STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 337 (VP 24)

Some remarks on mixtures of distributions

Preliminary report

by

W. Molenaar

December 1964

Printed at the Mathematical Centre at Amsterdam,49,2nd Boerhaavestraat.
The Netherlands.

The Mathematical Centre, founded the 11th of February 1946, is a non - profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Scientific Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

1. Summary

The connection between one-parameter mixtures and generalizations (i.e. $\underline{y}$-fold [1] convolutions, with stochastic $\underline{y}$) given by GURLAND (1957) is used to derive some simple properties of both, partly found with different proofs in TEICHER (1960). The one-parameter mixture is compared with one of its components.
Examples are given in an appendix.

2. Definitions

A compound Poisson distribution is a Poisson distribution with a random parameter: its parameter is not a positive constant but a random variable assuming positive real values. A generalized Poisson distribution is the $\underline{n}$-fold convolution of an arbitrary distribution, where $\underline{n}$ has a Poisson distribution. These definitions introduced by FELLER (1943) have been extended by GURLAND (1957) and TEICHER (1960) to the Definitions 1 and 2 given below. As some authors, e.g. FELLER (1957), use "compound" for what is here called "generalized", we shall henceforth replace "compound" by the less ambiguous "mixed".

Definition 1. If F_θ is a distribution function for each parameter value $\theta \in T \subset R^1$, and H is a distribution function which assigns probability 1 to T, then the H-mixture of F_θ is the distribution function $F_\theta \underset{\theta}{\wedge} H$ given by

(1) $$(F_\theta \underset{\theta}{\wedge} H)(x) = \int F_\theta(x)\, dH(\theta).$$

---

1) Random variables are underlined.

We shall denote by $\underline{x}_\theta$ the random variable with distribution function F_θ and by $\underline{\underline{x}}_\theta$ the random variable with distribution function $F_\theta \underset{\theta}{\bigwedge} H$.[2)]

This is a special case of TEICHER's definition of m-parameter mixtures. In this report the distribution function H is always one-dimensional, though F_θ may have more than one parameter. The extra θ under the sign "$\bigwedge$" is convenient in this case. Sometimes we shall write $F_{c\theta} \underset{\theta}{\bigwedge} H$, though we could have included the constant c in the distribution function H. The symbol $\underset{\theta}{\bigwedge}$ is also used between names of distributions. Several well-known mixtures are listed in Appendix 1. Two examples of mixtures are

$$\text{(2)} \qquad \begin{aligned} &\text{Binomial } (n,p) \underset{n}{\bigwedge} \text{Poisson } (\mu) = \text{Poisson } (\mu p); \\ &\text{Poisson } (k\mu) \underset{\mu}{\bigwedge} \text{Poisson } (\lambda) = \text{Neyman } A(\lambda,\mu). \end{aligned}$$

If H is a degenerate distribution function the mixture is just one F_θ, and if all F_θ are degenerate it is just a trivial modification of H. If we want to exclude these cases we shall call a mixture <u>non-trivial</u>.

It is obvious that for example the characteristic function and the moments about zero (if existing) are <u>mix-linear</u> with respect to (w.r.t.) $F_\theta \underset{\theta}{\bigwedge} H$, i.e. if ϕ_θ is characteristic function of F_θ and ϕ of $F_\theta \underset{\theta}{\bigwedge} H$, then

$$\text{(3)} \qquad \begin{aligned} &\phi(t) = \int \phi_\theta(t) \, dH(\theta); \\ &\mathcal{E}\, \underline{\underline{x}}_\theta^k = \int \mathcal{E}\, \underline{x}_\theta^k \, dH(\theta). \end{aligned}$$

<u>Definition 2.</u> If the random variable $\underline{x}$ has distribution function F and the non-negative integer-valued random variable $\underline{y}$ has distribution function G, we define the <u>G-generalized F-distribution</u>, with distribution function denoted by F^{G*}, as the distribution of

---

2) GURLAND uses $F \wedge H$ and $\underline{x} \wedge \underline{\theta}$. We have preferred to use another notation, more suggestive of the underlying idea and avoiding the possibly confusing suggestion of symmetry.

