ANNALES
UNIVERSITATIS MARIAE CURIE - SKŁODOWSKA
LUBLIN - POLONIA

VOL. LI. 1, 16 SECTIO A 1997

JÓZEF L. TEUGELS (Leuven)

On the Pólya–Eggenberger Distribution and its Applicability in Insurance Mathematics

ABSTRACT. Limiting cases of the Pólya–Eggenberger distribution are discussed and some links of PED to actuarial mathematics are given.

1. Introduction. The Pólya-Eggenberger urn model has always been a favourite illustration of elementary probabilistic concepts. In this modest contribution we show that a variety of classical distributions appear naturally as special cases or approximations of the Pólya-Eggenberger distribution. We also show that the distribution might be used as a model for the claim number distribution in insurance portfolios. The connection with recursive schemes, popular among insurance mathematicians, will be highlighted. This paper is dedicated to my friend Dominik Szynal on the occasion of his sixtieth birthday.

2. The Pólya-Eggenberger distribution. Denote by N_1 the initial number of red balls in an urn while $N_2 = N - N_1$ denotes the number of green balls. At each of m successive trials a ball is chosen at random from the urn; after denoting its colour, the ball drawn is added to the urn together with Δ balls of the same colour as the ball drawn.

If R_m denotes the number of red balls after these m trials then the distribution of the random variable R_m is given by the so-called *Pólya-*

Eggenberger distribution (PED), $0 \le n \le m$,

$$p_n := P\{R_m = n\} = \binom{m}{n} \frac{\prod_{i=1}^{n-1}(N_1 + i\Delta)\prod_{j=0}^{m-n-1}(N_2 + j\Delta)}{\prod_{k=0}^{m-1}(N + k\Delta)}, \tag{1}$$

where we assume that $(m-1)\Delta + N > 0$.

For a survey of the properties of PED we refer to section 6.2.4 in Johnson e.a. (1992) where the reference to Eggenberger-Pólya (1923) as a model for contagious distributions is clarified. We refer to this publication for historical information as well as for further references.

See also Berg's contribution in Kotz e.a. (1988), Jordan (1927) and Bricas (1949) where the connection with the Pearson system has been treated in great detail.

A few typical cases are obtained by taking $\Delta \in \{0, -1, +1\}$:

- For $\Delta = 0$ we obtain the classical case of *trias with replacement*. The resulting r.v. will be denoted by $B_m(p)$ since

$$p_n = P\{B_m(p) = n\} = \binom{m}{n} p^n (1-p)^{m-n}$$

is the *binomial distribution* with parameter $p = N_1/N$.

- For $\Delta = -1$ we recover *trials without replacement* governed by the *hypergeometric distribution*

$$p_n = \frac{\binom{N_1}{n}\binom{N_2}{m-n}}{\binom{N}{m}}.$$

- For $\Delta = 1$ we obtain the *negative hypergeometric distribution*

$$p_n = \frac{\binom{-N_1}{n}\binom{-N_2}{m-n}}{\binom{-N}{m}}.$$

From now on we assume that $\Delta \neq 0$ since the case where $\Delta = 0$ is standard. We introduce new parameters

$$\lambda = \frac{N_1}{\Delta}, \quad \mu = \frac{N_2}{\Delta}, \quad \nu = \lambda + \mu = \frac{N}{\Delta}$$

so that R_m essentially depends on m, λ and μ. A simple calculation shows that

$$p_n = P\{R_m(\lambda, \mu) = n\} = \binom{m}{n} \frac{B(\lambda + n, \mu + m - n)}{B(\lambda, \mu)}, \tag{2}$$

where $B(\alpha,\beta)$ is the classical beta-function. Note that whenever necessary, we will clearly indicate the dependence of $R_m := R_m m(\lambda,\mu)$ on the newly introduced parameters. An alternative way of writing the PED is then

$$p_n = \frac{\binom{-\lambda}{n}\binom{-\mu}{m-n}}{\binom{-\nu}{m}}.$$

The generating function of the PED is well-known

$$E\{z^{R_m(\lambda,\mu)}\} = \frac{F(\lambda,-m,1-\mu-m;z)}{F(\lambda,-m,1-\mu-m;1)}, \tag{3}$$

where $F(\alpha,\beta,\gamma;x)$ is the familiar hypergeometric function. Replacing z by $\exp(it)$ we get the *characteristic function* that will prove to be useful as well. Relation (3) can be used to obtain closed expressions for the *factorial moments* of $R_m(\lambda,\mu)$ for $k \geq 0$

