Aktuárské vědy

Miloš Vacek

Sur la loi de Polya régissant les faits corrélatifs. II

Aktuárské vědy, Vol. 3 (1932), No. 2, 49–61

Persistent URL: `http://dml.cz/dmlcz/144566`

Sur la loi de Polya régissant les faits corrélatifs.

Dr. *Miloš Vacek.*

VII. Application du développement de Charlier.

Les fonctions génératrices de toutes les lois de la probabilité considérées peuvent être amenées à la forme

$$e^{-h(1-z)}\{1+\gamma_2(1-z)^2+\gamma_3(1-z)^3+\ldots\}=G(z). \tag{16}$$

Il en résulte

$$W_r=\sum_{i=0}^{\infty}\gamma_i\Delta^i Q_r, \quad \gamma_0=1, \quad \gamma_1=0 \tag{17}$$

ou Δ désigne la différence inférieure. - Car si nous mettons $F(z)=e^{-h(1-z)}$, nous obtiendrons

$$G(z)=F(z)\sum_{i=0}^{\infty}\gamma_i(1-z)^i=\sum_{i=0}^{\infty}\gamma_i\sum_{k=0}^{i}(-1)^k\binom{i}{k}z^k F(z).$$

Le coefficient de z^r dans $z^k F(z)$ est Q_{r-k}. Exprimons que les coefficients de z^r sont égaux:

$$W_r=\sum_{i=0}^{\infty}\gamma_i\sum_{k=0}^{i}(-1)^k\binom{i}{k}Q_{r-k}=\sum_{i=0}^{\infty}\gamma_i\Delta^i Q_r,$$

$\dfrac{\Delta^k Q_r}{Q_r}=q_k(r)$ sont les polynomes „G“. L'orthogonalité est leur propriété principale:

$$\left.\begin{aligned}&\sum_{r=0}^{\infty}q_i(r)q_k(r)Q_r=0, \quad i\neq k\\ &\sum_{r=0}^{\infty}q_k^2(r)Q_r=\frac{k!}{h^k}\,.\end{aligned}\right\} \tag{18}$$

Le développement de la fonction des nombres entiers donnée U_r

$$U_r=\gamma_0 Q_r+\gamma_1\Delta Q_r+\gamma_2\Delta^2 Q_r+\ldots$$

$$\sum_{r=0}^{\infty} q_k(r) U_r = \gamma_k \sum_{r=0}^{\infty} q_k^2(r) \, . \, Q_r = \gamma_k \frac{k!}{h^k}$$

$$\gamma_k = \frac{h^k}{k!} \sum_{r=0}^{\infty} q_k(r) \, U_r. \tag{19}$$

Nous trouverons les développements des lois envisagées. Nous pouvons utiliser la formule (19), mais nous pouvons aussi prendre pour point de départ le développement (16). La convergence du développement (16) se démontre facilement par la méthode des fonctions majorantes appliquée à chaque cas. De la convergence absolue (16) résulte la convergence absolue de (19), car (16) peut être développée dans une série double à convergence absolue. Par la suppression de certains membres dans cette série on obtient la série (19).

	γ_2	γ_3	γ_k
Q, Q^*	0	0	0
$\overline{Q}$	$\frac{1}{2s}\sum_{i=1}^{s}(h-h_i)^2$	$\frac{1}{6s}\sum_{i=1}^{s}(h-h_i)^3$	$\frac{1}{k!\,s}\sum_{i=1}^{s}(h-h_i)^k$
P	$\frac{hd}{2}$	$-\frac{hd^2}{3}$	$\sum_{i=0}^{k}\frac{h^{k-i}}{(k-i)!}\binom{-\frac{h}{d}}{i}d^i$
$\overline{P}$	$\frac{\sum_{i=1}^{s}(h-h_i)^2}{2s} + \frac{\sum_{i=0}^{s}h_i d_i}{2s}$	$\frac{1}{6s}\sum_{i=1}^{s}(h-h_i)^3 - \frac{1}{3s}\sum_{i=1}^{s}h_i d_i^2 + \frac{1}{2s}\sum_{i=1}^{s}h_i d_i (h-h_i)$	$\sum_{i=1}^{k}\frac{h^{a-i}}{(k-i)!}\cdot\frac{1}{s}\sum_{\gamma=1}^{s}\binom{-\frac{h_\gamma}{d_\gamma}}{i}d_\gamma^i$
P^*	$\frac{1}{2}\sum_{i=1}^{s}h_i d_i$	$-\frac{1}{3}\sum_{i=1}^{s}h_i d_i^2$	$\sum\frac{A_2^{i_2}A_3^{i_3}\dots A_k^{i_k}}{i_2!\;i_3!\;\dots\;i_k!}$
			où $2i_2+3i_3+\dots+ki_k=k$, $A_k=\frac{(-1)^k}{k}\sum_{i=1}^{s}h_i d_i^{k-1}$, $i_2,\dots,i_k$ entiers >0

