1999

Fisher Information in Weighted Distributions

Satish Iyengar

Paul H. Kvam
University of Richmond, pkvam@richmond.edu

Harshinder Singh

FISHER INFORMATION IN WEIGHTED DISTRIBUTIONS

Satish Iyengar*
Department of Statistics
University of Pittsburgh
Pittsburgh, PA 15260

Paul Kvam
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Harshinder Singh
Department of Statistics
Panjab University
Chandigarh, India 160014

May 1998

Corresponding author:
Satish Iyengar
Department of Statistics
518 Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260
Phone: 412-624-8341
Fax: 412-648-8814
email: si@satish.stat.pitt.edu

* Supported by NIMH grant MH-55123.

FISHER INFORMATION IN WEIGHTED DISTRIBUTIONS

ABSTRACT

Standard inference procedures assume a random sample from a population with density $f_\mu(x)$ for estimating the parameter μ. However, there are many applications in which the available data are a biased sample instead. Fisher modeled biased sampling using a weight function $w(x) \geq 0$, and constructed a weighted distribution with a density $f^w_\mu(x)$ that is proportional to $w(x)f_\mu(x)$. In this paper, we assume that $f_\mu(x)$ belongs to an exponential family, and study the Fisher information about μ in observations obtained from some commonly arising weighted distributions: (i) the k^{th} order statistic of a random sample of size m, (ii) observations from the stationary distribution of the residual lifetime of a renewal process, and (iii) truncated distributions. We give general conditions under which the weighted distribution has greater Fisher information than the original distribution, and specialize to the normal, gamma, and Weibull distributions. These conditions involve the distributions' hazard rate and the reversed hazard rate functions.

Key words and phrases: exponential family, hazard rate function, meta-analysis, order statistics, residual lifetime, reversed hazard rate function, selection model.

AMS 1991 subject classifications: 62B99, 62F10.

Running head: INFORMATION IN WEIGHTED DISTRIBUTIONS

1. Introduction

Let a random variable X have probability density function (pdf) $f_\theta(x)$, where θ is an unknown parameter taking on values in the parameter space Θ. The usual inference procedures assume that a random sample is drawn from a population with pdf $f_\theta(x)$. However, there are many situations in which the available data are a biased sample instead. Using a weight function, $w(x) \geq 0$, to model ascertainment bias, Fisher (1934) constructed the weighted distribution corresponding to $f_\theta(x)$, and having the form

$$f_\theta^w(y) = \frac{w(y) f_\theta(y)}{E_\theta\{w(X)\}}, \tag{1}$$

where the expectation in (1) is assumed to exist. For example, in meta-analysis $w(x) = I\{|x| \geq 1.96\}$ models an extreme form of publication bias, in which only statistically significant results are reported: see Hedges (1992), and Iyengar and Greenhouse (1988). In reliability theory, $w(x) = x$ models size-biased sampling for lifetime distributions. In general, the weight function may also depend upon another parameter, and upon θ itself. Rao (1965, 1985) presented a unified theory of weighted distributions, identifying various sampling situations which can be modeled using them.

Several studies have compared the information about θ in an experiment yielding a random sample from $f_\theta(x)$ with that in an experiment yielding a random sample of the same size from $f_\theta^w(x)$. Patil and Taillie (1987) calculated the Fisher information for certain exponential families, focusing primarily on $w(x) = x$ for nonnegative random variables. Bayarri and DeGroot (1987a, 1987b) and Bayarri et al. (1987) investigated Fisher information when observations are obtained only from selected portions of the underlying population, which was assumed to belong to an exponential family. They gave some explicit results for the normal and gamma distributions; for the gamma distribution, they noticed the role played by the hazard rate function when studying the Fisher information.