$$\alpha_k(R_m(\lambda,\mu)) := E\left\{\frac{R_m(\lambda,\mu)!}{(R_m(\lambda,\mu)-k)!}\right\} = \frac{m!}{(m-k)!}\,\frac{\Gamma(\lambda+k)}{\Gamma(\lambda)}\,\frac{\Gamma(\nu)}{\Gamma(\nu+k)} \tag{4}$$

and consequently for the *factorial moment ratio*

$$c_k(R_m(\lambda,\mu)) := \frac{\alpha_{k+1}(R_m(\lambda,\mu))}{\alpha_k(R_m(\lambda,\mu))} = (m-k)\,\frac{\lambda+k}{\nu+k}, \quad k \geq 0. \tag{5}$$

The ordinary *moments* can then be expressed in terms of the *Stirling numbers of the second kind* defined for ≥ 1 and $1 \leq m \leq r$ by

$$x^r = \sum_{m=1}^{r} \left\{ {r \atop m} \right\} x(x-1)\ldots(x-m+1) \tag{6}$$

so that

$$\mu_k(R_m(\lambda,\mu)) := E\{R_m^k(\lambda,\mu)\} = \sum_{j=1}^{k} \left\{ {k \atop j} \right\} \alpha_j(R_m(\lambda,\mu)). \tag{7}$$

In particular we easily derive that $E\{R_m(\lambda,\mu)\} = (m\lambda)/\nu$ and $\mathrm{Var}\{R_m(\lambda,\mu)\} = \{(m\lambda\mu)/(\nu^2)\}\{(\nu+m)/(\nu+1)\}$.

3. Limits and approximations. The above urn scheme admits of a large number of limits and approximations when the four available parameters N_1, N, m and Δ take particular boundary values. An attempt to cover a wide variety of cases is due to [4] although many special cases can be found in standard textbooks on statistics. See also [3].

We restrict attention to non degenerate limits and approximations. We will show that the following discrete and continuous distributions appear as limits or approximations for $R_m(\lambda,\mu)$.

- The *binomial variable* $B_m(p)$ mentioned above. Recall that its characteristic function is given by $(pe^{it}+q)^m$ and that the factorial moment ratio is $c_k(B_m(p)) = (m-k)p$.
- The *Poisson variable* $P(\tau)$ with distribution

$$P\{P(\tau)=n\} = e^{-\tau}\,\frac{\tau^n}{n!}\,,\quad n \geq 0\,,$$

with characteristic function $\exp(\tau(e^{it}-1))$ and factorial moment ratio $c_k(P(\tau)) = \tau$.

- The *negative binomial* or *Pascal variable* $Y(\lambda,\rho)$ with distribution

$$P\{Y(\lambda,\rho)=n\} = \binom{\lambda+n-1}{n}\rho^n(1-\rho)^\lambda\,,\quad n \geq 0\,,$$

with characteristic function $\left\{1+\rho(1-e^{it})/(1-\rho)\right\}^{-\lambda}$ and factorial moment ratio $c_k(Y(\lambda,\rho)) = (\lambda+k)\rho/(1-\rho)$.

- The *gamma variable* $G(\lambda)$ with distribution

$$P\{G(\lambda)\leq x\} = \frac{1}{\Gamma(\lambda)}\int_0^x e^{-t}t^{\lambda-1}dt\,,\quad x \geq 0\,,$$

with characteristic function $(1-it)^{-\lambda}$ and moments $\mu_k(G(\lambda)) = \Gamma(\lambda+k)/\Gamma(\lambda)$.

