SANKHYĀ
THE INDIAN JOURNAL OF STATISTICS
SERIES B, VOLUME 38, PART 1, 1976

ON SIZE-BIASED SAMPLING AND RELATED FORM-ININVARIANT WEIGHTED DISTRIBUTIONS

By
G. P. PATIL
and
J. K. ORD

STATISTICAL PUBLISHING SOCIETY
CALCUTTA
ON SIZE-BIASED SAMPLING AND RELATED FORM-ININVARIANT WEIGHTED DISTRIBUTIONS

By G. P. PATIL

The Pennsylvania State University

and

J. K. ORD

The University of Warwick

SUMMARY. The purpose of this paper is to inter-relate further the concepts of size biased sampling and weighted distributions by identifying some of the situations where the underlying models retain their form, even when they are subjected to size-biased sampling in preference to the usual random sampling.

1. BACKGROUND

Consider a finite or infinite population of units carrying values of a non-negative random variable $X(\theta)$ having distribution with probability (density) function (pdf) $f(x; \theta)$ where $x \geq 0$ and $\theta \in \Omega$, the parameter space defined for $f$. In the usual random sampling on $X = X(\theta)$, the probability of selection of each unit is the same, regardless of the value of $x$ it carries, so that the pdf at the observation $x$ is $f(x; \theta)$. In the size-biased sampling on $X = X(\theta)$, the probability of selection of a unit is proportional to a predetermined weight-function such as $w(x) = x$ corresponding to the value of $x$ that the unit carries, implying the pdf at the observation $x$ to be $f^*(x; \theta) = w(x)f(x; \theta)/\Sigma xw(x)f(x; \theta)$ or $w(x)f(x; \theta) \int w(\xi)f(x; \theta)d\xi$ depending on whether $X$ is discrete or continuous. In this situation, we shall say that $X$ is size-biased and we shall denote the size-biased version by $X^*$, which we will write sometimes as $(X(\theta))^s$, when the parameter $\theta$ needs special attention. It is clear that the pdf of $X^*$ is $f^*$, the pdf of the weighted distribution defined by the original distribution having pdf $f$ together with the identity weight-function $w$. We may also emphasize here that the $x$-value of the unit is not the well-known ancillary variable of the $ggs$ sampling, but is itself the variable observed and recorded.


Prepared with partial support of the NATO Scientific Affairs Division.
SIZE-BIASED SAMPLING AND WEIGHTED DISTRIBUTIONS

The publications listed at the end of the paper will indicate that the concept of size-biased sampling has begun only recently to receive formal attention in both the statistical and the applied literature. The concept of weighted distributions has been introduced and formalized recently, see for example, Rao (1965). The purpose of this paper is to inter-relate further the concepts of size-biased sampling and weighted distributions by identifying some of the situations where the underlying models retain their form, even when they are subjected to size-biased sampling in preference to the usual random sampling.

2. Introduction

If the non-negative random variable \( X \) has pdf \( f(x) \), then the first moment distribution for \( X \) may be defined to have the pdf

\[
f^*(x) = f_1(x) = x f(x)/\mu,
\]

provided that \( \mu = \mu_1 = E[X] \) exists. The extension to the \( j \)-th moment distributions with pdf \( f_j(x) \) is easily made, provided that \( \mu_j = E[X^j] < \infty \). For discrete distributions, the factorial moments arise more naturally; that is, we could use \( \mu_{(j)} = E[X^j] = E[X(X-1) \ldots (X-j+1)] \), defining the corresponding pdf for the \( j \)-th factorial moment distribution.

Moment distributions arise through size-biased sampling in many areas of application, as the following examples show.

(i) In the investigation of families through the detection of abnormal children, size-biased sampling is introduced, since the probability of selection of a unit is itself a function of the character under study. Rao (1965) observes this and refers to distributions with pdf of the form (2.1) as weighted distributions. He also allows fractional powers for discrete as well as continuous models. Schaeffer (1972) provides an example for fractional power \( \frac{2}{3} \) in a situation, where \( \alpha \) is the volume of a spherical element, and the sampling mechanism selects elements with probabilities proportional to surface area.

(ii) Exclusive angle sampling of trees to estimate the volume of timber in a forest; see Kendall and Moran (1963), Warren (1975).

