqrt{p_2 - 3}} \qquad (1) \tag{30}$$

Nous avons reporté dans la Table 3 les fréquences ajustées au moyen du procédé qui a été décrit: on peut constater que l'adaptation des valeurs calculées aux valeurs observées n'est pas très satisfais ante. Toutefois cette méthode fournit une indication suffisamment significative pour les fins pratiques, dans l'hypothèse posée où on ne disposerait d'aucune donnée relative aux fréquences susdites de la Table 3.

La moyenne, la variance et les paramètres de la fonction $\varphi_2(u)$ ont résulté être les suivants:

	Hommes	*Femmes*
A_2	46,54	39,91
$\sigma_2{}^2$	1.704,88	1.353,92
a_2	105,65	87,00
p_2	4,27	4,18

Il est opportun d'observer que la construction de la fonction $\psi(u)$ permet la détermination des facteurs de réduction rapportés aux assurés, au lieu de ceux rapportés aux cas de maladie, comme parfois cela s'avère nécessaire dans la technique des assurances envisagées.

[1]) Cfr. l'ouvrage cité en page 53.

Tableau I

Distributions par durée des cas de maladie en distinguant d'après l'ordre dans lequel ils ont eu lieu pour un même assuré

Durée de la maladie (en jours)	Nombre des cas de maladie									
	Premiers cas	Deuxièmes cas	Troisièmes cas	Quatrièmes cas	Cinquièmes cas et au delà	Premiers cas	Deuxièmes cas	Troisièmes cas	Quatrièmes cas	Cinquièmes cas et au delà
	Hommes					Femmes				
1	1.479	431	147	44	21	438	125	53	27	6
2	144	42	17	5	1	25	12	7	—	2
3	1.317	442	150	52	29	402	144	72	30	21
4	363	132	36	18	1	90	37	12	8	3
5	694	208	74	14	9	176	68	27	13	9
6	688	218	56	17	6	172	67	16	6	2
7	1.500	293	95	29	10	200	78	24	10	2
8	613	182	57	18	9	132	40	20	7	3
9	429	143	50	11	10	91	31	19	2	2
10	406	110	36	8	4	91	41	14	3	1
11 — 15	1.554	477	175	44	17	376	122	36	11	9
16 — 20	891	271	97	25	6	185	83	23	8	6
21 — 25	598	218	61	28	6	152	53	17	8	2
26 — 30	433	151	50	14	2	94	41	12	3	1
31 — 40	544	186	59	22	7	112	43	19	2	2
41 — 50	312	104	41	13	2	75	23	9	5	1
51 — 60	193	84	27	7	1	33	13	6	1	3
61 — 90	346	125	47	16	6	50	24	7	5	1
91 — 120	142	69	13	6	1	30	13	5	—	—
121 — 150	85	17	6	3	1	13	4	2	—	2
Au delà de 150 jours	127	21	6	1	—	17	2	2	—	—
	12.363	3.924	1.300	395	149	2.954	1.064	402	149	78

Tableau II

Distribution par durée des cas de maladie en distinguant d'après leur multiplicité par rapport à un même assuré

Durée de la maladie (en jours)	Nombre des cas de maladie									
	Cas uniques	Cas doubles	Cas triples	Cas quadruples	Cas quintuples et au delà	Cas uniques	Cas doubles	Cas triples	Cas quadruples	Cas quintuples et au delà
	Hommes					Femmes				
1	1.013	582	291	148	88	284	168	84	61	52
2	105	51	34	14	5	16	9	9	6	6
3	960	579	248	121	82	288	162	111	62	46
4	256	163	83	35	13	55	44	22	15	14
5	473	272	154	58	42	118	75	48	28	24
6	447	299	159	55	25	118	75	36	22	12
7	685	422	196	84	45	122	101	50	26	15
8	404	276	123	44	32	83	56	36	16	11
9	294	184	97	48	20	55	43	31	6	10
10	285	160	68	34	17	58	39	30	12	11
11 — 15	1.028	653	376	138	72	217	179	101	33	24
16 — 20	594	369	210	90	27	103	98	55	22	27
21 — 25	415	258	147	67	24	96	66	39	17	14
26 — 30	282	199	105	55	9	62	47	24	13	5
31 — 40	368	242	136	55	17	69	60	35	6	8
41 — 50	225	122	73	41	11	45	34	18	11	5
51 — 60	125	96	58	27	6	21	16	8	7	4
61 — 90	229	167	91	36	17	36	30	9	8	4
91 — 120	86	85	41	15	4	23	14	8	3	—
121 — 150	58	34	13	6	1	9	5	3	1	3
Au delà de 150 jours	107	35	12	1	—	12	3	2	1	3
	8.439	5.248	2.715	1.172	557	1.890	1.324	759	376	298

TABLEAU III

Distributions des assurés d'après le nombre des cas et des jours de maladie dont ils ont été frappés au cours de l'année

Nombre de jours de maladie	Nombre des assurés frappés de:									
	1 cas	2 cas	3 cas	4 cas	5 cas et plus	1 cas	2 cas	3 cas	4 cas	5 cas et plus
	Hommes					Femmes				
1	1.013	—	—	—	—	284	—	—	—	—
2	105	72	—	—	—	16	24	—	—	—
3	960	2	14	—	—	288	1	4	—	—
4	256	54	1	2	—	55	24	1	2	—
5	473	26	5	—	—	118	2	3	—	—
6	448	111	4	2	—	118	24	1	—	—
7	684	50	8	1	—	122	13	8	—	2
8	404	63	7	6	2	83	19	2	2	1
9	294	60	17	1	2	55	15	4	—	2
10	285	100	6	5	—	58	38	1	2	—
11 — 15	1.028	424	70	12	3	217	97	19	8	1
16 — 20	594	314	70	7	1	103	85	27	7	3
21 — 25	415	229	77	23	9	96	49	25	7	3
26 — 30	282	156	74	25	4	62	43	23	9	4
31 — 40	368	267	131	24	13	69	74	37	12	4
41 — 50	225	169	86	34	12	45	46	28	8	2
51 — 60	125	110	65	27	10	21	35	21	6	4
61 — 90	229	196	119	43	15	36	35	30	14	14
91 — 120	86	93	67	33	17	23	20	9	12	8
121 — 150	58	49	34	23	6	9	5	4	3	—
Au delà de 150 jours	107	79	50	25	8	12	13	6	2	7
	8.439	2.624	905	293	102	1.890	662	253	94	55

https://doi.org/10.1017/S0515036100007613 Published online by Cambridge University Press