- The *beta variable* $B(\lambda,\mu)$ with distribution

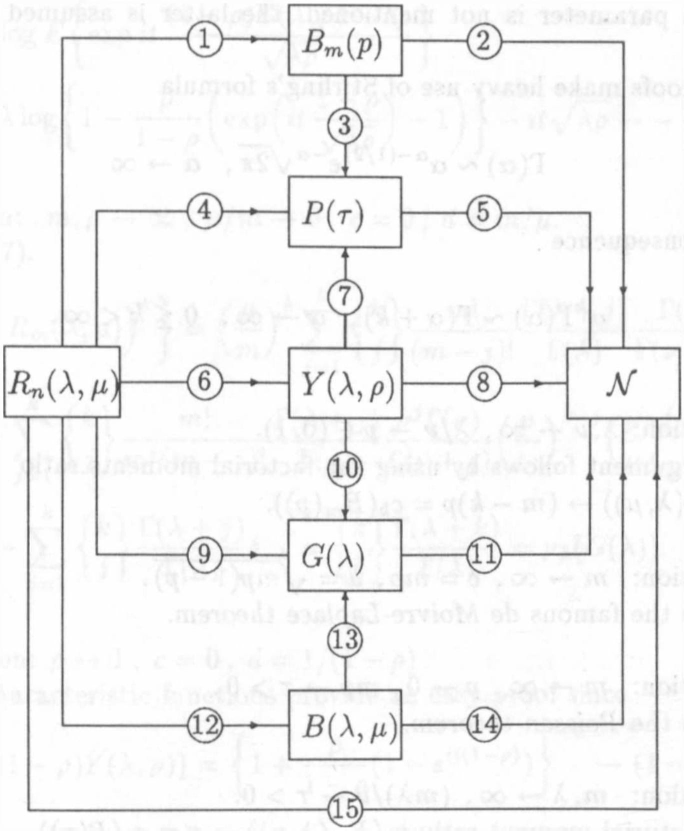
$$P\{B(\lambda,\mu)\leq x\} = \frac{1}{B(\lambda,\mu)}\int_0^x u^{\lambda-1}(1-u)^{\mu-1}du\,,\quad 0 \leq x \leq 1$$

and with moments $\mu_k(B(\lambda,\mu)) = \{\Gamma(\lambda+k)\Gamma(\nu)\}\{\Gamma(\lambda)\Gamma(\nu+k)\}$.

- The *standard normal variable* $\mathcal{N}$ with distribution

$$P\{\mathcal{N}\leq x\} = \Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-y^2/2}dy\,,\ -\infty < x < \infty$$

and with characteristic function $\exp(-t^2/2)$.



For a number of the limits we need a normalisation before the limit in distribution should be evaluated. When a r.v. Z has to be normalized, we will write this in the form $(Z - c)/d$ and clarify the explicit expressions for c and d in terms of the parameters λ, μ and m.

To indicate the interdependence of the PED with the above mentioned distributions we display them in a layout. The accompanying numbers refer to the parts of the proofs in the next section. Here the limiting operations and the eventual normalizations are spelled out explicitly for each of the possible approximating distributions.

4. Proofs. Most of the proofs in the next section are very classical. We nevertheless provide outlines for all parts of the proof using a variety of different methods: moment convergence, convergence of characteristic functions or immediate convergence in distribution. The theoretical results needed to underbuild the proofs can be found i.a. in [2].

For each case we formulate the explicit conditions and normalizations. In case a parameter is not mentioned, the latter is assumed to be kept constant.

The proofs make heavy use of Stirling's formula

$$\Gamma(\alpha) \sim \alpha^{\alpha-(1/2)} e^{-\alpha}\sqrt{2\pi}\,, \quad \alpha \to \infty \tag{8}$$

and its consequence

$$\alpha^k \Gamma(\alpha) \sim \Gamma(\alpha+k)\,, \quad \alpha \to \infty\,, \quad 0 \le k < \infty\,. \tag{9}$$

1. Condition: $\lambda, \nu \to \infty\,,\ \lambda/\nu \to p \in (0,1)$.
   The argument follows by using the factorial moments ratio
   $c_k(R_m(\lambda,\mu)) \to (m-k)p = c_k(B_m(p))$.

2. Condition: $m \to \infty\,,\ c = mp\,,\ d = \sqrt{mp(1-p)}\,.$
   This is the famous *de Moivre-Laplace theorem.*

3. Condition: $m \to \infty\,,\ p \to 0\,,\ mp \to \tau > 0\,.$
   This is the *Poisson theorem.*

4. Condition: $m, \lambda \to \infty\,,\ (m\lambda)/\nu \to \tau > 0.$
   The factorial moment ratio $c_k(R_m(\lambda,\mu)) \to \tau = c_k(P(\tau))$
   which provides the required link.

5. Condition: $\tau \to \infty\,, c = \tau\,,\ d = \sqrt{\tau}.$
   Use characteristic functions (or their logaritms)

$$\log E\left\{\exp it\frac{P(\tau)-\tau}{\sqrt{\tau}}\right\} = -it\sqrt{\tau} + \tau\left(e^{it\tau^{-1/2}} - 1\right) \to -\frac{t^2}{2}\,.$$

6. Condition: $m, \mu \to \infty\,,\ m/\mu \to \rho/(1-\rho)\,, 0 < \rho < 1.$
   Again the factorial moment ratio can be used since
   $c_k(R_m(\lambda,\mu)) \to [\rho/(1-\rho)](\lambda + k) = c_k(Y(\lambda,\rho))$.