In this paper, we study Fisher information for two commonly arising classes of weighted distributions, which are conventionally not regarded as weighted distributions: the k^{th} order statistic from a random sample of size m from $f_\theta(x)$, and observations from the stationary distribution of the residual lifetime from a renewal process, when $f_\theta(x)$ belongs to a certain exponential family. We also extend the results of Bayarri et al. (1987) on selection samples to other distributions, and selection sets.

After establishing our notation in Section 2, in Section 3, we compare the Fisher information in the k^{th} order statistic, $X_{(k)}$, in a sample of size m from $f_\mu(x)$ with the information in a single observation from $f_\mu(x)$. At first blush, this comparison may not seem to be a fair one; however, there are many instances in which it is fair. For example, in order to reduce costs, many industries have curtailed their statistical testing. As a result, the availability of test data on the lifetimes of individual components of interest is reduced. Instead, the available data are lifetimes of systems that are made up of those components. If a system with identical working components is configured in series (parallel) the lifetime of the system corresponds to the first (last) order statistic. In general, if it is a k-out-of-m system, the lifetime of the system is the $(m - k + 1)^{th}$ order statistic. In a medical setting, the measurements of a patient's physical attributes such as the total breath volume for lung capacity are measured several times, with only the best measurement recorded. And in ballistic experiments, observations are obtained from experiments in which several projectiles are fired at a target; often only the worst shot, as measured by distance from the target, is used for further analysis. In this section, we first note that the pdf of $X_{(k)}$ is a weighted version of the pdf of X. We then provide a general condition for $X_{(k)}$ to have greater Fisher information than X, and illustrate the result using the normal location family, and the gamma and Weibull scale families.

In Section 4, we express the stationary distribution of the residual lifetime of a component as a weighted version of the lifetime of the component, with the weight function being the reciprocal of the hazard rate function of that component. If a renewal process is in operation for a long time, the remaining lifetime of an existing component may be regarded as an observation from the stationary residual lifetime distribution. In this section, we study the Fisher information about the unknown scale parameter of the gamma and Weibull distributions when the observations are drawn from a stationary residual distribution.

Bayarri et al. (1987) studied the Fisher information in selection models, in which $w(x) = I(x \in S)$, and S is the selection set. Such models arise in meta-analysis, where S represents statistically significant results. For positive random variables, selection in the lower tail $S = \{x : x < a\}$ also arises when a manufacturer decides to burn in the items for a units of time before sending them to market. Other examples of selection models are given in Bayarri and DeGroot (1987a), and Rao (1985). In Section 5, we extend the study of Bayarri et al. (1987) to selection samples from the gamma, Weibull, and exponential distributions.

There are, of course, many other applications of weighted distributions, most notably in sampling (see Rao, 1985) and nonresponse (see Little and Rubin, 1987). We do not pursue these matters here; we do note, however, that our results on selection samples below apply to the truncation problems discussed in Rao (1985).

2. Preliminaries

Under standard regularity conditions, the Fisher information $I_X(\theta)$ based on an observation X with pdf $f_\theta(x)$ is

$$I_X(\theta) = E_\theta\left\{-\frac{d^2 \log f_\theta(X)}{d\theta^2}\right\};$$

and the Fisher information in a random sample $X_1, \ldots, X_n$ from $f_\theta(x)$ is $nI_X(\theta)$. We assume that the pdf $f_\theta(x)$ belongs (perhaps after reparametrization) to the following exponential family of distributions,

$$f_\theta(x) = a(x) \exp\{\theta T(x) - C(\theta)\}; \tag{2}$$

in which case $I_X(\theta) = C''(\theta)$. For the weight function $w_\theta(y)$, let Y have the weighted distribution with pdf

$$f_\theta^{w_\theta}(y) = \frac{w_\theta(y) a(y) \exp\{\theta T(y) - C(\theta)\}}{E_\theta\{w_\theta(X)\}}.$$