$$x^{y*} \overset{\text{def}}{=} x_1 + x_2 + \ldots + x_y, \qquad (4)$$

where x_i are independent and have distribution function F. If the characteristic function ϕ of F is such that $\{\phi(t)\}^y$ is a characteristic function for all values y in the carrier [3] of G, we extend the definition to arbitrary distribution functions G with $G(0-) = 0$ and define x^{y*} and F^{G*} by their characteristic function

$$g(\phi(t)) \overset{\text{def}}{=} \int \{\phi(t)\}^y \, dG(y). \qquad (5)$$

This clearly includes the definition by (4). GURLAND (1957) uses the notation $y \vee x$, and $G \vee F$, and defines it, for non-negative x and y, by stating that its generating function is $g(f(z))$, where $f(z) \equiv \mathcal{E}\, z^{x}$ and $g(z) \equiv \mathcal{E}\, z^{y}$. In GURLAND's examples, where F is infinitely divisible, $g(f(z))$ is always a generating function; it does not seem quite clear whether this is true in all cases with $F(0-) = 0$ and $G(0-) = 0$. Compared with GURLAND's, our definition includes negative values for x but excludes the cases where $\{\phi(t)\}^y$ is not a characteristic function for some y in the carrier of G. In the list of examples in the appendix we have always a nonnegative integer-valued y or an infinitely divisible x with a nonnegative y; in both cases x^{y*} is defined.

3. Simple properties

Lemma 1. Generalizing and mixing are associative operations, i.e.

$(H^{G*})^{F*} = H^{(G^{F*})*}$ and $(F_\theta \underset{\theta}{\wedge} G_\eta) \underset{\eta}{\wedge} H = F_\theta \underset{\theta}{\wedge} (G_\eta \underset{\eta}{\wedge} H)$

in all cases where the distributions are defined.

---

3) The carrier of a distribution function G is the set of all values y for which $\varepsilon > 0$ implies $G(y + \varepsilon) - G(y - \varepsilon) > 0$.

<u>Proof.</u> If both sides of the first formula are meaningful, then $F(0-) = 0$ and $G(0-) = 0$ and $f(g(z))$ must be a generating function (corresponding to G^{F*}). Now if $\phi(t)$ is the characteristic function of H, then by (5) both sides have characteristic function $f(g(\phi(t)))$.

For the second formula, one finds from TULCEA's theorem (LOÈVE (1963), p. 137)

$$(6) \qquad \iint F_\theta(x)\, dG_\eta(\theta)\, dH(\eta) = \int F_\theta(x)\, d_\theta \int G_\eta(\theta)\, dH(\eta).$$

<u>Remark.</u> It is trivial that for a two-parameter family $\{F_{\theta,\eta}\}$ we have

$$(7) \qquad (F_{\theta,\eta} \underset{\theta}{\wedge} G) \underset{\eta}{\wedge} H = (F_{\theta,\eta} \underset{\eta}{\wedge} H) \underset{\theta}{\wedge} G.$$

This might be extended by splitting any two-dimensional distribution of θ and η in the two possible ways in a conditional and a marginal distribution.

<u>Definition 3.</u> The family $\{F_\theta \mid \theta \in T\}$ of distribution functions is <u>additively closed</u> (w.r.t. θ) if we have, for all θ, $\eta \in T$:

$$(8) \qquad \theta + \eta \in T \text{ and } F_\theta(x) * F_\eta(x) \equiv F_{\theta+\eta}(x).$$

It is <u>strongly additively closed</u> if there exists a characteristic function $\phi_1(t)$ independent of θ such that for each $\theta \in T$ the characteristic function ϕ_θ of F_θ is $\phi_\theta(t) = \{\phi_1(t)\}^\theta$.

If T consists of the positive integers or rationals the two notions coincide; if $T = (0,\infty)$ an additively closed family is strongly additively closed if $\phi_\theta(t)$ is a continuous function of θ or if $\phi_\theta(t)$ is real-valued for real t (see TEICHER (1954), where also additively closed families in more than one parameter are investigated). PYKE (1960) has shown for additively closed families with $T = [0,\infty)$ that $\phi_\theta(t) = \{\phi_1(t)\}^\theta \exp\{itc(\theta)\}$, where the real-valued function $c(\theta)$ on $[0,\infty)$ has $c(\theta) = 0$ for all rational θ and $c(\theta) + c(\eta) = c(\theta + \eta)$ for all $\theta \geq 0$ and $\eta \geq 0$. For a strongly additively closed family with for T the positive reals or rationals ϕ_1 is of course infinitely divisible.