7. Condition: $\rho \to 0,\ \lambda \to \infty\,,\ \rho\lambda \to \tau > 0.$
   Clearly $c_k(Y(\lambda,\rho)) \to \tau = c_k(P(\tau))$.

8. Condition: $\lambda \to \infty\,,\ \rho \to 1\,,\ c = \lambda\rho/(1-\rho)\,,\ d = \sqrt{\lambda\rho}/(1-\rho)$.

Use characteristic functions

$$\log E\left\{\exp it\,\frac{(1-\rho)Y(\lambda,\rho)-\lambda\rho}{\sqrt{\lambda\rho}}\right\}$$

$$= -\lambda\log\left\{1-\frac{\rho}{1-\rho}\left(\exp\left(it\,\frac{1-\rho}{\sqrt{\lambda\rho}}\right)-1\right)\right\}-it\sqrt{\lambda\rho}\to-\frac{t^2}{2}.$$

9. Condition: $m,\mu\to\infty$, $\mu/m\to 0$, $c=0$, $d=m/\mu$.
We use (7).

$$E\left\{\left(\frac{\mu}{m}R_m(\lambda,\mu)\right)^k\right\}=\left(\frac{\mu}{m}\right)^k\sum_{j=1}^{k}\left\{{k \atop j}\right\}\frac{m!}{(m-j)!}\,\frac{\Gamma(\lambda+j)}{\Gamma(\lambda)}\,\frac{\Gamma(\nu)}{\Gamma(\nu+j)}$$

$$=\sum_{j=1}^{k}\left\{{k \atop j}\right\}\frac{m!}{m^j(m-j)!}\,\frac{\Gamma(\lambda+j)}{\Gamma(\lambda)}\,\frac{\nu^j\Gamma(\nu)}{\Gamma(\nu+j)}\left(\frac{\mu}{m}\right)^{k-j}\left(\frac{\mu}{\nu}\right)^{j}$$

$$\to\sum_{j=1}^{k}\left\{{k \atop j}\right\}\frac{\Gamma(\lambda+j)}{\Gamma(\lambda)}\delta_{j,k}=\left\{{k \atop k}\right\}\frac{\Gamma(\lambda+k)}{\Gamma(\lambda)}=\mu_k(G(\lambda)).$$

10. Condition: $\rho\to 1$, $c=0$, $d=1/(1-\rho)$.
Again, characteristic functions provide an easy proof since

$$E\{\exp it(1-\rho)Y(\lambda,\rho)\}=\left\{1+\frac{\rho}{1-\rho}\left(1-e^{it(1-\rho)}\right)\right\}^{-\lambda}\to(1-it)^{-\lambda}.$$

11. Condition: $\lambda\to\infty$, $c=\lambda$, $d=\sqrt{\lambda}$.
As before

$$\log E\left\{\exp it\frac{G(\lambda)-\lambda}{\sqrt{\lambda}}\right\}=-\lambda\log\left(1-\frac{it}{\sqrt{\lambda}}\right)-it\sqrt{\lambda}\to-\frac{t^2}{2}.$$

12. Condition: $m\to\infty$, $c=0$, $d=m$.
Start from (2) and write

$$(10)\qquad p_n=\frac{1}{B(\lambda,\mu)}\int_0^1 y^{\lambda-1}(1-y)^{\mu-1}\left\{\binom{m}{n}y^n(1-y)^{m-n}\right\}dy$$

and hence

$$P\left\{\frac{R_m(\lambda,\mu)}{m}\le x\right\}=\frac{1}{B(\lambda,\mu)}\int_0^1 y^{\lambda-1}(1-y)^{\mu-1}P\{B_m(y)\le mx\}dy$$

where we used the binomial distribution on the right. By the weak law of large numbers however the probability in the integrand converges to 0 for $x<y$ and to 1 for $x>y$. This leads to the required result.