(iii) The analysis of size differences in economics (size of firms, income distribution, and so on); see Aitchison and Brown (1957).

(iv) The analysis of particle size; see Schultz (1975).

(v) The analysis of life-length studies in reliability; see Schaeffer (1972), Cox (1962), Wallach and Siegel (1963).
The analysis of life-length studies in biomedicine; see Zelen (1971), Simon (1975).

In ecological and environmental work; see Pielou (1969), Belle and Schinssiermann (1973), Ord (1975).

In these situations, it is of interest to know whether $f(x)$ and $f_2(x)$ correspond to distributions of the same form, parameter(s) apart, so that data collected under size-biased sampling can readily relate back to the original population. At a theoretical level, it is of interest to know whether such a form-invariant property is characteristic of certain (classes of) distributions. The term of form-invariance requires a more rigorous definition. Clearly, the class of all pdf’s contains $f(x)$ and all valid $f_2(x)$ relating to it, while at the other extreme, if we require $f(x) = f_2(x)$ for all $x$, then we are left with a degenerate distribution. In this paper, we shall consider situations where $f_2(x)$ can be reduced to $f(x)$ by changing one or more existing parameter(s) $\theta$ of the model $f(x) = f_1(x; \theta)$. Our main emphasis will be on $f_1(x)$, equivalently written as $f^*(x)$.

3. Notation, terminology and preliminary results

3.1. Let the original distribution of the non-negative random variable $X$ have the pdf $f(x; \theta)$ with a scalar or vector parameter $\theta$. Sometimes we may call $X = X(\theta)$ the original random variable. Let $E[X] = \mu(\theta) < \infty$. The observed random variable under size-biased sampling of order one with $w(x) = x$ will be denoted by $X^* = (X(\theta))^*$. We will write the pdf of the observed distribution as $f^*(x; \theta)$. Clearly,

$$f^*(x; \theta) = x f(x; \theta) / \mu(\theta).$$  \hfill (3.1)

More generally, to define the size-biased sampling of order $\alpha$ on $x$ with parameter $\theta$, let $E[X^\alpha] = \mu'_\alpha(\theta) < \infty$. The corresponding pdf is then

$$f^*_\alpha(x; \theta) = x^\alpha f(x; \theta) / \mu'_\alpha(\theta).$$  \hfill (3.2)

We will denote the corresponding observed variable by $(X(\theta))^{*\alpha}$. It is clear that when $\alpha = 1$, $f^*_1(x; \theta) = f^*(x; \theta)$, and further $(X(\theta))^{*1} = (X(\theta))^*$.  

3.2. As in 3.1, let $X(\theta)$ have the pdf $f(x; \theta)$. We say that the distribution with pdf $f(x; \theta)$ is form-invariant under size-biased sampling of order $\alpha$ if the observed variable $(X(\theta))^{*\alpha}$ is equivalent in distribution to the variable $X(\eta)$, the original variable with parameter $\eta$. That is, in symbols, $(X(\theta))^{*\alpha} = X(\eta)$ implying

$$f^*_\alpha(x; \theta) = f(x, \eta),$$ \hfill (3.3)
where $f^*$ and $f$ are defined in (3.2). Thus, the form-invariance defined by (3.3) retains the functional form of the pdf $f$ allowing for the possibility of a change in the value of the parameter from $\theta$ to $\eta$.

3.3. It is clear that, for a given $f$, $\eta$ is a function of $\alpha$ with $\eta = \theta$ when $\alpha = 0$. We will say that $f(x; \theta)$ satisfies the continuity condition relative to $\alpha$ when $\eta \to \theta$ if and only if $\alpha \to 0$. That is, $0 < \lim_{\alpha \to 0} \left( \frac{\eta - \theta}{\alpha} \right) < \infty$ and $0 < \lim_{\eta \to \theta} \left( \frac{\alpha}{\eta - \theta} \right) < \infty$. Also, assuming the usual regularity condition of interchange of orders of integration with those of differentiation and limit-taking, we note that

$$
\lim_{\alpha \to 0} \left( \frac{\log \mu_\alpha(\theta)}{\alpha} \right) = \lim_{\alpha \to 0} \left( \frac{d}{d\alpha} \int_0^\infty x^\alpha f(x; \theta)dx \right)
$$

$$
= \lim_{\alpha \to 0} \int_0^\infty x^\alpha \log x f(x; \theta)dx
$$

$$
= \int \log x f(x; \theta)dx
$$

$$
= \mathbb{E}[\log X]. \quad \ldots \quad (3.4)
$$

We will say that $f(x; \theta)$ satisfies the expectation condition relative to $\alpha$ if

$$
\lim_{\alpha \to 0} \left( |(\log \mu_\alpha(\theta))/\alpha| \right) = |\mathbb{E}[\log X]| < \infty.
$$