Exemple: Calcul de la constante γ_i pour la loi P_r.

$$H(z) = e^{-h(1-z)}\, e^{+h(1-z)} \left[1 + d\,(1-z)\right]^{-\frac{h}{d}} =$$

$$= e^{-h(1-z)} \sum_{r=0}^{\infty} \frac{1}{r!} h^r (1-z)^r \sum_{i=0}^{\infty} \binom{-\frac{h}{d}}{i} d^i (1-z)^i.$$

Multiplions les expressions à droite; le coefficient de $(1-z)^k$ sera γ_k. Mettons $r + i = k$ et exprimons l'égalité des coefficients de $(1-z)^k$.

Dans les autres cas, le procédé est pareil. Un peu plus compliqué est le calcul de γ_k pour la loi P^*, car ce calcul exige l'emploi du théorème des polynomes.

Si la non-homogénéité ou le degré de contagibilité est petite, les coefficients supérieurs γ_k sont eux-mêmes aussi petits. Dans ce cas on peut se borner dans le développement de Charlier aux 2 ou 3 membres. Ceci est possible d'autant plus que la dispersion et l'assymétrie dépendent uniquement de γ_2 et γ_3 comme nous allons le démontrer plus loin. On peut alors constater à cette loi approximative le déplacement du maximum à gauche (Eggenberger).

Voyons quelle est la différence entre la somme $\sum_{r=0}^{[h]} Q_r$ de la somme $\sum_{r=0}^{[h]} W_r$, si $\gamma_i = 0$ pour $i \geq 4$

$$\frac{G(z)}{1-z} - \frac{F(z)}{1-z} = \gamma_2 (1-z) F(z) + \gamma_3 (1-z)^2 F(z).$$

Exprimons l'égalité des coefficients de z^r:

$$\sum_{i=0}^{r} W_i - \sum_{i=0}^{r} Q_i = Q_r (\gamma_2 + \gamma_3) - Q_{r-1} (\gamma_2 + 2\gamma_3) + Q_{r-2}\gamma_3.$$

Remplaçons d'après la relation:

$$Q_{r-i} = \frac{r\,(r-1)\,\ldots\,(r-i+1)}{h^i} Q_r.$$

Supposons que $\gamma_i = 0$ pour $i \geq 4$. En effectuant les opérations, il vient:

$$\sum_{i=0}^{r} W_i - \sum_{i=0}^{r} Q_i = Q_r \left\{\gamma_2 \frac{h-r}{h} - \gamma_3 \frac{r-(h-r)^2}{h^2}\right\}.$$

Mettons $r = [h]$, $h - [h] = \varepsilon$, donc $0 \leq \varepsilon < 1$,

$$\sum_{i=0}^{[h]} W_i - \sum_{i=0}^{[h]} Q_i = Q_{[h]} \left\{-\frac{\gamma_3}{h} + \frac{\varepsilon}{h}\left(\gamma_2 + \frac{\gamma_3}{h} + \frac{\gamma_3 \varepsilon}{h}\right)\right\}.$$

De (21) on déduit:

1⁰ Si la grandeur de ε est de tel ordre qu'on peut supprimer le crochet, le déplacement ne dépend que de γ_3. $\gamma_3 < 0$ exprime l'acroissement, $\gamma > 0$ la diminution de la somme $\sum_{i=0}^{[h]}$.

2⁰ Si par contre $\varepsilon \doteq 1$, on peut écrire $\frac{\varepsilon}{h} \gamma_2 \doteq \frac{\gamma_2}{h}$. Ce terme peus alors avoir une importance considérable. Dans ce cas nous ne pouvont rien dire du déplacement d'une façon générale.

(Le résultat d'Eggenberger n'a pas donc une valeur précise.)