13. Condition: $\mu \to \infty$, $c = 0$, $d = 1/\mu$.
We use the moments.

$$\mu_k(\mu B(\lambda,\mu)) = \frac{\Gamma(\lambda+k)}{\Gamma(\lambda)} \frac{\nu^k \Gamma(\nu)}{\Gamma(\nu+k)} \left(\frac{\mu}{\nu}\right)^k \to \frac{\Gamma(\lambda+k)}{\Gamma(\lambda)} = \mu_k(G(\lambda)).$$

14. Condition: $\lambda, \mu \to \infty$, $c = \lambda/\nu$, $d = \sqrt{(\lambda\mu)/\nu^3}$.
Define

$$F(x) := P\left\{\frac{\nu B(\lambda,\mu) - \lambda}{\sqrt{(\lambda\mu)/\nu}} \le x\right\}.$$

Then

$$F'(x) = \frac{1}{B(\lambda,\mu)} \sqrt{\frac{\lambda\mu}{\nu^3}} \left(\frac{\lambda}{\nu} + x\sqrt{\frac{\lambda\mu}{\nu^3}}\right)^{\lambda-1} \left(\frac{\mu}{\nu} - x\sqrt{\frac{\lambda\mu}{\nu^3}}\right)^{\mu-1}$$

$$= \frac{\lambda^{\lambda-\frac{1}{2}}}{\Gamma(\lambda)} \frac{\mu^{\mu-\frac{1}{2}}}{\Gamma(\mu)} \frac{\Gamma(\nu)}{\nu^{\nu-\frac{1}{2}}} \left(1 + x\sqrt{\frac{\mu}{\lambda\nu}}\right)^{\lambda-1} \left(1 - x\sqrt{\frac{\lambda}{\mu\nu}}\right)^{\mu-1}.$$

A little calculation and Stirling's formula show that under the given limiting operation, the right hand side tends to the density of a standard normal distribution.

15. Condition: $\lambda, \mu, m \to \infty$, $\lambda/m\mu \to 0$, $m/\nu \to 0$, $c = (m\lambda)/\nu$, $d = \sqrt{m\lambda\mu}/\nu$.
From (10) we see that

$$P\{R_m(\lambda,\mu) \le x\} = \frac{1}{B(\lambda,\mu)} \int_0^1 y^{\lambda-1}(1-y)^{\mu-1}\phi_m(x,y)dy$$

where

$$\phi_m(x,y) := \sum_{r=0}^{x} \binom{m}{r} y^r (1-y)^{m-r}. \tag{11}$$

Put $y = q_m + p_m z$, where the sequences $\{q_m\}$ and $\{p_m\}$ will be determined shortly. Also put $\overline{q}_m := 1 - q_m$. Then changing y into z we find that

$$P\{R_m(\lambda,\mu) \le x\} = I_1(m) \int_{-q_m/p_m}^{\overline{q}_m/p_m} I_2(z,m)\phi_m(x, q_m + p_m z)dz,$$

where in turn

$$I_1(m) = \frac{1}{B(\lambda,\mu)} q_m^{\lambda-1} \overline{q}_m^{\mu-1} p_m$$

and

$$I_2(z,m) = \left\{1 + \frac{p_m}{q_m} z\right\}^{\lambda-1} \left\{1 - \frac{p_m}{\overline{q}_m} z\right\}^{\mu-1}.$$

When $\lambda, \mu \to \infty$ then $I_2(z,m)$ will have a useful limit if and only if we make the following choice

$$q_m = \frac{\lambda - 1}{\nu - 2}, \qquad \overline{q}_m = \frac{\mu - 1}{\nu - 2}, \qquad p_m^2 = \frac{(\lambda-1)(\mu-1)}{(\nu-2)^3}.$$

With these choices, Stirling's formula yields $I_1(m) \to 1/\sqrt{2\pi}$ when $\lambda, \mu \to \infty$. If further $p_m/q_m \to 0$ and $p_m/\overline{q}_m \to 0$ together with $\lambda, \mu \to \infty$ then uniformly on compact z-intervals,

$$I_2(z,m) \to \exp -\frac{z^2}{2}.$$

It remains to look at the factor ϕ_m. However, as in the proof of 12

$$\phi_m(x, q_m + p_m z) = P\{B_m(q_m + p_m z) \le x\}.$$

Clearly, as $m \to \infty$, B_m needs centering and normalization. Let us first look at compact z-intervals and define

$$T_m := \frac{B_m(q_m + p_m z) - m q_m}{\sqrt{m q_m \overline{q}_m}}.$$

Then by the binomial character of B_m we find that

$$E\{\exp itT_m\} = \left\{\exp\left(-it\frac{q_m}{d_m}\right)\left\{1 + (q_m + p_m z)(e^{it/d_m} - 1)\right\}\right\}^m,$$

where $d_m := m q_m \overline{q}_m$ for convenience.