Then

$$\begin{aligned} I_Y(\theta) &= C''(\theta) + \frac{d^2}{d\theta^2} \log E_\theta\{w_\theta(X)\} - E_\theta\left\{\frac{d^2}{d\theta^2} \log w_\theta(Y)\right\} \\ &= I_X(\theta) + \frac{d^2}{d\theta^2} \log E_\theta\{w_\theta(X)\} - E_\theta\left\{\frac{d^2}{d\theta^2} \log w_\theta(Y)\right\}. \end{aligned} \tag{3}$$

Two quantities that arise in our study are the hazard rate and reversed hazard rate functions. For a random variable X with pdf $f_\theta(x)$, cumulative distribution function (cdf)

$F_\theta(x)$, and $\bar{F}_\theta(x) = 1 - F_\theta(x)$, its hazard rate function is

$$\lambda_\theta(x) = \frac{f_\theta(x)}{\bar{F}_\theta(x)} = \frac{f_\theta(x)}{1 - F_\theta(x)};$$

for all x such that $F_\theta(x) < 1$, and its reversed hazard rate function is

$$\mu_\theta(x) = \frac{f_\theta(x)}{F_\theta(x)};$$

for all x such that $F_\theta(x) > 0$. Both of these functions are important in reliability theory and survival analysis. For positive random variables, $\bar{F}_\theta$ is known as the survival function, and in survival analysis the reversed hazard rate function is better known as the retrohazard rate function. For nonnegative random variables, F_θ is called an IFR (DFR) distribution if its hazard rate, or failure rate function is increasing (decreasing) on its interval of support; it is called a DRHR distribution if its reversed hazard rate function is decreasing on its interval of support. These monotonicity properties of the hazard rate and reversed hazard rate functions play a role in each of the sections below. For further discussion of these functions, see Anderson et al. (1993), Barlow and Proschan (1975), Block et al. (1998), and Shaked and Shantikumar (1994).

3. Order Statistics

Let $X_1, \ldots, X_m$ be independent and identically distributed observations with pdf $f_\theta(x)$ which has the exponential form in (2). Then the pdf of the k^{th} order statistic, $Y = X_{(k)}$, is (see David, 1970)

$$g_\theta(y) = k \binom{m}{k} F_\theta(y)^{k-1} \bar{F}_\theta(y)^{m-k} f_\theta(y).$$

Notice that $g_\theta(y)$ is a weighted version of $f_\theta(y)$, with weight function

$$w(y) = w_\theta(y) = F_\theta(y)^{k-1} \bar{F}_\theta(y)^{m-k}$$

depending on θ, and that the normalizing constant $E_\theta\{w_\theta(X)\}$ is independent of θ. From (3), the Fisher information about θ in Y is

$$I_Y(\theta) = I_X(\theta) - E_\theta\left[\frac{d^2}{d\theta^2}\left\{(k-1)\log F_\theta(Y) + (m-k)\log \bar{F}_\theta(Y)\right\}\right].$$

Thus, we have the following theorem and examples illustrating its use.

Theorem 3.1 Let $X_1, \ldots, X_m$ be i.i.d. observations with pdf $f_\theta(x)$ which has the exponential family form in (2). Let $Y = X_{(k)}$ be the k^{th} order statistic from that sample. Then $I_Y(\theta) \geq (\leq) I_X(\theta)$ if for every y the function

$$(k-1)\log F_\theta(y) + (m-k)\log \bar{F}_\theta(y) \tag{4}$$

is a concave (convex) function of θ. The inequality is strict if in addition, for each θ the function in (4) is strictly concave (convex) for y in a set of positive probability under g_θ.