If ϕ_1 is not degenerate, the parameter of a strongly additively closed family can assume only nonnegative values, as we have $|\phi_\theta(t)| = |\phi_1(t)|^\theta \leq 1$ for all θ and t.

A few examples of strongly additively closed families are [4]:

Normal $(0, \sigma^2)$ w.r.t. σ^2 $(\sigma > 0)$;

Normal $(\mu\theta, \sigma^2\theta)$ w.r.t. θ $(\theta > 0)$;

Poisson (θ) w.r.t. θ $(\theta > 0)$;

Binomial (n, p) w.r.t. n (integer $n > 0$);

Pascal (γ, p) w.r.t. γ ($\gamma > 0$, or integer $\gamma > 0$);

in all cases the parameter value 0 corresponds to the degenerate distribution in 0 and may be included.

The following basic theorem is a slightly modified form of a theorem by GURLAND (1957):

Lemma 2. If $\{F_\theta \mid \theta \in T\}$ is strongly additively closed and H assigns probability 1 to T, then

$$F_\theta \underset{\theta}{\wedge} H = F_1^{H*}. \tag{9}$$

Proof. On both sides the characteristic function is $\int \{\phi_1(t)\}^\theta \, dH(\theta)$.
It is not necessary to assume $1 \in T$, as ϕ_1 and F_1 are defined by Definition 3.

Several relations of the type (9) are found in Appendix 1.
As an example we mention

$$\text{Pascal } (\gamma, p) = \text{Poisson } (\lambda) \underset{\lambda}{\wedge} \text{Gamma } (pq^{-1}, \gamma) = \{\text{Poisson } (1)\}^{\text{Gamma } (pq^{-1}, \gamma)*}.$$

Lemma 3. If $\{F_\theta\}$ is strongly additively closed, the operations $F_\theta \underset{\theta}{\wedge} \circ$ and $\circ^{G*}$ may be interchanged.

Proof. By lemmas 1 and 2 we have

---

4) A specification of the distributions is given in Appendix 2.

(10) $F_\theta \underset{\theta}{\frown} (H^{G*}) = F_1{}^{(H^{G*})*} = (F_1{}^{H*})^{G*} = (F_\theta \underset{\theta}{\frown} H)^{G*}$.

By lemma 2, each mixed Poisson distribution Poisson $(\theta) \underset{\theta}{\frown} H$ is also $\{\text{Poisson }(1)\}^{H*}$. Some generalized Poisson distributions $G^{\text{Poisson}*}$ are at the same time mixed Poisson (e.g. Neyman A and Pascal).
MACEDA (1948) proves that all distributions of the form $(\text{Poisson} \frown H)^{\text{Poisson}*}$ are both generalized Poisson and mixed Poisson.
This is now a consequence of lemma 3, as we have

(11) $(\text{Poisson} \frown H)^{\text{Poisson}*} = \text{Poisson} \frown (H^{\text{Poisson}*})$.

<u>Lemma 4.</u> One-sided distributivity of generalizing and mixing with respect to convolution holds in the following sense: whenever both sides are defined we have

(12) $H^{(F*G)*} = (H^{F*}) * (H^{G*})$,

and

(13) $F_\theta \underset{\theta}{\frown} (H * G) = (F_\theta \underset{\theta}{\frown} H) * (F_\theta \underset{\theta}{\frown} G)$ if $\{F_\theta\}$ is strongly additively closed.

<u>Proof.</u> If H has characteristic function ϕ , both sides of (12) have characteristic function $f(\phi(t))\ g(\phi(t))$. From (12) and lemma 2 follows (13).

(13) was already proved by TEICHER (1960), by writing out the integrals. For F_θ = Poisson (θ) it is mentioned by FELLER (1943) and MACEDA (1948). The last author proves also

$$\{H_1{}^{\text{Poisson }(\lambda)*}\} * \{H_2{}^{\text{Poisson }(\mu)*}\} = \left\{\frac{\lambda H_1 + \mu H_2}{\lambda + \mu}\right\}^{\text{Poisson }(\lambda + \mu)*}.$$

Only the special form of the Poisson characteristic function makes this extension of (12) to different H_i possible.

One can easily see that the other two distributive laws

$(H^{F*}) * (G^{F*}) = (H * G)^{F*}$ and $(F_\theta * F_\eta) \wedge H = (F_\theta \wedge H) * (F_\eta \wedge H)$

hold only in the trivial cases where at least one of the distributions is degenerate.