If we require $q_m/d_m \to 0$ and $mp_m/d_m \to 0$ on top of the previous conditions, then an easy calculation yields that $E(\exp itT_m) \to \exp -\frac{t^2}{2}$. Hence if

$$(12) \qquad \lambda, \mu, m \to \infty, \quad \frac{p_m}{q_m} \to 0, \quad 2\frac{p_m}{\overline{q}_m} \to 0, \quad \frac{p_m}{d_m} \to 0, \quad \frac{mp_m}{d_m} \to 0,$$

then on compact z-intervals we can apply a result in Ibragimov & Linnik (1971) to the extent that

$$\phi_m(mq_m + d_m x, q_m + p_m z) \to \Phi(x).$$

Take m in accordance with (12) and pick a fixed T such that $0 < T < min(q_m/p_m\,,\bar{q}_m/p_m)$. Then

$$P\left\{\frac{R_m(\lambda,\mu)-mq_m}{d_m}\le x\right\}=K+L+M:=$$

$$I_1(m)\left\{\int_{-q_m/p_m}^{-T}+\int_{-T}^{T}+\int_{T}^{\bar{q}_m/p_m}\right\}I_2(z,m)\phi_m(mq_m+d_mx,q_m+p_mz)dz.$$

By the uniformity on $|z|<T$ we have under (12) that

$$L\to\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}e^{-x^2/2}\Phi(x)dx.$$

We have to estimate the two remaining quantities K and M. We only deal with K. Returning to the variable y we see that

$$\begin{aligned}K&=\frac{1}{B(\lambda,\mu)}\int_{q_m+p_mT}^{1}y^{\lambda-1}(1-y)^{\mu-1}\phi_m(mq_m+d_mx,y)dy\\&\le\frac{1}{B(\lambda,\mu)}\int_{q_m+p_mT}^{1}y^{\lambda-1}(1-y)^{\mu-1}dy\\&\le\frac{1}{B(\lambda,\mu)}\int_{q_m+p_mT}^{1}y^{\lambda-1}(1-y)^{\mu-1}\left(\frac{y-q_m}{p_mT}\right)^2dy\\&\le\frac{1}{B(\lambda,\mu)}\frac{1}{p_m^2T^2}\int_0^1y^{\lambda-1}(1-y)^{\mu-1}(y-q_m)^2dy=\frac{1}{T^2}(1+o(1)).\end{aligned}$$

Treating M in the same way, we see that under (12) the result is proved by taking ultimately T as large as we like.

It remains to simplify the conditions in (12). But this is easy, since the choice of q_m and p_m reduce all conditions to the simpler set

$$\lambda\,,\mu\,,m\to\infty\,,\qquad\frac{m}{\nu}\to0\,,\qquad\frac{\lambda}{m\mu}\to0$$

m as in the statement of the condition. This finishes the proof of the table. Let us make a couple of remarks.

1. As mentioned before, a large number of the above approximations (1,2,3, 4,5,7,8,10,11,13 and 14) are standard. But also the other have been derived before. We mention in particular the papers by Bosch (1963) and the monograph by Bricas (1949) where we find the results covered in

6, 9 and 12. Finally 15 has been known for quite some time. See Gouet (1993) for references and a functional limit theorem. Let us however mention that for a number of cases, our conditions are (some- times much) simpler than those found in the mentioned literature.

2. As far as we can see, we obtained all possible limits for the PED which are non degenerate. We surely rederived all special cases that we could find in the literature.
3. There are still more special cases that could have been written down as they appear by rescaling the limiting random variable. For example we could easily obtain the gamma distribution with two parameters, or the beta-distribution on $\mathbb{R}_+$, or the non-standard normal distribution. Also we could have included for example the geometric, the exponential, the logaritmic, the uniform and the arc-sine distributions. As these distributions are obtainable by properly chosing the remaining parameter(s), we have not included them.
4. There is a remarkable similarity between the given table and a subset of the *Askey table* for orthogonal polynomials. For example see Koekoek e.a. (1994) or (Nokiforov e.a. (1990) and the references contained in them. This unexpected connection opens up intruiging questions about the interplay between the two seemingly different mathematical topics.
5. There also have been a few attempts to extend the PED to bivariate and multivariate settings. See for example Marshall e.a. (1990) and its references.
6. A slightly different survey table as given above has appeared in preprint format in Teugels (1975) and again in lecture notes on insurance mathematics in Teugels (1985).