3.4. Let the moment generating function (mgf) of $X$ with pdf $f(x; \theta)$ be $M_x(t) = \mathbb{E}[e^{tx}]$. Then the mgf of $X^*+t$ is given by

$$
M_{x+t}(t) = \left[ \frac{d}{dt} M_x(t) \right] \left[ \frac{d}{dt} M_x(t) \right]_{t=0}. \quad \ldots \quad (3.5)
$$

Clearly

$$
M_{X^*}(t) = \left[ \frac{d}{dt} M_X(t) \right] \left[ \frac{d}{dt} M_X(t) \right]_{t=0} = \frac{M_X(t)}{M_X(0)}. \quad \ldots \quad (3.6)
$$

3.5. Let $X$ be a non-negative integer-valued random variable with zero in its range. Clearly, $X^*$ will have no zero in its range. Let us define the displacement of $X^*$ by one to the left by the random variable $X_\alpha = X^* - 1$. Thus, $X_\alpha$ would be realized by size-biased sampling on $X$ with displacement as indicated.
Let the probability generating function (pgf) of $X$ with pdf $f(x; \theta)$ be $G_X(t) = E[t^X]$, for $0 < t < 1$. Then the pgf of $X^* \sim \theta$ is given by

$$G_{X^*}(t) = \left[ \frac{\partial}{\partial \theta} G_X(t) \right] \left[ \frac{\partial}{\partial \theta} G_X(t) \right]^{-1} = \frac{G_X'(t)}{G_X(1)}. \quad \ldots \ (3.7)$$

3.6. Let $X$ have the power series distribution in $\theta \in \Omega$ with series function $g(\theta) = \sum_{x=0}^{\infty} a(x) \theta^x$. Then its pgf is $G_X(t) = \frac{g(t\theta)}{g(\theta)}$ and $E[X] = \theta g'(\theta)/g(\theta)$. It follows from (3.7) that $X^*$ has the power series distribution in $\theta$ with series function $\theta g'(\theta)$, whereas $X^* \sim \theta$ has the power series distribution in $\theta$ with series function $g(\theta)$, where $'$ denotes differentiation with respect to $\theta \in \Omega$, where $\Omega = [0, \rho)$, $\rho$ being the radius of convergence of $g(\theta)$.

4. Form-invariance under size-bias of order $\alpha$ and log-exponential family

In this section, we show that the form-invariance under size-bias of any order $\alpha$ implies under certain regularity conditions that the common pdf belongs to what we call the log-exponential family.

Theorem 1: Let the pdf $f(x; \theta)$ of $X$ satisfy the regularity conditions of continuity and expectation relative to $\alpha$ as defined in 3.3. Then a necessary and sufficient condition (nasc) for $X$ to be form-invariant under size-bias of order $\alpha$ is that its pdf is of the form

$$f(x; \theta) = x^\alpha \lambda(x)/m(\theta) = \exp(\theta \log x + A(x) - B(\theta)) \quad \ldots \ (4.1)$$

where $\lambda(x) = \exp[A(x)], m(\theta) = \exp[B(\theta)], E[ log X] = m'(0)\lambda(0) = B'(0)$. \ 

Proof: It is straightforward to show that the pdf defined by (4.1) satisfies the continuity and expectation conditions relative to $\alpha$, as also the property of form-invariance under size-bias of order $\alpha$ as defined by (3.3).