VIII. Le calcul des moments.

Une méthode connue du calcul des moments des factorielles a pour base la relation

$$Er\,(r-1)\ldots(r-k+1) = \left\{\frac{d^k}{dz^k}\,G(z)\right\}_{z=1}.$$

$G(z)$ est la fonction génératrice. Nous trouverons par cette méthode facilement:

loi de prob.	$k^{\text{ième}}$ moment des factorielles
Q, Q^*	h^k
$\overline{Q}$	$\frac{1}{s}\sum_{i=1}^{s} h_i^k$
P	$h\,(h+d)\ldots(h+[k-1]\,d)$
$\overline{P}$	$\frac{1}{s}\sum_{i=1}^{s} h_i\,(h_i+d_i)\ldots(h_i+[k-1]\,d_i)$

Le calcul des moments dans la loi P^* est beaucoup plus difficile.

$$H^*(z) = \prod_{i=1}^{s}\{1+d_i\,(1-z)\}^{-\frac{h_i}{d_i}}$$

$$\log H^*(z) = -\sum_{i=1}^{s}\frac{h_i}{d_i}\log\{1+d_i\,(1-z)\}.$$

Désignons la dérivée logarithmique de $H^*(z)$ par $K(z) = K$

$$K = \frac{H^{*\prime}}{H^*} = +\sum_{i=1}^{s}\frac{h_i}{1+d_i\,(1-z)},$$

$$H^{*\prime} = H^* \,.\, K,$$

$$H^{*(k)} = \sum_{i=0}^{k-1} \binom{k-1}{i} H^{*(i)} K^{(k-i-1)}, \tag{α}$$

$$K'(z) = \sum_{i=1}^{s} \frac{h_i d_i}{\{1 + d_i(1-z)\}^2}.$$

Par l'induction complète nous obtiendrons:

$$K^{(n)}(z) = \sum_{i=1}^{s} \frac{h_i d_i^n \, n!}{\{1 + d_i(1-z)\}^{n+1}}. \tag{β}$$

$$K^{(n)}(1) = n! \sum_{i=1}^{s} h_i d_i^n.$$

Mettons $H^{*(i)}(1) = N_i$ et introduisons cette expression dans (α). Il vient:

$$N_k = \sum_{i=0}^{k-1} \binom{k-1}{i} (k-i-1)! \sum_{\gamma=1}^{s} h_\gamma d_\gamma^{k-i-1} N_i. \tag{γ}$$

Mettons $\displaystyle \binom{k-1}{i} (k-i-1)! \sum_{\gamma=1}^{s} h_\gamma d_\gamma^{k-i-1} = \beta_{ki}.$

Ecrivons la relation (γ) pour $k = 1, 2, \ldots, k$ et effectuons la substitution $N_0 = 1$.

$$\left.\begin{array}{lr} 1 \,.\, N_1 & = \beta_{10} \\ -\beta_{21} N_1 + 1 \,.\, N_2 & = \beta_{20} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \ldots \\ -\beta_{k1} N_1 \; -\beta_{2k} N_2 - \ldots \; -\beta_{k,k-1} N_{k-1} + 1 N_k & = \beta_{k0}. \end{array}\right\}$$

En résolvant ce système nous obtiendrons le $k^{\text{ième}}$ moment

$$N_k = \begin{vmatrix} 1 & 0 & \ldots & 0 & \beta_{10} \\ -\beta_{21} & 1 & \ldots & 0 & \beta_{20} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -\beta_{k1} & -\beta_{k2} & \ldots & -\beta_{k,k-1} & \beta_{k0}. \end{vmatrix}$$

Pour calculer des moments l'un après l'autre nous appliquerons la relation

$$\beta_{ki} = \binom{k-1}{i} \beta_{k-i,0}, \quad \beta_{n0} = (n-1)! \sum_{\gamma=1}^{s} h_\gamma \, d_\gamma^{n-1}.$$

Nous obtiendrons p. ex.

$$N_1 = h, \; N_2 = \sum_{i=1}^{s} h_i d_i^2 + h^2.$$

Enfin, on peut démontrer, par le calcul des moments de la loi de probabilité (17), la relation qui existe entre les moments des factorielles et γ_i.

$$G(z) = e^{-h(1-z)} \sum_{\nu=0}^{\infty} \gamma_\nu (1-z)^\nu$$

$$G^{(k)}(z) = \left\{ e^{-h(1-z)} + \sum_{\nu=0}^{\infty} \gamma_\nu (1-z)^\nu \right\}^{(k)} =$$

$$= \sum_{i=0}^{\infty} \binom{k}{i} \frac{d^{k-i}}{dz^{k-i}} \left(e^{-h(1-z)} \right) \cdot \frac{d^i}{dz^i} \left(\sum_{\nu=0}^{\infty} \gamma_\nu (1-z)^\nu \right).$$

Nous obtiendrons:

$$\frac{d^i}{dz^i} \sum_{\nu=0}^{\infty} \gamma_\nu (1-z)^\nu = (-1)^i \,.\, i! \sum_{\nu=i}^{\infty} \binom{\nu}{i} \gamma_\nu (1-z)^{\nu-i},$$

$$\left[\frac{d^i}{dz^i} \sum_{\nu=0}^{\infty} \gamma_\nu (1-z)^\nu \right]_{z-1} = (-1)^i \, i! \, \gamma_i.$$