Example 3.2: Normal. When $f_\theta(x)$ has the form (2) and is also a location family $f_\theta(x) = f_0(x-\theta)$, f_0 must be a normal pdf. This fact follows from methods similar to those used to prove Theorem 8.5.1 of Kagan et al. (1973); we omit the proof. We assume that the variance, σ^2, is known; setting $\sigma = 1$, the second derivative of (4) is

$$\frac{d^2}{d\theta^2}\left\{(k-1)\log F_\theta(y) + (m-k)\log \bar{F}_\theta(y)\right\} = (k-1)\mu_0'(y-\theta) - (m-k)\lambda_0'(y-\theta);$$

where μ_0 and λ_0 are the reversed hazard rate and hazard rate functions, respectively, of the standard normal distribution with cdf Φ. Because of symmetry of the normal pdf, the hazard rate and reversed hazard rate functions are related thus:

$$\mu_0(x) = \frac{\phi(x)}{\Phi(x)} = \frac{\phi(-x)}{1-\Phi(-x)} = \lambda_0(-x) = \frac{1}{M(-x)};$$

where $M(x) = 1/\lambda(x)$ is Mills' ratio for the standard normal distribution. It is known that M is a strictly decreasing function for all x: see Iyengar (1986). Thus, μ_0 is strictly decreasing, and λ_0 is strictly increasing everywhere. Therefore, the k^{th} normal order statistic of a random sample of size $m > 1$, $1 \leq k \leq m$ always gives greater Fisher information about the mean than an individual observation.

Example 3.3: Gamma. For a positive random variable, if $f_\theta(x)$ has the exponential form (2) and is also a scale family $f_\theta(x) = \theta f_1(\theta x)$ for all $\theta > 0$ and $x > 0$, then f_1 must be a gamma pdf; as before, this fact follows from from methods similar to those used to prove Theorem 8.5.2 of Kagan et al. (1973); we omit the proof. Thus, let

$$f_\theta(x) = \frac{1}{\Gamma(\alpha)}\theta^\alpha x^{\alpha-1}e^{-\theta x} \text{ for } x \geq 0. \tag{5}$$

In this case,

$$\frac{d^2}{d\theta^2}\left\{(k-1)\log F_\theta(y) + (m-k)\log \bar{F}_\theta(y)\right\} = (k-1)y^2\mu_1'(\theta y) - (m-k)y^2\lambda_1'(\theta y)$$

where μ_1 and λ_1 are the reversed hazard rate and hazard rate functions for F_1, respectively. From Theorem 3.1, we then have the following results. If F_1 is IFR, $I_{X_{(1)}}(\theta) \geq I_X(\theta)$ for all θ; if F_1 is DRHR, $I_{X_{(m)}}(\theta) \geq I_X(\theta)$ for all θ; and if F_1 is both IFR and DRHR, for $2 \leq k \leq m-1$, $I_{X_{(k)}}(\theta) \geq I_X(\theta)$ for all θ. The gamma is an IFR (DFR) distribution if $\alpha > (<) 1$, and it reduces to the exponential distribution for $\alpha = 1$: see Barlow and Proschan (1975). Since Block et al. (1998) have shown that f_1 is DRHR for every α, Theorem 3.1 then yields the following results for the gamma:

(i) For $\alpha = 1$ and all θ, $I_{X_{(1)}}(\theta) = I_X(\theta)$, and $I_{X_{(k)}}(\theta) > I_X(\theta)$ for $2 \leq k \leq m-1$.

(ii) For $\alpha < 1$ and all θ, $I_{X_{(1)}}(\theta) < I_X(\theta)$.

(iii) For all α and θ, $I_{X_{(m)}}(\theta) > I_X(\theta)$.

(iv) For $\alpha > 1$ all θ, $I_{X_{(k)}}(\theta) > I_X(\theta)$ for $1 \leq k \leq m-1$.