To prove the sufficiency, we observe that by taking the logarithms in (3.3),

$$\left( \frac{\alpha \log x}{\eta - \theta} \right) - \left( \frac{\log \lambda'(\theta)}{\alpha} \right) \left( \frac{\alpha}{\eta - \theta} \right) = -\left( \frac{\log f(x; \eta) - \log f(x; \theta)}{\eta - \theta} \right).$$

Taking limits of both sides as $\eta \to \theta$ and applying the regularity conditions of continuity and expectation relative to $\alpha$, we get

$$b(\theta) \log x - c(\theta)b(\theta) = \frac{\partial}{\partial \theta} \log f(x; \theta),$$

from which the form of (4.1) follows. Here $b(\theta)$ and $c(\theta)$ are the appropriate limiting functions,
Comment 1: Since log x is the coefficient of \( b \) in the exponential form of the pdf defined by (4.1), this pdf may be said to define "Log-Exponential Family" analogous to the well-known "Linear Exponential Family". Some important members of the log-exponential family are listed in Table 1. From the theorem, it is evident that truncations at either or both ends of the range do not affect membership of the family.

<table>
<thead>
<tr>
<th>Table 1. Some members of the log-exponential family</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>distribution</strong></td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>lognormal</td>
</tr>
<tr>
<td>Pareto</td>
</tr>
<tr>
<td>gamma</td>
</tr>
<tr>
<td>beta (first kind)</td>
</tr>
<tr>
<td>beta (second kind)</td>
</tr>
<tr>
<td>Pearson type V</td>
</tr>
</tbody>
</table>

Comment 2: A feature of the log-exponential family is that the geometric mean is a sufficient statistic for \( \theta \) (cf. the same property of the arithmetic mean in the linear exponential family). This raises questions about the efficient estimation of the mean, as the sample mean is unbiased, but usually inefficient; see Aitchison and Brown (1957) who consider this problem for the lognormal distribution.

Corollary: Let \( Y = \log X \). Then the form-invariance under size-biased sampling according to \( \exp(\alpha Y) \) characterizes the linear exponential family.

5. Form-invariance of n-fold sums of iid variables under size-bias of order one

In this section, we characterize the gamma distribution within the class of distributions of \( n \)-fold sums of identically distributed independent random variables by the property of form-invariance relative to the parameter \( n \) under size-biased sampling of order one. We also characterize the Poisson, binomial and negative binomial distributions within the class of distributions of \( n \)-fold sums of identically distributed independent non-negative integer-valued random variables by the property of form-invariance relative to the parameter \( n \) under size-biased sampling of order one with displacement. We prove the following:

Theorem 2: Let \( X(n) = \sum_{i=1}^{n} Z_i \), where \( Z_i \)'s are independently and identically distributed with common mean \( \mu \). Then a nasc for \( X(n) \) to be form-invariant
under size-bias of order one given by \((X(n)) \overset{d}{=} X(n+1)\) is that the \(Z_i\)'s have the negative exponential distribution with parameter \(\mu\), implying that \(X(n)\) has the gamma distribution with parameters \(\mu\) and \(n\).

**Proof:** It is straightforward to prove the necessity. For sufficiency, let \(M(t)\) be the common mgf of \(Z_i\)'s. Then,

\[
M_{X(n)}(t) = [M(t)]^n.
\]

Using (3.8), we get

\[
\frac{\partial}{\partial t} [M(t)]^n = \frac{\partial}{\partial t} [M(t)]^{n+1}.
\]

The solution upon simplification gives \(M(t) = (c-\mu t)^{-1} = (1-\mu t)^{-1}\). Hence the theorem.

On the same lines as above, but using the pgf's, we obtain the following theorem for size-biased sampling with displacement.

**Theorem 3:** Let \(X(n) = \sum_{i=1}^{n} Z_i\), where \(Z_i\)'s are independently and identically distributed non-negative integer-valued random variables with a common mean \(\mu\).

1. A necessary for \((X(n)) \overset{d}{=} X(n+1)\) for some \(n \geq 1\) is that the \(Z_i\)'s have the geometric distribution with parameter \(p = 1/(1+\mu)\), implying that \(X(n)\) has the negative binomial distribution with parameters \(n\) and \(p\).
2. A necessary for \((X(n)) \overset{d}{=} X(n-1)\) for some \(n \geq 1\) is that the \(Z_i\)'s have Bernoulli distribution with parameter \(p = \mu\), implying that \(X(n)\) has the binomial distribution with parameters \(n\) and \(p\).
3. A necessary for \((X(n)) \overset{d}{=} X(n)\) for some \(n \geq 1\) is that the \(Z_i\)'s have the Poisson distribution with parameter \(\mu\), implying that \(X(n)\) has the Poisson distribution with parameter \(n\mu\).