Par conséquent le $k^{\text{ième}}$ moment des factorielles de la loi générale $G(z)$ est

$$G^{(k)}(1) = \sum_{i=0}^{k} (-1)^i \, k(k-1) \ldots (k-i+1) \, h^{k-i} \gamma_i. \tag{21}$$

De cette relation on peut calculer les moments des puissances:

$$E\,(r-h)^2 = \sigma^2 = h + 2\gamma_2, \qquad E\,(r-h)^3 = h + 6\gamma_2 - 6\gamma_3.$$

D'une façon générale γ_z dépend uniquement des moments dont le rang est inférieur ou égal à k.

De ce résultat nous pouvons tirer tout de suite deux conséquences; d'une part dans toutes les lois considérées γ_2 est supérieur à zéro, donc la dispersion est hypernormale; d'autre part le nombre mesurant l'assymétrie $\varkappa = \dfrac{h + 6\gamma_2 - 6\gamma_3}{(h + 2\gamma_2)^{3/2}}$ dépend de γ_3 de façon que $\gamma_3 > 0$ fait diminuer et $\gamma_3 < 0$ fait augmenter l'assymétrie.

Si γ_2 reste constant, $|\gamma_3|$ ett plus petit pour P que pour P^*. (Démonstration par l'inégalité de Cauchy). L'inégalité $\gamma_3 < 0$ donne lieu au théorème suivant:

La non-homogénéité interne dans l'acroissement des probabilités mène à l'agrrandissement de l'assymétrie.

IX. La non-homogénéité et la contagibilité.

La non-homogénéité et la contagibilité ont le même effet, si leur influence est faible. Car dans tous les deux cas on peut employer la loi approximative

$$W_\nu = Q_\nu + \gamma_2 \Delta^2 Q_\nu, \quad \gamma_2 > 0$$

expression qui n'est autre chose que le commencement du développement de Charlier, si nous nous bornons à deux membres.

Pour différencier les deux cas, Eggenberger a employé la méthode connue de la combinaison des séries. — Ecrivons l'écart quadratique moyen de la loi $\overline{Q_\nu}$ sous la forme de Bortkiewicz:

$$\sigma^2 = h\left\{1 + \frac{1}{hs}\sum_{i=1}^{s}(h - h_i)^2\right\} = h\,(1 + \lambda^2),$$

où λ désigne l'excédent relatif des erreurs. Si nous faisons augmenter le nombre des expériences dans chaque série k-fois (les conditions à l'intérieur des séries restant les mêmes), nous constatons que chacune des valeurs h_γ augmente k-fois, de sorte que l'expression $h = \frac{1}{s}\sum_{i=1}^{s} h_i$ augmente elle-même k-fois, en supposant toujours que le nombre des séries s ne change pas. (Eggenberger ne parle pas de cette hypothèse). Nous voyons que la valeur λ^2 est directement proportionelle à la longueur de la série k. Si σ_k appartient à s séries k-fois prolongées il vient

$$\sigma_k = kh\,(1 + k\lambda^2), \tag{22}$$

σ_k^2 est la fonction quadratique de k.

Le prolongement est appelé continu si nous combinons les séries de telle façon que chaque série (schéma de l'urne) reprend l'état de l'urne dans lequel il se trouvait à la fin de la série précédente. A cette façon de combiner les séries répond la combinaison des séries voisines dans le temps, si les séries sont les segments d'une série temporelle. Dans le cas du prolongement discontinu la série ajoutée commence dans les mêmes conditions que la série précédente; ceci a lieu pour les séries éloignées entre elles dans le temps.

S'il s'agit du prolongement discontinu dans la loi de Polya, l'écart quadratique moyen des séries k-fois prolongées est évidemment la fonction linéaire de k:

$$\sigma_k^2 = k\,.\,\sigma^2.$$

S'il s'agit du prolongement continu nous avons le même résultat que dans le cas d'une série, seulement au lieu de h, il y aura $kh = \lim_{n\varrho \to h} kn\,.\,\varrho$, au lieu de d $kd = \lim_{n\delta = d} kn\delta$. Donc σ_k^2 est la fonction quadratique de k:

$$\sigma^2 = kh\,(1 + kd) = k\,\frac{1 + kd}{1 + d}\,.\,\sigma^2.$$

Nous résumons les résultats dans le tableau suivant:

prolongement continu	σ_k est fonction lin. de k	l'accroissement de prob., prolongement continu
	σ_k est fonction quadr. de k	l'accr. de prob. prol. discontinu
prol. discontinu	σ_k est fonction lin. de k	l'accr. de prob.
	σ_k est fonction quadr. de k	non-homogénéité

C'est Eggenberger qui a résolu le problème de la façon précédente. Ses exemples pratiques montrent qu'il ne fait pas la supposition que le nombre des séries dans le cas de la non-homogénéité est le même avant et après la combinaison des séries. Nous allons montrer que dans certaines conditions son procédé non précis a une valeur pour l'approximation.