Lemma 5. If G is infinitely divisible, then so is H^{G*} for each H and $F_\theta \underset{\theta}{\wedge} G$ for each strongly additively closed family $\{F_\theta\}$.

Proof. For each positive integer n there is a distribution function G_n such that $G_n^{n*} = G$. Thus by lemmas 4 and 2

$$H^{G*} = H^{(G_n^{n*})*} = (H^{G_n*})^{n*} \tag{14}$$

and

$$F_\theta \underset{\theta}{\wedge} G = F_1^{G*} = (F_1^{G_n*})^{n*} = (F_\theta \underset{\theta}{\wedge} G_n)^{n*} . \tag{15}$$

The second half of lemma 5 is stated by TEICHER (1960), with a different proof.

4. Moments of mixtures

For the investigation of the moments of $\underline{x}_\theta$ we introduce

$$\begin{aligned} m_\theta &\overset{\text{def}}{=} \mathcal{E}\, \underline{x}_\theta = \int x \, dF_\theta(x); \\ s_\theta^2 &\overset{\text{def}}{=} \sigma^2(\underline{x}_\theta) = \int (x - m_\theta)^2 \, dF_\theta(x). \end{aligned} \tag{16}$$

By well-known formulae we have

$$\begin{aligned} \mathcal{E}\, \underline{x}_\theta &= \mathcal{E}\{\mathcal{E}(\underline{x}_\theta \mid \underline{\theta})\} = \mathcal{E}\, m_{\underline{\theta}} \; ; \\ \sigma^2(\underline{x}_\theta) &= \sigma^2\{\mathcal{E}(\underline{x}_\theta \mid \underline{\theta})\} + \mathcal{E}\{\sigma^2(\underline{x}_\theta \mid \underline{\theta})\} = \sigma^2(m_{\underline{\theta}}) + \mathcal{E}\, s^2_{\underline{\theta}}. \end{aligned} \tag{17}$$

Let us assume that F_θ has expectation $m_\theta = \theta$, then $\mathcal{E}\, \underline{x}_\theta = \mathcal{E}\, \underline{\theta}$ and

$\sigma^2(\underline{x}_\theta) = \sigma^2(\underline{\theta}) + \mathcal{E}\, s^2_{\underline{\theta}}$. If F_θ is non-degenerate for some set of θ-values assumed with positive probability, then the variance of $\underline{x}_\theta$ is larger than the variance of $\underline{\theta}$. In the special case F_θ = Poisson (θ) we have also $s^2_\theta = \theta$ and, as proved by FELLER (1943), $\sigma^2(\underline{x}_\theta) = \sigma^2(\underline{\theta}) + \mathcal{E}\, \underline{x}_\theta$: the variance of a non-trivial mixture of Poisson distributions is always larger than its expectation.

Another special case is $F_\theta \underset{\theta}{\wedge}$ Poisson (λ). Here it is not very realistic to assume $m_\theta = \theta$, as in most mixtures the expectation of F_θ will not be integer-valued for all θ. If $m_\theta = k\theta$ for some constant k and integer θ, then

$$\sigma^2(\underline{x}_\theta) = k^2\lambda^2 + \mathcal{E}\, s^2_{\underline{\theta}} > k\lambda = \mathcal{E}\, \underline{x}_\theta,$$

provided we have $k\lambda > 1$. This condition $\mathcal{E}\, \underline{x}_\theta > 1$ is not necessary for a Poisson-mixture of distributions F_θ with expectation $k\theta$ to have larger variance than expectation. The Poisson Binomial [5] distribution Binomial $(n\theta, p) \underset{\theta}{\wedge}$ Poisson (λ) has expectation λnp and variance $\lambda n^2p^2 + \lambda npq$; for $n > 1$ the variance exceeds the expectation even in the case $\mathcal{E}\, \underline{x}_\theta \leq 1$.