Example 3.4: Weibull. Let X be a Weibull random variable with unknown scale parameter θ and known shape parameter α. The pdf of X is

$$f_\theta(x) = \alpha\theta^\alpha x^{\alpha-1} e^{-(\theta x)^\alpha} \text{ for } x > 0. \tag{6}$$

If $\alpha > (<) 1$ the Weibull has an IFR (DFR) distribution, and for $\alpha = 1$, it reduces to the exponential: see Barlow and Proschan (1975). This pdf does not belong to the exponential family (2); however, after the reparametrization $\beta = \theta^\alpha$, the Weibull pdf becomes

$$h_\beta(x) = \alpha\beta x^{\alpha-1} e^{-\beta x^\alpha} \text{ for } x > 0;$$

which does belong to (2). In general, if the pdf of a random variable Z is written $p_\theta(z) = q_\beta(z)$, where $\beta = \beta(\theta)$, the Fisher information transforms thus (see Lehmann, 1983):

$$I_Z(\theta) = I_Z^*(\beta(\theta))\left(\frac{d\beta}{d\theta}\right)^2; \tag{7}$$

where $I_Z(\theta)$ and $I_Z^*(\beta)$ denote the Fisher information about θ and β, respectively, based on Z. Comparisons of the Fisher information about θ between two experiments can thus be recast as comparisons of the Fisher information about β because the term $(d\beta/d\theta)^2$

cancels out. For the Weibull case, $I^*_X(\beta) = 1/\beta^2$, and $I_X(\theta) = \alpha^2/\theta^2$. Now, the pdf of the k^{th} order statistic $Y = X_{(k)}$ from a sample of size m is

$$g_\beta(y) = k\begin{pmatrix} m \\ k \end{pmatrix}\alpha\beta y^{\alpha-1}(1-e^{-\beta y^\alpha})^{k-1}e^{-(m-k+1)\beta y^\alpha} \text{ for } y > 0.$$

The Fisher information about β in $X_{(k)}$ is

$$I^*_{X_{(k)}}(\beta) = E_\beta\left[-\frac{d^2\log g_\beta\{X_{(k)}\}}{d\beta^2}\right] = \frac{1}{\beta^2} - (k-1)E_\beta\left\{\frac{d}{d\beta}\frac{X^\alpha_{(k)}}{e^{\beta X^\alpha_{(k)}}-1}\right\}. \qquad (8)$$

Since $1/(e^{\beta x^\alpha}-1)$ is a decreasing function of β for all $x > 0$, Theorem 3.1 together with (7) and (8) yields the following results for the Weibull with known shape parameter α:

(i) For $k = 1$, $I_{X_{(1)}}(\theta) = I_X(\theta)$ for all θ and α.

(ii) For $1 < k \le m$, $I_{X_{(k)}}(\theta) > I_X(\theta)$ for all θ and α.

In contrast to the gamma distribution, the first order statistic from a sample from a Weibull distribution provides the same Fisher information about the scale parameter as does a single observation, regardless of the value of α.

In this Section, we have not addressed the problem of computing how much more (or less) Fisher information is contained in certain order statistics; this knowledge would be useful for planning purposes, for example to pick out the most informative order statistic when a choice is available. Such a study will involve a numerical investigation of $I_{X_{(k)}}(\theta)$ for specific models, such as the related work of Park (1996), who derived recurrence relations for the Fisher information in sets of consecutive order statistics.

4. Residual Lifetime of a Stationary Renewal Process

Consider a renewal process where a component with lifetime $X \ge 0$ is replaced upon failure by an identical new component. For successive lifetimes $X_1, X_2, \ldots$, let $N(t)$ denote the number of renewals up to time t, and let

$$R_t = X_1 + \cdots + X_{N(t)+1} - t$$

denote the residual lifetime of the component that is working at time t. The limiting distribution of R_t as $t \to \infty$ exists when X has a finite mean. If X has pdf $f(x)$, cdf $F(x)$, survival function $\bar{F}(x) = 1 - F(x)$, and mean μ, then this stationary distribution has pdf

$$g(y) = \frac{\bar F(y)}{\mu} \text{ for } y > 0.$$

Since

$$g(y) = \frac{f(y)}{\mu \lambda(y)} \text{ for } y > 0,$$

where $\lambda(y)$ is the of the hazard rate function of X, $g(y)$ is a weighted version of $f(y)$ with weight function $w(y) = 1/\lambda(y)$.