Soit s divisible par m, $s/m = t$. Introduisons la double notations pour les valeurs moyennes:

$$\{h_1, \ldots, h_t; h_{t+1}, \ldots, h_{2t}; \ldots; h_{t(m-1)+1}, \ldots, h_{tm}\} =$$
$$= \{h_1^{(1)},, \ldots, h_t^{(1)}; h_1^{(2)}, \ldots, h_t^{(2)}; \ldots; h_1^{(m)}, \ldots, h_t^{(m)}\}.$$

Nous combinerons m séries dans une série nouvelle de sorte que

$$h_k^{(1)} + \ldots + h_k^{(m)} = \overline{h}_k \quad k = 1, 2, \ldots, t,$$

$$h = \frac{1}{s}\sum_{i=1}^{s} h_i = \frac{1}{s}\sum_{i=1}^{t}(h_i^{(1)} + h_i^{(2)} + \ldots + h_i^{(m)}) = \frac{1}{m}\cdot\frac{m}{s}\sum_{i=1}^{t}\overline{h}_i = \frac{1}{m}h_{(m)},$$

$$\sum_{i=1}^{s}(h - h_i)^2 = \sum_{i=1}^{t}\left\{(h - h_i^{(1)})^2 + (h - h_i^{(2)})^2 + \ldots + (h - h_i^{(m)})^2\right\},$$

$$(h_{(m)} - h_i)^2 = \{mh - h_i^{(1)} - \ldots - h_i^{(m)}\}^2 =$$
$$= \sum_{k=1}^{m}(h - h_i^{(k)})^2 + 2\sum\sum_{k \neq j}(h - h_i^{(k)})(h - h_i^{(j)}).$$

Puisque $\sum_{i=1}^{t}\sum_{k=1}^{m}(h - h_i^{(k)})^2 = \sum_{i=1}^{s}(h - h_i)^2$, par l'addition pour $i = 1, \ldots, t$ il vient:

$$\sum_{i=1}^{t}(h_{(m)} - \overline{h}_i)^2 = \sum_{i=1}^{s}(h - h_i)^2 + 2\sum_{i=1}^{t}\left\{\sum\sum_{k \neq j}(h - h_i^{(k)})(h - h_i^{(j)})\right\}. \quad (23)$$

1⁰ Supposons que $h_i^{(1)} = h_i^{(2)} = \ldots = h_i^{(m)} = h_i$ pour $i = 1, \ldots, t$. Nous pouvons écrire $(h^{(m)} - \overline{h}_i)^2 = m^2 (h - h_i)^2$, donc

$$\sum_{i=1}^{t}(h^{(m)} - \overline{h}_i)^2 = m^2 \sum_{i=1}^{t}(h - h_i)^2 = m \sum_{i=1}^{s}(h - h_i)^2,$$

$$\lambda_{(m)}{}^2 = \frac{1}{h_{(m)} \cdot t} \sum_{i=1}^{t}(h_{(m)} - \overline{h}_i)^2 = \frac{m}{hs} \sum_{i=1}^{s}(h - h_i)^2 = m\lambda^2, \ h_{(m)} = mh. \quad (24)$$

Donc

$$\sigma_{(m)}{}^2 = h_{(m)}\,(1 + \lambda_{(m)}{}^2) = mh\,(1 + m\lambda^2),$$

$\sigma_m{}^2$ est bien la fonction quadratique de m.

2⁰ Si la condition 1⁰ n'est pas remplie la relation (24) n'est pas accomplie elle-même non plus. Car le membre à droite de (23) est la somme des expressions qui peuvent être des signes différents et qui peuvent se détruire de sorte que la valeur de ce membre peut être très petite. Dans ce cas nous aurions

$$\sum_{i=1}^{t}(h_{(m)} - \overline{h}_i)^2 = \sum_{i=1}^{s}(h - h_i)^2, \ \ \overline{\lambda^2} = \lambda^2,$$

donc

$$\sigma_m{}^2 = mh\,(1 + \lambda^2),$$

c'est à dire la fonction linéaire de m, ce qui caractérise l'accroissement des probabilités. Il en résulte la règle suivante: Nous pouvons combiner seulement des séries dont les espérances mathématiques sont approximativement égales.