**6. Mean-invariance and a characterization of the Poisson distribution**

In this section, we prove a modification of (3c) of Theorem 3. Instead of requiring the form-invariance, mean-invariance is considered. The class of distributions is assumed to consist of the power series distributions instead of the \(n\)-fold sum distributions.
Theorem 4: Let \( X \) have the power series distribution in \( \theta \) with series function \( g(\theta) \). Then \( \theta \) is a bone for \( X \) to be mean-invariant under size-bias of order one with displacement defined by \( E[X_\theta] = E[X] \) is that \( X \) has the Poisson distribution with parameter \( \theta \).

Proof: The necessity follows from the fact that for the Poisson distribution, \( g(\theta) = \exp(\theta) \) and hence \( g'(\theta) = g(\theta) \), implying the mean-invariance of \( X \) because of (3.6). To prove the sufficiency, we note that

\[
E[X] = \theta g'(\theta)/g(\theta) = \theta g'(\theta)/g' = E[X_\theta].
\]

Now, observe that

\[
\frac{d}{d\theta} \left( \frac{g'(\theta)}{g(\theta)} \right) = \frac{g''(\theta)}{g(\theta)} - \left( \frac{g'(\theta)}{g(\theta)} \right)^2 = 0,
\]

leading to the exponentiality of \( g(\theta) \) and hence to the Poissonness of the distribution.

7. Form-invariance for cluster models with Poisson frequency of clusters

In this section, we prove certain modifications of (3a) and (3c) of Theorem 3. While the form-invariance under size-bias of order one with displacement is used as before, the class of distributions is assumed to consist of the distributions of random sums instead of \( n \)-fold sums. We prove the following theorem.

Theorem 5: Let \( X(\lambda) = Y_1 + Y_2 + \cdots + Y_N \) where the cluster frequency \( N \) has the Poisson distribution with parameter \( \lambda \). The cluster sizes, \( Y_i \)'s, are identically and independently distributed and are independent of \( N \).

(5a) A bone for \( X(\lambda) \) to be form-invariant defined by \( (X(\lambda))_\theta \distr X(\lambda^*) \) where \( \lambda^* = \lambda \) is that \( X(\lambda) \) has Poisson distribution.

(5b) A bone for \( X(\lambda) \) to be form-invariant defined by \( (X(\lambda))_\theta \distr X(\lambda^*) \) where \( \lambda^* > \lambda \) is that \( X(\lambda) \) has negative binomial distribution.

Proof: It is straightforward to prove the necessity. To prove the sufficiency, we first observe that \( G(t) = e^{-\lambda t} \). Therefore, if \( G(t) = g(t) \),

\[
G(t) = e^{-\lambda t} g(t). \quad \text{and} \quad \frac{\lambda g'(t)}{g(1)} = e^{-\lambda t} \lambda g(t),
\]

leading to

\[
\frac{g'(t)}{a} = [e^{g(t) - 1}],
\]

where \( a = g'(1) > 0 \) and \( \sigma = \lambda^* - \lambda \).
Now, we note that both \( g(t) \) and \( g'(t) \) are monotone increasing in \( t \). Therefore, \( c > 0 \).

(a) When \( c = 0 \), we have the case for (5a), in which case \( g'(t) = \alpha \) implying that \( g(t) = q + pt \) with \( 0 < q, q < 1 \) and \( p + q = 1 \), in which case \( X(\lambda) \) has Poisson distribution with parameter \( \lambda p \).

(b) When \( c > 0 \), we have the case for (5b), in which case \( g'(t) = y^c \), where \( y = e^{s(t) - t} \), gives \( \frac{1}{y} \left( \frac{dy}{dt} \right) = ay^c \). Solving and resubstituting,

\[
g(t) = 1 - \left( \frac{1}{c} \right) \log (-bc - act).
\]

\( g(1) = 1 \) gives \( g(t) = 1 - \left( \frac{1}{c} \right) \log (1 + ac - act) \)

\[
= p_0 + (1-p_0) \left[ \frac{-\log(1-\theta t)}{-\log(1-\theta)} \right]
\]

where \( \theta = \frac{ac}{1+ac} \) and \( 0 < p_0 = 1 - \left( \frac{1}{c} \right) \log (1 + ac) < 1 \) when \((ac-1)c > a\).