If $\mathcal{E}\, \underline{\theta} = \int \theta \, dH(\theta)$ exists and is an admissible parameter value, it is interesting to compare the mixture $F_\theta \underset{\theta}{\wedge} H$ with the single component $F_{\mathcal{E}\underline{\theta}}$. When the expectation m_θ of F_θ is proportional to θ, this amounts to a comparison of the mixture to the component with the same expectation. Starting-point is the following result by FELLER (1943):

Lemma 6. For each non-trivial mixture Poisson $(\theta) \underset{\theta}{\wedge} H$, $\sigma^2(\underline{x})$ is larger, $P\{\underline{x} = 0\}$ is larger and $P\{\underline{x} = 1\} / P\{\underline{x} = 0\}$ is smaller than the corresponding quantity for the Poisson distribution with the same expectation.

---

[5] This distribution was first mentioned by NEYMAN (1939), later by FELLER (1943) and SKELLAM (1952). It probably got its name, which is commonly accepted now, from McGUIRE et al.(1957). Some authors use "Poisson Binomial" for the distribution of the number of successes in n independent experiments with different probabilities for succes.

Two possible extensions are stated in the following lemmas.

Lemma 7. For a non-trivial mixture $F_\theta \underset{\theta}{\wedge} H$ where $\mathcal{E}\underline{\theta}$ is an admissible parameter value and the expectation of F_θ is a linear function of θ ($m_\theta = k\theta$, $k \neq 0$) each of the following conditions is sufficient for $\sigma^2(\underline{x}_{\underline{\theta}}) > \sigma^2(\underline{x}_{\mathcal{E}\underline{\theta}})$:

(a) $s^2_\theta = \int (x - m_\theta)^2 \, dF_\theta(x)$ is a convex [6] function of θ;

(b) F_θ is Binomial (n,θ) for fixed $n > 1$.

Proof. In case (a) we have

$$\sigma^2(\underline{x}_{\underline{\theta}}) = \sigma^2(m_{\underline{\theta}}) + \mathcal{E} s^2_{\underline{\theta}} = k^2\sigma^2(\underline{\theta}) + \mathcal{E} s^2_{\underline{\theta}} > s^2_{\mathcal{E}\underline{\theta}},$$

by Jensen's inequality and the fact that the first term is positive. In case (b) $k = n$, $s^2_\theta = n\theta(1 - \theta)$, and

$$\sigma^2(\underline{x}_{\underline{\theta}}) = n^2\sigma^2(\underline{\theta}) + n\mathcal{E}\underline{\theta} - n\mathcal{E}\underline{\theta}^2 > n\mathcal{E}\underline{\theta} - n(\mathcal{E}\underline{\theta})^2 = s^2_{\mathcal{E}\underline{\theta}},$$

because $(n^2 - n)\sigma^2(\underline{\theta}) > 0$.

Remark. The lemma holds for Binomial (n,θ) though its variance is a strictly concave function of θ. Cases (a) and (b) together cover all mixtures listed in the appendix.

Definition 4. A functional v mapping a class of distribution functions into the real numbers is mix-concave w.r.t. $F_\theta \underset{\theta}{\wedge} H$, if it is defined at least for $F_\theta \underset{\theta}{\wedge} H$ and all F_θ, and satisfies

$$v(F_\theta \underset{\theta}{\wedge} H) \geq \int v(F_\theta) \, dH(\theta). \tag{18}$$

---

6) As usual "convex" includes "linear"; a function f is called "strictly convex" when $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$ for all x and y in the interval where f is defined and all $\lambda \in (0,1)$.

It is <u>strictly mix-concave</u> if we have inequality and <u>mix-linear</u> if we have equality in (18).

<u>Examples.</u> For all mixtures and all real a and b functionals like $v(F) = P\{\underline{x} = a \mid F\}$ and $v(F) = P\{a < \underline{x} \leq b \mid F\}$ are mix-linear. The expectation is mix-linear; as we have proved $\sigma^2(\underline{x}_\theta) = \sigma^2(m_{\underline{\theta}}) + \mathcal{E}\, s^2_{\underline{\theta}}$, the variance is always mix-concave, and strictly mix-concave unless m_θ does not depend on θ. If v is mix-linear, not constant, and assumes only positive values, then $\frac{1}{v}$ is strictly convex.

<u>Lemma 8.</u> Let $F_\theta \underset{\theta}{\wedge} H$ be a non-trivial mixture and $\mathcal{E}\,\underline{\theta}$ an admissible parameter value. If the functional v is mix-concave w.r.t. $F_\theta \underset{\theta}{\wedge} H$, and $v(F_\theta)$ is a convex function of θ, then $v(F_\theta \underset{\theta}{\wedge} H) \geq v(F_{\mathcal{E}\underline{\theta}})$. The last inequality is strict as soon as v is strictly mix-concave or strictly convex in θ.