X. Application aux collectifs des décès à cause des maladies contagieuses.

Polya et Eggenberger citent quelques exemples traitant les collectifs des accidents et des décès à cause des maladies d'infection. Nous allons contrôler leurs résultats au moyen des données de la statistique tchécoslovaque.

1⁰ Scarlatine.

Eggenberger, en employant les données de la statistique suisse, est arrivé à la conclusion que la différence entre la théorie et la pratique dans le cas de la scarlatine peut être expliquée par la non-homogénéité causée par la différence entre la mortalité dans la première et dans la seconde moitié de l'année. Pour contrôler ce résultat dressons le tableau suivant:

Nombres de décès par suite de scarlatine. 1921—29:

	janvier	février	mars	avril	mai	juin	juillet	août	sept.	oct.	nov.	déc.	moyenne
hommes ..	91	67	75	70	65	41	43	32	49	84	96	96	67·6
femmes...	96	65	74	56	67	53	44	35	54	87	83	86	66·6

Les mois de la saison d'été (avril — sept.) ont une mortalité inférieure à la moyenne; par contre les mois de la saison d'hiver ont une mortalité supérieur à la moyenne. Pour pouvoir examiner l'influence de cette non-homogénéité de série en série, nous sommes obligés de représenter à part les mois d'été et les mois d'hiver. Nous citerons les résultats totaux $\overline{P}_r$.

Scarlatine hommes 1921—29

r	f_r	f^*_r	$10^8 P_r$	$10^8 \overline{P}_r$	$10^8 Q_r$
0	4	2·7	1·6	1·9	0·1
1	5	3·9	4·2	4·5	0·5
2	2	5·3	6·7	6·9	1·7
3	8	8·2	8·8	8·7	4·2
4	12	10·2	10·0	9·7	7·9
5	17	11·0	10·5	10·0	11·8
6	4	9·2	10·2	9·8	14·8
7	8	8·3	9·5	9·2	15·8
8	9	7·4	8·5	8·4	14·8
9	8	7·0	7·4	7·4	12·4
10	4	5·9	6·3	6·3	9·3
11	5	5·0	5·2	5·3	6·3
12	6	4·2	4·3	4·4	4·0
13	2	3·7	3·4	3·5	2·3
14	2	3·2	2·7	2·8	1·2
15	6	3·0	2·1	2·2	0·6
16	1	2·2	1·6	1·7	0·3
17	1	1·5	1·3	1·3	0·1
18	3	0·7	1·0	1·0	0·1
19	0	0·4	0·7	0·7	0·0
20	0	0·3	0·5	0·5	0·0
21	1	0·3	0·4	0·4	0·0

femmes

r	f_r	f^*_r	$10^8 P_r$	$10^8 Q_r$
0	6	5·0	1·7	0·1
1	3	5·0	4·4	0·5
2	6	6·1	7·0	1·8
3	7	7·3	9·0	4·5
4	11	8·8	10·1	8·3
5	8	8·9	10·5	12·2
6	10	9·1	10·2	15·1
7	7	8·7	9·4	15·9
8	11	8·7	8·4	14·7
9	7	7·8	7·3	12·1
10	7	6·9	6·2	9·0
11	6	5·7	5·1	6·0
12	4	4·6	4·2	3·7
13	4	3·6	3·3	2·1
14	2	2·7	2·6	1·1
15	2	2·1	2·1	0·6
16	2	1·7	1·6	0·3
17	1	1·3	1·2	0·1
18	1	1·2	1·0	0·1
19	1	1·2	0·8	0·0
20	2	1·3	0·5	0·0

$h' = 7{\cdot}4907$, $d' = 1{\cdot}8808$, $A' = 3{\cdot}77$, $A_P = 3{\cdot}66$

Puisque la durée de 9 ans est relativement courte, il n'est pas surprenant que les valeurs de f_r balancent considérablement. Pour faciliter le passage en revue nous citons dans la colonne 3 les valeurs f^*_r ajustées au moyen de la formule

$$f^*_r = \tfrac{1}{9}(f_{r-2} + 2f_{r-1} + 3f_r + 2f_{r+1} + f_{r+2}).$$

La concordance de la théorie avec les résultats empiriques est satisfaisante et montre la justesse des résultats de la théorie (maximum, asymétrie, les valeurs extrèmes). En comparant les nombres P_r et $\overline{P}_r$ on constate que l'influence de la non-homogénéité est sans importance et qu'elle peut être supprimée. Plus précisément la loi de Polya est sensible même à la non-homogénéité de série en série, qu'elle décompense automatiquement.