Thus the distribution of \( Y_1 \) is a modified logarithmic series distribution, in which case \( X(\lambda) \) can be shown to have the negative binomial distribution with corresponding parameters.

8. Further results for discrete distributions

We have introduced the concept of factorial moment distributions earlier in Section 2. Here we propose to consider some of their properties. Also, in this section, we will study a problem of comparing the size-biased sampling with "damage" as defined by Rao (1965) and investigated further by Patil and Ratnaparkhi (1975), among others.

8.1. Factorial moment distributions. Let \( X \) be a non-negative integer valued random variable with pdf \( f(x; \theta) \) and pgf \( G_X(\theta) \). We define the corresponding \( j \)-th factorial moment distribution by the pdf

\[
\tilde{f}_{(j)}^*(x) = \frac{x_j \theta^j}{\mu_j^*(\theta)} \quad \cdots \quad (8.1)
\]

where \( \mu_j^*(\theta) = E[X^j\theta] < \infty \) is the \( j \)-th factorial moment of \( X \). If we denote the size-biased random variable of factorial order \( j \) by \( X^*(j) \), its pdf is given by (8.1). Further, its pgf can be obtained as

\[
G_{X^*(j)}(\theta) = \left[ \frac{\partial}{\partial \theta} G_X(\theta) \right]_{\theta=1} \quad \cdots \quad (8.2)
\]
It is clear that \( f_{ij}(x) = 0 \) for \( x < j \). Also, it follows from (8.2) that
\[ Y_j = X^* - j \]
has its pgf given by
\[ G_{Y_j}(t) = \frac{\frac{\partial}{\partial t} G_X(t)}{\frac{\partial^i}{\partial t^i} G_X(t)_{i=1}}. \]
\[ \ldots \quad (8.3) \]

It is clear that the range of \( Y_j \) can include zero and one can write its pdf in terms of the pdf of \( X \) as
\[ h_j(y) = \binom{y+j}{j} \frac{j^j}{\mu_{g_j}(\theta)} \cdot f(y+j; \theta). \]
\[ \ldots \quad (8.4) \]

Equation (8.4) incorporates a wide range of discrete distributions, including the hypergeometric series distributions, the Poisson and the binomials as shown in Table 2 (see also Rao, 1965). We may observe that all the tabulated distributions with the exception of the logarithmic series distribution are form-invariant under the size-bias of factorial order \( j \).

8.2. Damage and size-biased sampling. If the random variable \( X \) has pgf \( G_X(t) \) and the observations are subject to binomial damage (see Rao, 1966), the pgf of the recorded variate \( Y \) is \( G_Y(t) = G_X(q+pt) \), where \( p \) is the probability that the original observation is recorded. As noted in (3.7), the size-biased variate with displacement denoted by \( X_\ast \) has pgf \( G^{X_\ast}_X(t)/G_X(1) \).

If the question arises as to whether \( Y \overset{d}{=} X_\ast \), a partial answer is given in the following theorem. Note that we do not assume that \( X \overset{d}{=} Z \) or \( X_\ast \), as the result would then be trivial.

Theorem 6: Let \( X_t \) have the power series distributions in \( \theta_t \) with series functions \( g(\theta^i) \) satisfying \( g^{(i)}(0)|g^{(i)}(0) = g^{(i)}(0)|g^{(i)}(0) \) for \( i = 1, 2 \). Let \( X_1 \) be subjected to binomial damage with parameter \( p \) leading to \( Y \). Let \( X_2 \) be subjected to size-bias with displacement leading to \( Z = X_\ast \). Then a noise for \( Y \overset{d}{=} Z \) is that both \( X_1 \) and \( X_\ast \) (equivalently both \( Y \) and \( Z \) also) are Poisson distributed with \( \theta_2 = \theta_1 p \).