<u>Proof.</u> $v(F_\theta \underset{\theta}{\wedge} H) \geq \int v(F_\theta)\, dH(\theta) \geq v(F_{\mathcal{E}\underline{\theta}})$.

<u>Example 1.</u> Pascal (γ, p) = Poisson $(\theta) \underset{\theta}{\wedge}$ Gamma (pq^{-1}, γ).

The variance $s^2_\theta = \theta$ is linear in θ, so lemma 7 holds. In fact the variance γpq^{-2} for Pascal (γ, p) is larger than the variance of Poisson $(\mathcal{E}\,\underline{\theta})$, which is $\mathcal{E}\,\underline{\theta} = \gamma pq^{-1}$.
$v(F_\theta) = P\{\underline{x} = 0 \mid F_\theta\} = e^{-\theta}$ is strictly convex in θ and mix-linear, so lemma 8 is applicable. In fact $P\{\underline{x} = 0\}$ is q^γ for Pascal (γ, p); this is larger than $\exp(-\gamma pq^{-1})$ for Poisson $(\mathcal{E}\,\underline{\theta})$ as we have $\gamma \log(1 - p) > -\gamma p(1 - p)^{-1}$.
$v(F) = P\{\underline{x} = 1 \mid F\} / P\{\underline{x} = 0 \mid F\}$ can be shown to be strictly mix-convex and $v(F_\theta) = \theta$ is linear, so by an obvious modification of lemma 8

$$v(F_{\mathcal{E}\underline{\theta}}) = \int v(F_\theta)\, dH(\theta) > v(F_\theta \underset{\theta}{\wedge} H),$$

namely

$$\gamma pq^{-1} = \mathcal{E}\,\underline{\theta} > \gamma p.$$

Example 2. Polya (n,r,s) = Binomial $(n,p) \underset{p}{\wedge}$ Beta (r,s).

By lemma 7, the variance of the Polya distribution must exceed for $n > 1$ that of Binomial (n,r / (r + s)). In fact we have

$$\frac{nrs(n + r + s)}{(r + s)^2(r + s + 1)} > \frac{nrs}{(r + s)^2}.$$

$v(F_p) = P\{\underline{x} = 0 \mid F_p\} = (1 - p)^n$ is convex in p (strictly so for $n > 1$), and mix-linear. In fact $P\{\underline{x} = 0\}$ is

$$\frac{(s + n - 1)(s + n - 2) \ldots s}{(r + s + n - 1)(r + s + n - 2) \ldots (r + s)} \quad \text{and} \quad \left(\frac{s}{r + s}\right)^n$$

for Polya (n,r,s) and Binomial (n,r / (r + s)) respectively. They are clearly equal for $n = 1$ and for $n > 1$ the Polya distribution has larger $P[\underline{x} = 0]$.

The possibility is considered of extending this study, for example by considering the moments of $\underline{x}^{\underline{y}*}$ for which we have $\mathcal{E}\underline{x}^{\underline{y}*} = \mathcal{E}\underline{x} \cdot \mathcal{E}\underline{y}$ and $\sigma^2(\underline{x}^{\underline{y}*}) = \mathcal{E}\underline{y}\sigma^2(\underline{x}) + (\mathcal{E}\underline{x})^2\sigma^2(\underline{y})$.

I want to thank Professor Hemelrijk and Dr. van Zwet, who have read the manuscript and have suggested various improvements.

Appendix 1

List of mixtures and generalizations

For the definitions of the distributions see appendix 2.
This list gives some well-known examples; it is far from being complete.
As usual, q denotes 1 - p.