Toutes ces conclusions sont les mêmes pour la mortalité des femmes. Ici la concordance est encore meilleure. Nous pouvons donc conclure que les nombres mensuels des décès causés par la scarlatine sont régis par la loi de Polya.

2⁰ La rougeole en Bohème 1921—29, hommes.

Vu la place restreinte, le tableau est réduit.

r	0—1	2—3	4—5	6—7	8—9	10—11	12—13	14—15	16—17	18—19	20—21	22—23	24—25	26—27	28—53
f_r	4	11	9	11	12	16	7	4	7	4	0	5	5	1	12
f^*_r	5.1	8·4	9·8	11·4	14·4	12·4	8·9	5·7	4·2	3·6	3·2	3·5	3·4	2·4	11·6
$P_r.108$	7.5	10·3	10·8	10·4	9·6	8·6	7·6	6·6	5·7	4·9	4·2	3·6	3·0	2·5	12·7
$Q_r.108$	0.0	0·1	0·6	3·2	9·2	12·2	22·6	21·9	16·3	9·6	4·6	1·8	0·6	0·2	0·1

$h' = 13{\cdot}7\dot{5}9\dot{2}$, $d' = 8{\cdot}9466$, $A' = 8{\cdot}46$, $A_P = 8{\cdot}85$, $A_Q = 2{\cdot}96$

La deuxième colonne comprend les fréquences ajustées. Remarquons que la loi de Poisson place le maximum trop à droite et en haut, tandis que la loi de Polya conduit à une faute dans le sens inverse.

Nous pouvons supposer que la courbe idéale de la mortalité causée par la rougeole en Bohème ressemble dans sa forme grossièrement à la courbe de Polya, mais pourtant quelle différe de celle-ci d'une façon caractéristique.

3⁰ La diphtérie.

Nous constatons ici encore ce fait important que la même différence que dans le cas précédent apparaît ici et d'une façon encore plus prononcée. (V. tab. p. 22.)

En tenant compte des résultats de VIII il est impossible d'expliquer dans ce cas cette différence entre la théorie et la pratique par la non-homogénéité interne car le déplacement se produit dans le sens inverse. De même la non-homogénéité de série en série ne peut élever le maximum. D'ailleurs nous avons montré dans le cas de la scarlatine, l'influence insignificante de cette non-homogénéité.

Conclusion. La méthode inventée par Polya donne l'exemple de l'application du calcul de probabilité à la statistique. Le but de ces méthodes diffère de celui des méthodes exclusivement descriptives (p. ex. la méthode de Charlier). Dans l'application d'un schéma nous faisons des hypothèses sur les causes des faits et nous les vérifions. Le but terminal est d'examiner la structure du fait et de trouver la loi statistique.

Si nous trouvons une discordance (comme p. ex. dans le cas de la diphtérie) il s'agit de l'expliquer. Ces discordances peuvent être causées p. ex.:

La diphtérie en Bohème 1921—29, femmes.

I	II	III	IV	V	VI
r	f_r	$108\,P_r$	Σf_r	$108\,\Sigma P_r$	V—IV
0	1	4·7	1	4·7	+3·7
1	1	5·1	2	9·9	+7·9
2	3	5·2	5	15·1	+10·1
3	4	5·1	9	20·2	+11·2
4	7	5·0	16	25·1	+9·1
5	6	4·8	22	29·9	+7·9
6	9	4·6	31	34·4	+3·4
7	10	4·4	41	38·8	—2·2
8	7	4·1	48	42·9	—5·1
9	5	3·9	53	46·8	—6·2
10	3	3·7	56	50·6	—5·5
11	4	3·5	60	54·1	—5·9
12	1	3·3	61	57·4	—3·6
13	3	3·1	64	60·5	—3·5
14	3	3·0	67	63·5	—3·5
15	2	2·8	69	66·2	—2·8
16	3	2·6	72	68·9	—3·1
17	0	2·5	72	71·3	—0·7
18	2	2·3	74	73·6	—0·4
19	4	2·2	78	75·8	—2·2
20	2	2·0	80	77·8	—2·2
21	2	1·9	82	79·8	—2·2
22	1	1·8	83	81·6	—1·4
23	1	1·7	84	83·3	—0·8
24	2	1·6	86	84·8	—1·2
25	1	1·5	87	86·3	—0·7
26	3	1·4	90	87·7	—2·3
27	0	1·3	90	89·0	—1·0
28	0	1·2	90	90·2	+0·2
29	3	1·1	93	91·3	—1·7
30	1	1·1	94	92·4	—1·6
31	0	1·0	94	93·4	—0·6
32	2	0·9	96	94·3	—1·7
33	0	0·9	96	95·2	—0·8
34	0	0·8	96	96·0	+0·0
35	1	0·8	97	96·8	—0·2
36	0	0·7	97	97·5	+0·5
37	1	0·7	98	98·2	+0·2
38	1	0·6	99	98·8	—0·2
39	0	0·6	99	99·4	+0·4
40	0	0·6	99	100·0	+1·0
41	1	0·5	100	100·5	+0·5
42	1	0·5	101	101·0	—0·0
43	1	0·5	102	101·4	—0·6
44	2	0·4	104	101·9	—2·2
.	.	.	.	.	.
54	1	0·2	105	104·9	—0·1
55	0	0·2	105	105·1	+0·1
56	1	0·2	106	105·2	—0·8
.	.	.	.	.	.
70	1	0.1	107	107·0	—0·1
.	.	.	.	.	.
93	1	0·01	108	107·9	—0·1