In this section, we investigate the Fisher information in a sample from this stationary residual distribution about the scale parameter, θ, of the gamma and Weibull distributions. In both cases, we assume that the shape parameter α is known. In brief, our results are the following. Let X denote the component lifetime, and let Y denote the corresponding residual lifetime, with the stationary distribution. Then for both the gamma and Weibull, if $\alpha > 1$ (IFR), $I_Y(\theta) < I_X(\theta)$, and if $\alpha < 1$ (DFR), $I_Y(\theta) > I_X(\theta)$; of course, if $\alpha = 1$, both families reduce to the exponential distribution, for which $I_Y(\theta) = I_X(\theta)$ by its memorylessness property. We now prove these statements.

Let X have the gamma distribution with pdf $f_\theta(x) = \theta f_1(\theta x)$ given in (5) and cdf $F_\theta(x) = F_1(\theta x)$, so that the mean is $\mu = \alpha/\theta$, and $I_X(\theta) = \alpha/\theta^2$. Then Y has the stationary distribution with pdf

$$g_\theta(y) = \frac{\bar F_\theta(y)}{\mu} = \frac{\theta}{\alpha}\bar F_\theta(y) = \frac{\theta}{\alpha}\bar F_1(\theta y).$$

In this case,

$$-\frac{d^2}{d\theta^2}\log g_\theta(y) = \frac{1}{\theta^2} + \frac{y^2 f_1'(\theta y)}{\bar F_1(\theta y)} + \frac{y^2 f_1^2(\theta y)}{\bar F_1^2(\theta y)} = \frac{1}{\theta^2}\left\{1 + y^2\frac{f_\theta'(y)}{\bar F_\theta(y)} + \frac{y^2 f_\theta^2(y)}{\bar F_\theta^2(y)}\right\}; \qquad (9)$$

where $f_\theta'(y)$ is the derivative with respect to y. An integration by parts yields

$$E_\theta\left\{Y^2\frac{f_\theta'(Y)}{\bar F_\theta(Y)}\right\} = \frac{\theta}{\alpha}\int_0^\infty y^2 f_\theta'(y)\,dy = -\frac{2\theta}{\alpha}\int_0^\infty y f_\theta(y)\,dy = -2. \qquad (10)$$

Next,

$$\begin{aligned} E_\theta\left\{\frac{Y^2 f_\theta^2(Y)}{\bar F_\theta^2 Y}\right\} &= \frac{\theta}{\alpha}\int_0^\infty y^2 \frac{\left\{\frac{1}{\Gamma(\alpha)}\theta^\alpha y^{\alpha-1}e^{-\theta y}\right\}^2}{\left\{\frac{1}{\Gamma(\alpha)}\int_y^\infty \theta^\alpha x^{\alpha-1}e^{-\theta x}\,dx\right\}}\,dy \\ &= \frac{\theta^{\alpha+1}}{\Gamma(\alpha+1)}\int_0^\infty \frac{y^{2\alpha}e^{-2\theta y}}{\int_y^\infty x^{\alpha-1}e^{-\theta x}\,dx}\,dy. \end{aligned} \qquad (11)$$

For $\alpha > (<) 1$ and all $y > 0$,

$$\int_y^\infty x^{\alpha-1} e^{-\theta x}\,dx > (<) \frac{y^{\alpha-1} e^{-\theta y}}{\theta};$$

so that for $\alpha > (<) 1$

$$E_\theta \left\{ \frac{Y^2 f_\theta^2(Y)}{\bar{F}_\theta^2(Y)} \right\} < (>) \frac{\theta^{\alpha+2}}{\Gamma(\alpha+1)} \int_0^\infty y^{\alpha+1} e^{-\theta y}\,dy = \alpha + 1. \tag{12}$$

Thus for $\alpha > (<) 1$, (9)-(12) imply that

$$I_Y(\theta) = E_\theta \left\{ -\frac{d^2}{d\theta^2} \log g_\theta(Y) \right\} < (>) \frac{1}{\theta^2}\{1 - 2 + (\alpha+1)\} = \frac{\alpha}{\theta^2} = I_X(\theta);$$

which completes the proof for the gamma.