$$\text{Neyman A}(\lambda,\mu) = \text{Poisson }(k\mu)\underset{k}{\wedge}\text{Poisson }(\lambda) = [\text{Poisson }(\mu)]^{\text{Poisson }(\lambda)*} =$$
$$= [\text{Binomial }(n,p)\underset{n}{\wedge}\text{Poisson }(\mu p^{-1})]^{\text{Poisson }(\lambda)*}$$

$$\text{Pascal }(\gamma,p) = \text{Poisson }(\lambda)\underset{\lambda}{\wedge}\text{Gamma }(pq^{-1},\gamma) = [\text{Poisson }(1)]^{\text{Gamma }(pq^{-1},\gamma)*} =$$
$$= [\text{Poisson }(pq^{-1})]^{\text{Gamma }(1,\gamma)*} = [\text{Log }(p)]^{\text{Poisson }(-\gamma \log q)*}$$

$$\text{Poisson Pascal }(\lambda,\gamma,p) = \text{Pascal }(k\gamma,p)\underset{k}{\wedge}\text{Poisson }(\lambda) = [\text{Pascal }(\gamma,p)]^{\text{Poisson }(\lambda)*}$$

$$\text{Poisson Binomial }(\lambda,n,p) = \text{Binomial }(kn,p)\underset{k}{\wedge}\text{Poisson }(\lambda) =$$
$$= [\text{Binomial }(n,p)]^{\text{Poisson }(\lambda)*}$$

$$\text{Pascal }(cy,p)\underset{y}{\wedge}\text{Gamma }(\beta,\gamma) = [\text{Log }(p)]^{\text{Pascal }\left(\gamma,\frac{-c\beta \log q}{1-c\beta\log q}\right)*} =$$
$$= [\text{Pascal }(c,p)]^{\text{Gamma }(\beta,\gamma)*}$$

$$\text{Poisson }(\mu p) = \text{Binomial }(n,p)\underset{n}{\wedge}\text{Poisson }(\mu) = [\text{Binomial }(1,p)]^{\text{Poisson }(\mu)*}$$

$$\text{Polya }(n,r,s) = \text{Binomial }(n,p)\underset{p}{\wedge}\text{Beta }(r,s)$$

$$\text{Gamma }(1,\gamma) = \text{Gamma }(a^{-1},\gamma + j)\underset{j}{\wedge}\text{Pascal }(\gamma,1 - a^{-1})$$

$$\text{Gurland }(\alpha,\beta,\mu) = \text{Poisson }(\mu p)\underset{p}{\wedge}\text{Beta }(\alpha,\beta) = \text{Polya }(n,\alpha,\beta)\underset{n}{\wedge}\text{Poisson }(\mu) =$$
$$= (\text{Binomial }(n,p)\underset{p}{\wedge}\text{Beta }(\alpha,\beta))\underset{n}{\wedge}\text{Poisson }(\mu)$$

$$\text{Laplace}(1)=\text{Normal}(0,\sigma^2)\underset{\frac{1}{2}\sigma^2}{\wedge}\text{Gamma }(1,1)$$

Appendix 2

List of distributions

It will be obvious what is meant by Binomial (n,p), Degenerate in a, Normal (μ,σ^2) or Poisson (λ). The other distributions used are listed here.

NAME RESTRICTIONS	DENSITY OR PROBABILITIES [7]	EXPECTATION; VARIANCE; CHARACTERISTIC FUNCTION
Beta (r,s) $r > 0,\ s > 0$	$\frac{x^{r-1}(1-x)^{s-1}}{B(r,s)}$ $(0 \leq x \leq 1)$	$\frac{r}{r+s}$; $\frac{rs}{(r+s+1)(r+s)^2}$; [8]
Gamma (β,γ) $\beta > 0,\ \gamma > 0$	$\frac{e^{-x/\beta}\, x^{\gamma-1}}{\beta^{\gamma}\Gamma(\gamma)}$ $(x \geq 0)$	$\beta\gamma$; $\beta^2\gamma$; $(1-\beta it)^{-\gamma}$
Gurland (α,β,μ) $\alpha > 0,\ \beta > 0,\ \mu > 0$	[9] $(x = 0,\ 1,\ \ldots)$	$\frac{\alpha\mu}{\alpha+\beta}$; $\frac{\alpha\mu}{\alpha+\beta} + \frac{\alpha\beta\mu^2}{(\alpha+\beta+1)(\alpha+\beta)^2}$ ${}_1F_1(\alpha,\alpha+\beta,\mu(e^{it}-1))$
Laplace (β) $\beta > 0$	$\frac{1}{2\beta}\exp\left\{-\frac{\lvert x\rvert}{\beta}\right\}$	0; $2\beta^2$; $(1+\beta^2t^2)^{-1}$
Log (p) $0 < p < 1$	$\frac{p^x}{x\lvert\log q\rvert}$ $(x = 1,\ 2,\ \ldots)$	$\frac{p}{q\lvert\log q\rvert}$; $\frac{p(p+\log q)}{q^2\lvert\log q\rvert^2}$; $\log(1-pe^{it})/\log q$
Neyman A (λ,μ) $\lambda > 0,\ \mu > 0$	$\frac{e^{-\lambda}\mu^x}{x!}\sum_{k=0}^{\infty}\frac{k^x\lambda^k e^{-\mu k}}{k!}$ $(x = 0,\ 1,\ \ldots)$	$\lambda\mu$; $\lambda\mu(\mu+1)$; $\exp(\lambda\{e^{\mu(e^{it}-1)}-1\})$