Proof: It is straightforward to prove the necessity. For sufficiency, we must solve the equation
\[ g(\theta_1(q+pt))|g(\theta_1) = g'(\theta_1)|g'(\theta_2), \quad 0 \leq t \leq 1. \]
\[ \ldots \quad (8.5) \]
Comparing coefficients of \( \theta^i \) and writing \( \phi = \theta q \), we find that
\[ (\theta_1 p)^{x} \left( \sum_{t=0}^{\infty} \frac{(s+x)}{x} a(s+x) \phi^t \right) / g(\theta_1) = (x+1) a(x+1) \theta_2^x / g'(\theta_2), \]
\[ \ldots \quad (8.6) \]
where \( g(\theta) = \Sigma a(x) \theta^x \).
But, we observe that
\[ \sum_{s=0}^{s=x} \binom{s}{x} \alpha(s+1)(\phi^s)/x! = g^{(\phi)}(\phi)/x!. \]
Thus (8.6) simplifies to
\[ (\theta_1 g')^{s=x} g^{(\phi)}(\phi)/x! g(\theta_1) = (x+1)\alpha(x+1)\beta_2 g'(\theta_2). \]  ... (8.7)
Writing \( x = 0 \), we get \( g(\phi)/g(\theta_1) = \alpha(1)g'(\theta_1) \), and therefore (8.7) can be simplified to
\[ (\theta_1 g')^{s=x} g^{(\phi)}(\phi)/x! = (x+1)\alpha(x+1)\beta_2 g(\phi)/\alpha(1). \]  ... (8.8)
Using \( x = 0, 1, 2 \) in (8.8) and eliminating \( \theta_2 \) yields
\[ g(\phi)/g(\phi) = c g'(\phi)/g(\phi) \]  ... (8.9)
where \( c = 3\alpha(3)\alpha(1)/2!\alpha(2)^2 \) = 1 because of the assumed property of the power series distributions.

The solution of (8.9) with \( c = 1 \) is of the form \( g(\phi) = b e^{k\phi} \) implying the required Poissonness.

To show that \( \theta_3 = \theta_1 g' \), observe that \( E[Y] = \theta_1 g'(\theta_1)/g(\theta_1) \) and \( E[Z] \theta_2 g''(\theta_2)/g'(\theta_2) \). Substituting \( g(\phi) = b e^{k\phi} \) in \( E[Y] = E[Z] \), we obtain \( \theta_3 = \theta_1 g' \).

**Table 2. Some examples of weighted discrete distributions**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>f(x)</th>
<th>( f(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binomial, ( B(n, p) )</td>
<td>( \binom{n}{x} p^x (1-p)^{n-x} )</td>
<td>( B(n-j, 2) )</td>
</tr>
<tr>
<td>Negative Binomial, ( NB(k, q) )</td>
<td>( \frac{k+x-1}{x} q^x )</td>
<td>( NB(k+j, q) )</td>
</tr>
<tr>
<td>Poisson, ( P(\lambda) )</td>
<td>( e^{-\lambda} \lambda^x/x! )</td>
<td>( P(\lambda) )</td>
</tr>
<tr>
<td>Logarithmic Series, ( L(\alpha) )</td>
<td>( -\log (1-\alpha)^{-1} \alpha^x/x )</td>
<td>( NB(j, \alpha) )</td>
</tr>
<tr>
<td>Hypergeometric, ( H(n, M, N) )</td>
<td>( \binom{n}{x} H^* (N-M(n-x))/\binom{n}{x} )</td>
<td>( H(n-j, M-j, N-j) )</td>
</tr>
<tr>
<td>Binomial Beta, ( BB(\alpha, \gamma) )</td>
<td>( \binom{n}{x} \beta(\alpha+x, \gamma+n-x)/\beta(\alpha, \gamma) )</td>
<td>( BB(n-1, \alpha, \gamma) )</td>
</tr>
<tr>
<td>Pascal Beta, ( PB(k, \alpha, \gamma) )</td>
<td>( \binom{k+x-1}{x} \beta(\alpha+x, \gamma+k)/\beta(\alpha, \gamma) )</td>
<td>( PB(k+1, \alpha, \gamma) )</td>
</tr>
</tbody>
</table>