5. Recursive class. We finish with some observations that link the PED to actuarial mathematics. Consider a homogeneous portfolio where claims arrive according to a time process that we denote for convenience by $\{N(t), t \geq 0\}$. In the actuarial literature one finds. a substantial amount of papers on possible probabilistic models for the counting process of claims. For a recent treatment, see Panjer and Willmot (1992).

Most often actuaries look for a distribution of $N(t)$ at a specific time point, like one year. We can then drop the dependence on t if the time behaviour is of secundary interest. We therefore write N.

Let us denote by pn the probability that $P\{N = n\}$. One very popular model suggests that the sequence $\{p_n\}$ satisfies a first order recursion relation of the form

$$\frac{p_n}{p_{n-1}} = a + \frac{b}{n}, \quad n = 1, 2, 3, \ldots . \tag{13}$$

For a full treatment see section 6.6 in Panjer and Willmot (1992). Among the distributions that satisfy the above relation we mention the binomial, Poisson and negative binomial distributions. For the calculation of the sequence of moments for distributions that satify (13) we refer to Szynal and Teugels (1995).

Generalizations of (13) abound. We mention a few.

- One can start the recursive relation with $n = 2$ rather than with $n = 1$. This leads to the (a, b)*class* treated in Ch.7 of Panjer and Willmot (1992). It is easy to see that the latter case introduces an extra para- meter p_0 into the analysis. Hence the solutions are of a *truncated* form with a potential weight at the zero-value.
- A generalization treated by Willmot and Panjer (1987) starts from the relation

$$(an+1)p_n = \{b(n-1)+c\}p_{n-1}\,, \quad n \geq 1.$$

  This generalization leads i.a. to a Warring distribution, a shifted logaritmic distribution and a hyperPoisson distribution.
- Panjer and Willmot in (1982) use quadratic rather than linear functions in the recursion

$$n(n-1)p_n = \{b(n-1)(n-2)+c(n-1)+d\}p_{n-1}\,, \; n \geq 1.$$

  Among the new members are a generalised Warring distribution as well as a form of the Pólya-Eggenberger distribution. We return to this model later.
- A pleasant generalization has been given by Sundt (1992). He takes as relation

$$p_n = \sum_{i=1}^{k}\left\{a_i + \frac{b_i}{n}\right\}p_{n-i}\,, \; n > k$$

  and finds classes of distributions that enjoy some nice algebraic properties. For the case where $k = 2$, see Schröter (1990).

We focus on the third approach. Assume that the portfolio has two types of policies, N_1 of the *claim causing type*, the remaining $N_2 = N - N_1$ of the *non-claim type*. At time points $0, 1, 2, ..., m$ a policy is picked at random from the portfolio and its claim is recorded, if any. The policy drawn is then put back in the portfolio together with Δ of the same type. We call Δ the *contamination*. The number of claims recorded after m trials is then R_m and follows a PED. With this kind of interpretation, the PED can be seen as a model for the claim number distribution.

The potential applications of the PED in insurance has already been noticed by Kupper (1962) where a number of special cases are treated as well.

Let us now look at the recursive properties of the PED. From (1) it follows that

$$\frac{p_n}{p_{n-1}} = \frac{m-n+1}{n} \frac{N_1 + (n-1)\Delta}{N_2 + (m-n)\Delta}, \; n \geq 1 \tag{14}$$

where p_0 is given by

$$p_0 = \begin{cases} (N_2/N)^m & \text{if } \Delta = 0, \\ \frac{\Gamma(\mu+m)}{\Gamma(\mu)} \frac{\Gamma(\nu)}{\Gamma(\nu+m)} & \text{if } \Delta \neq 0. \end{cases}$$

We notice that (14) has a form rather akin to the third generalization above. However an identification does not seem possible even when the PED has the four parameters N, N_1, m and Δ while Schröter's has only three. Even a truncation does not lead to a better identification.

Let us finally point out that the three cases that satisfy (13) are limiting cases of (14) under precisely the same kind of conditions as specified in the table.