---

7) For arguments not mentioned the value is zero.

8) $\frac{\Gamma(p+q)}{\Gamma(p)}\sum_{j=0}^{\infty}\frac{\Gamma(p+j)(it)^j}{\Gamma(p+q+j)\Gamma(j+1)}$

9) Recurrence relations for $P[\underline{x} = x]$ are given by GURLAND (1958).

NAME RESTRICTIONS	DENSITY OR PROBABILITIES [7]	EXPECTATION; VARIANCE; CHARACTERISTIC FUNCTION
Pascal (γ, p) [10] $\gamma > 0,\ 0 < p < 1$	$\frac{\Gamma(\gamma+x)}{\Gamma(\gamma)x!} p^x q^\gamma$ $(x = 0, 1, \ldots)$	$\gamma p q^{-1}$; $\gamma p q^{-2}$; $q^\gamma (1-pe^{it})^{-\gamma}$
Poisson Binomial (λ, n, p) $\lambda > 0$, integer $n > 0$, $0 < p < 1$	$e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \binom{n\,k}{x} p^x q^{nk-x}$ $(x = 0, 1, \ldots)$	λnp; $\lambda n^2 p^2 + \lambda npq$; $\exp(\lambda\{(pe^{it}+q)^n - 1\})$
Poisson Pascal (λ, γ, p) $\lambda > 0,\ \gamma > 0,\ 0 < p < 1$	[11] $(x = 0, 1, \ldots)$	$\gamma\lambda p q^{-1}$; $\gamma\lambda p q^{-2}(\gamma p+1)$; $\exp(\lambda\{q^\gamma(1-pe^{it})^{-\gamma} - 1\})$
Polya (n, r, s)	$\binom{n}{j} \frac{B(r+x, s+n+x)}{B(r,s)}$	$\frac{nr}{r+s}$; $\frac{nrs(n+r+s)}{(r+s)^2(r+s+1)}$

References

W. FELLER (1943), On a general class of "contagious" distributions, Ann.Math.Stat. **14**, 389-400.

W. FELLER (1957), An introduction to probability theory and its applications, Vol.1, Second edition, Wiley, New York.

J. GURLAND (1957), Some interrelations among compound and generalized distributions, Biometrika **44**, 265-268.

J. GURLAND (1958), A generalized class of contagious distributions, Biometrics **14**, 229-249.

---

[10] This is the negative binomial, but we prefer the shorter name. For integer γ it is the distribution of the number of successes preceding the γ^{th} failure.

[11] Recurrence relations for $P[\underline{x} = x]$ are given by KATTI & GURLAND (1961).

S.K. KATTI & J. GURLAND (1961), The Poisson Pascal distribution, Biometrics 17, 527-538.

M. LOÊVE (1963), Probability theory, Third edition, Van Nostrand.

E. CANSADO MACEDA (1948), On the compound and generalized Poisson distributions, Ann.Math.Stat. 19, 414-416.

J.U. McGUIRE, T.A. BRINDLEY & T.A. BANCROFT (1957), The distribution of European corn borer larvae Pyrausta Nubilalis (Hbn.), in field corn, Biometrics 13, 65-78 and 14, 432-434.

J. NEYMAN (1939), On a new class of "contagious" distributions applicable in entomology and bacteriology, Ann.Math.Stat. 10, 35-57.

R. PYKE (1960), On centering infinitely divisible processes, Ann.Math.Stat. 31, 797-800.

J.G. SKELLAM (1952), Studies in statistical ecology, Biometrika 39, 346-362.

H. TEICHER (1954), On the convolution of distributions, Ann.Math.Stat. 25, 775-778.

H. TEICHER (1960), On the mixture of distributions, Ann.Math.Stat. 31, 55-73.