Next, let X have the Weibull distribution with pdf $f_\theta(x)$ given in (6), cdf $F_\theta(x)$, survival function $\bar{F}_\theta(x) = e^{-(\theta x)^\alpha}$, and $I_X(\theta) = \alpha^2/\theta^2$. Since the mean is

$$\mu = E_\theta(X) = \frac{\Gamma(\alpha^{-1} + 1)}{\theta},$$

the stationary distribution of the residual lifetime, Y, has pdf

$$g_\theta(y) = \frac{\theta}{\Gamma(\alpha^{-1} + 1)} e^{-(\theta y)^\alpha}.$$

Thus,

$$I_Y(\theta) = E_\theta \left\{ -\frac{d^2 \log g_\theta(Y)}{d\theta^2} \right\} = \frac{1}{\theta^2} + \alpha(\alpha - 1)\theta^{\alpha-2} E_\theta(Y^\alpha).$$

Since $E_\theta(Y^\alpha) = 1/\alpha\theta^\alpha$, $I_Y(\theta) = \alpha/\theta^2$, so that $I_X(\theta) = \alpha I_Y(\theta)$.

5. Selection Samples

Let X have pdf $f_\theta(x)$. Bayarri et al. (1987) considered problems in which observations are restricted, or truncated, to a subset S of the sample space, so that $w(x) = I(x \in S)$. Let Y have the truncated pdf

$$h_\theta(y) = \frac{f_\theta(y)}{P_\theta(X \in S)} \text{ for } y \in S; \tag{13}$$

and zero otherwise. A random sample from (13) is called a selection sample, and S is called a selection set. If $f_\theta(x)$ belongs to the exponential family of distributions of the form (2), then it follows from Corollary 2.2 of Bayarri et al. (1987) that $I_X(\theta) \geq I_Y(\theta)$ for all θ if and only if $\log P_\theta(X \in S)$ is a concave function of θ. They used this fact to derive information orderings between the normal location, normal scale, and gamma scale families and their weighted counterparts using various truncation sets. In this section, we present further results on selected samples from the gamma, Weibull, and exponential distributions.

First, for the gamma with known shape parameter α, unknown scale θ, and cdf $F_\theta(x)$, let $S = \{x : x \leq a\}$ for some $a > 0$. Then

$$\frac{d}{d\theta} \log P_\theta(X \in S) = \frac{d}{d\theta} \log F_\theta(a) = a\frac{f_1(a\theta)}{F_1(a\theta)} = a\mu_1(a\theta);$$

where μ_1 is the reversed hazard rate function of F_1. Since the gamma distribution is DRHR for all values of α, a selection sample from the lower tail of the gamma distribution has smaller Fisher information about the scale parameter than an unrestricted sample, for all values of α.

Next we consider the Weibull distribution. In Section 3, we noted that it belongs to the exponential family (2) only after reparametrization. Let $f_\theta(x)$ be given by (6), with $\alpha > 0$ known, and θ unknown. Let $S = \{x : x < a\}$ for $a > 0$, and let Y be an observation selected from this lower tail. Since $P_\theta(X \in S) = 1 - e^{-(\theta a)^\alpha}$,

$$g_\theta(y) = \frac{\theta^\alpha y^{\alpha-1} e^{-(\theta y)^\alpha}}{1 - e^{-(\theta a)^\alpha}} \text{ for } 0 < y < a.$$