- First the *binominal distribution*. The ratio $\{N_1 + (n-1)\Delta\}/\{N_2 + (m-n)\Delta\}$ should be tackled under the conditions of 1, i.e. $\lambda, \nu \to \infty$, $\lambda/\nu \to p \in (0,1)$. It easily follows that (13) is satisfied with $b = [p/(1-p)](m+1)$ and $a = -p/(1-p)$.
- Next the Poisson distribution. The conditions for case 4 were $m, \lambda \to i$, $m\lambda/\nu \to \tau > 0$ and hence

$$(m-n+1)\frac{N_1 + (n-1)\Delta}{N_2 + (m-n)\Delta} \sim m\frac{N_1}{N} \to \tau$$

  v so that (13) follows with $a = \tau$ and $b = 0$.
- Finally the *Pascal distribution* follows in precisely the same fashion by using the condition of case 6, i.e. $m, \mu \to \infty$, $m/\mu \to \rho/(1-\rho)$, $0 < \rho < 1$.
  Now there results that $b = \rho(\lambda - 1)$ while $a = \rho$.

6. Acknowledgements. We would like to thank A.J. Bosch for providing us with a number of references that were unfamiliar to us at the time of writing the report Teugels (1975). We also thank W. Schoutens for helpful criticism on an earlier draft of this paper.

References

[1] Berg, S., *Urn models, Encyclopedia of Statistical Sciences*, 9, S. Kotz, N.L. Johnson and C.B. Read (editors), 1988.

[2] Billingsley, P., *Probability and Measure*, 3rd ed. New York: Wiley, 1995.

[3] Bosch, A.J., *The Plya distribution*, Statist. Neerlandica **17** (1963), 201- 213.

[4] Bricas, M.A., *Le Systéme de Courbes de Peczrson et le Schéma d'Urne de Pólya*, Athenes: G.S. Christou & Fils, 1949.

[5] Eggenberger, F. and G. Pólya, *Über die Statistik verketteter Vorgänge*, Z. Angew. Math. Mech. **3** (1923), 279-289.

[6] Gouet, R., *Martingale functional central limit theorems for a generalized Pólya urn*, Ann. Probab. **21** (1993), 1624-1639.

[7] Ibragimov, I.A. and Yu. V. Linnik, *Independent and Stationary Sequences of Random Variables*, Groningen: Noordhoff, 1971.

[8] Jordan, C., *Sur un cas généralisé de la probabilité des épreuves répétées, C. R. Acad. Sci. Sér. I Math.*, Paris **184**, 1927.

[10] Johnson, N.L., S. Kotz and A. W. Kemp, *Univariate Discrete Distributions*, 2nd ed. New York: Wiley, 1992.

[11] Koekoek, R. and R. F. Swarttouw, *The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue*, Report Faculty of Technical Mathematics and Informatics (1994), 94-05, Delft.

[12] Kupper, J., *Wahrscheinlichkeitstheoretische Modelle in der Schadenversicherung.* Teil I*: Die Schadenzahl*, Blätter der Deutschen Gesellschaft für Versicherungsmathematik, 5, 1962.

[13] Marshall, A.W. and I. Olkin, *Bivariate distributions generated from Pólya-Eggenberger models*, J. Multivariate Anal. **35** (1990), 48-65.

[14] Nikiforov, A.F., S. K. Suslov and V. B. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*, Berlin: Springer-Verlag, 1990.

[15] Panjer, H. H. and G. E. Willmot, *Recursions for compound distributions*, Astin Bulletin **13** (1982), 1-11.

[16] Panjer, H.H. and G. E. Willmot, *Insurance Risk Models*, Society of Actuaries, Schaumburg, 1992.

[17] Schröter, K.J., *On a family of counting distributions and recursions for related distributions*, Scand. Actuar. J. (1990), 161-175.

[18] Sundt, B., *On some extensions of Panjer's class of counting distributions*, Astin Bulletin **22** (1991), 61-80.

[19] Szynal, D. and J. L. Teugels, *Om moments of a class of counting distributions*, Annal. Univ. Mariae Curie-Sklodowska Sect. A **49** (1995), 199-211.

[20] Teugels, J.L., *Limits and approximations for the Pólya-Eggenberger distribution*, Med. uit het Wiskundig Instituut, Katholieke Universiteit Leuven **62** (1975).

[21] ———, *Selected Topics in Insurance Mathematics*, Wiskundig Instituut, Katholieke Universiteit Leuven, vii + p.150 (1985).

[22] Willmot, G.E. and H. H. Panjer, *Difference equation approaches in evaluation on compound distributions*, Insurance: Mathematics & Economics **6** (1987), 43-56.

Katholieke Universiteit Leuven
Departement Wiskunde
Celestijnenlaan 200 B B-3001 Heverlee
Belgium

received October 14, 1996