9. Form-invariance for continuous order statistics

Let \( Y \) be an absolutely continuous non-negative random variable with distribution function \( G(y) \). Let \( X(r, n) \) be the \( r \)-th order statistic based on a random sample of size \( n \) on \( Y \). Let \( E[X(r, n)] = \mu(r, n) < \infty \). We want to study the form-invariance of \( X(r, n) \) under size-bias of order one defined by
\[ (X(r, n))^n \overset{d}{=} X(r+j, n-k) \]  ... (9.1)
where \( j \) and \( k \) are suitably chosen non-negative integers satisfying \( 0 < r+j \leq n-k \), \( 0 < r < n \), \( k \leq n \), \( n > 1 \). We ask the question if (9.1) uniquely determines \( \mathcal{G} \) and hence the distribution of \( X = X(r, n) \). In this section, we attempt to answer this question for a few special cases involving \( j = 0 \) and \( j = 1 \). First of all, we note that for given \( j \) and \( k \), the pdf of \( X(r, n) \) is given by

\[
\varrho_{r, m}(x) = \frac{1}{B(r, n-r+1)} [G(x)]^{r-1}[1-G(x)]^{n-r}G'(x) \quad \ldots \quad (9.2)
\]

where \( B(r, n-r+1) \) is the beta function with arguments \( r \) and \( n-r+1 \). It is clear that in terms of the involved pdf's, (9.1) is equivalent to

\[
xG_{r, m}(x)/\mu(r, n) = g_{r+n-1-k}(x). \quad \ldots \quad (9.3)
\]

Quite easily, (9.3) simplifies to

\[
c(r, n; j, k)x^{j+k} = [G(x)]^j. \quad \ldots \quad (9.4)
\]

Because (9.4) is an identity which is true for arbitrary \( r, n \) to have started with, \( c(r, n; j, k) \) does not depend on \( r \) and \( n \), and therefore \( c(r, n; j, k) = c(j, k) \), say. (9.4) then reduces to

\[
c(j, k)x^{j+k} = [G(x)]^j. \quad \ldots \quad (9.5)
\]

We now state and prove the following results in Theorems 7 and 8 involving uniform, beta and Pareto distributions.

**Theorem 7**: Let \( X(r, n) \) be the \( r \)-th order statistic based on a random sample of size \( n \) on a non-negative continuous random variable \( Y \).

(7a) Then a nasc for \( (X(r, n))^* \xrightarrow{d} X(r, n-k) \) for \( k \in \{1, 2, \ldots, n-1\} \) is that the \( k \)-th power of \( X(r, n) \) has a beta distribution of the first kind.

(7b) Then a nasc for \( (X(r, n))^* \xrightarrow{d} X(r+1, n+1) \) is that a scalar multiple of \( X(r, n) \) has a beta distribution of the first kind.

(7c) Then a nasc for \( (X(r, n))^* \xrightarrow{d} X(r+1, n) \) is that a translation of \( Y \) has a Pareto distribution.

**Proof**: Follows from (9.5) with needed substitutions and simplifications.

**Theorem 8**: Let \( X(r, n) \) be the \( r \)-th order statistic based on a random sample of size \( n \) on a non-negative continuous random variable \( Y \). Then, \( (X(r, n))^* \xrightarrow{d} X(r+1, n-k) \) implies that \( P(Y \leq y) < 1 - \left( \frac{\alpha}{y} \right)^{1+k} \), which is the d.f. of the Pareto distribution with parameters \( \alpha \) and \( 1/(1+k) \), where \( \alpha \) is some constant.
Proof: Writing \( c(1, k) = \alpha \), (9.5) provides
\[
\alpha x [1 - G(x)]^{1 + k} = G(x).
\]
Differentiations of the two sides give
\[
\alpha [1 - G(x)]^k [1 - G(x) - \alpha (1 + k) G'(x)] = G'(x).
\]
It is obvious that \( 1 - G(x) - \alpha (1 + k) G'(x) > 0 \), implying
\[
\frac{G'(x)}{1 - G(x)} < -\frac{1}{\alpha (1 + k)}
\]
from which it follows that \( G(x) < 1 - \left( \frac{\alpha}{x} \right)^{1 + k} \), where \( \alpha \) is an arbitrary constant.

REFERENCES


SIZE-BIASED SAMPLING AND WEIGHTED DISTRIBUTIONS