$h' = 15{\cdot}8796$, $d' = 13{\cdot}6274$ $A' = 10{\cdot}87$, $A = 11{\cdot}34$.

1° Par le hasard et elles disparaîtront dans un temps plus ou moins long.

2° La mortalité est régie au fond par la loi de Polya mais par les dispositions spéciales dans le cas de grandes épidémies la courbe des fréquences est modifiée artificiellement.

3° La mortalité à cause du crou est régie par une autre loi que par celle de Polya.

A première vue il semble que l'explication 2° n'est pas justifiée, car les dispositions spéciales modifient les fréquences extrêment grandes. Par contre il faut remarquer qu'une plus grande discordance des valeurs extrèmes ne peut pas avoir lieu, car nous ajustons les deuxièmes moments, sur lesquels justement les valeurs extrèmes ont la plus grande influence. On peut dire même que l'intervention sur n'importe quel endroit se manifeste toujours de la façon la plus prononcée à l'endroit le plus sensible, c'est à dire aux environs de la moyenne.

Nous ne pouvons décider si ces explications suffisent à éclaircir les divergences considérées. Cette possibilité sera montrée par les données des années prochaines. Mais il est probable que la loi de Polya n'est pas une loi universelle régissant les faits corrélatifs. C'est qu'il est facile de construire des schémas encore plus généraux des faits corrélatifs et qui semblent être justifiés au même titre que le schéma de Polya, mais que nous ne savons pas interpréter mathématiquement.

Zum versicherungstechnischen Aufbau des neuen Pensionsversicherungsgesetzes in der Tschechoslowakei.

Von Prof. Dr. E. Schoenbaum.

Unter diesem Titel hat Herr Prof. Dr. G. Rosmanith im sechsten Hefte der unter seiner Redaktion erscheinenden Zeitschrift „Versicherungswissenschaftliche Mitteilungen" einen Aufsatz veröffentlicht, welcher sich mit dem versicherungstechnischen Aufbau des neuen tschechoslowakischen Pensionsversicherungsgesetzes beschäftigt. Die Abhandlung enthält außer wertvollen Anregungen einige meistens auf Mißverständnissen beruhende kritische Bemerkungen, welche teilweise bereits durch die inzwischen in den Aktuárské Vědy erschienene Abhandlung von Herrn Dr. Havlík[1]) eine Beantwortung gefunden haben. Trotzdem halte ich es im Interesse der Mathematik der Sozialversicherung für erforderlich, mich mit einigen durch Herrn Prof. Rosmanith aufgeworfenen Fragen zu beschäftigen.

I.

Herr Prof. Dr. Rosmanith kritisiert mit Recht den Anspruch auf die unbedingte Altersrente. In der Tat gehört die unbedingte Leibrente gar nicht in die Invaliden- und Altersversicherung der Arbeiter und Angestellten. Es wird ihm jedoch wohl bekannt sein, mit welcher Hartnäckigkeit die Dienstnehmer diese aus dem alten österreichischen Pensionsversicherungsgesetze stammende Leistung verteidigt haben, und daß alle Versuche, sie durch eine eventuell zu einem früheren Alter fällige bedingte Arbeitslosen-Altersrente zu ersetzen, welche ich für sozialpolitisch berechtigter als die vom Verfasser vorgeschlagene halte, erfolglos geblieben sind. Die Erfahrungen der letzten Jahre haben

[1]) Dr. V. Havlík. Die Reform der čechoslovakischen Pensionsversicherung der Privatangestellten in höheren Diensten und ihre Deckung. Aktuárské vědy I. Jahrgang, Nr. 3., II. Jahrg. Nr. 1.