After some calculations we have

$$I_Y(\theta) = \frac{\alpha^2}{\theta^2} - \left(\frac{\alpha a^\alpha \theta^{\alpha-1} e^{-(\theta a)^\alpha}}{1 - e^{-(\theta a)^\alpha}}\right)^2 = I_X(\theta) - \left(\frac{\alpha a^\alpha \theta^{\alpha-1} e^{-(\theta a)^\alpha}}{1 - e^{-(\theta a)^\alpha}}\right)^2;$$

so that $I_X(\theta) > I_Y(\theta)$ for all $\theta > 0$, for all α. However, if Y is an observation selected from the upper tail, $S = \{x : x > b\}$, of the Weibull distribution, its pdf is

$$g_\theta(y) = \alpha\theta^\alpha y^{\alpha-1} e^{-\{(\theta y)^\alpha - (\theta b)^\alpha\}} \text{ for } y > b.$$

In this case, $I_X(\theta) = I_Y(\theta)$ for all θ and all α. Thus, the gamma and Weibull differ in this respect, even though both of them are IFR for $\alpha > 1$ and DFR for $\alpha < 1$.

Bayarri et al. (1987) proved that an observation selected from both tails of the normal distribution with known variance has more information about the unknown mean θ for some values of θ than an observation from that normal distribution. For other values of

μ, it gives less information. We now prove an analogous result for the scale parameter of the exponential.

Theorem 5.1. Let X have an exponential distribution with pdf $f_\mu(x) = \mu e^{-\mu x}$ for $x > 0$. For $0 < a < b < \infty$, let $w(x) = I(x < a \text{ or } x > b)$, and let Y have the corresponding weighted distribution. Then

$$I_Y(\mu) > (<) I_X(\mu) \text{ for } 0 < \mu < \mu^* \; (\mu > \mu^*);$$

where μ^* is the unique positive solution of the equation

$$b^2 e^{\mu a} - a^2 e^{\mu b} = (b - a)^2.$$

Proof. Y has pdf

$$g_\mu(y) = \frac{\mu e^{-\mu y}}{1 - e^{-\mu a} + e^{-\mu b}} \text{ for } 0 < y < a \text{ or } y > b.$$

Thus,

$$I_Y(\mu) = E_\mu\left\{-\frac{d^2 \log g_\mu(Y)}{d\mu^2}\right\} = \frac{1}{\mu^2} + \frac{b^2 e^{\mu a} - a^2 e^{\mu b} - (b-a)^2}{e^{\mu(a+b)}(1 - e^{-\mu a} + e^{-\mu b})^2}.$$

Since $I_X(\mu) = 1/\mu^2$, $I_Y(\mu) > I_X(\mu)$ if and only if $h(\mu) = b^2 e^{\mu a} - a^2 e^{\mu b} - (b-a)^2$ is positive. Note that $h(0+) = 2a(b - a) > 0$ so that if $c = \log(b/a)/(b - a)$, h is strictly increasing in the interval $(0, c)$, and strictly decreasing in the interval (c, ∞); furthermore, h is concave and eventually negative. Thus, h is positive only when $0 < \mu < \mu^*$, establishing the result.

6. Acknowledgment

We thank two reviewers for their careful reading and helpful comments about this work.

7. References

Anderson, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). Statistical Models Based on Counting Processes. Springer-Verlag, New York.

Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing. Rineholt, Hart, and Winston, New York.

Bayarri, M. J. and DeGroot, M. H. (1987a). Bayesian Analysis of Selection Models. The Statistician, 36, 137-146.

Bayarri, M. J. and DeGroot, M. H. (1987b). Information in Selection Models. Probability and Bayesian Statistics, edited by R. Viertl. Plenum Press, New York, 39-51.

Bayarri, M. J., DeGroot, M. H., and Goel, P. K. (1987). Truncation, Information, and the Coe±cient of Variation. Contributions to Probability and Statistics, edited by L. J. Gleser, M. D. Perlman, S. J. Press, and A. R. Sampson. Springer-Verlag, New York, 412-428.

Block, H. W., Savits, T. H., and Singh, H. (1998). On the Reversed Hazard Rate Function. Probab. Engrg. Inform. Sci., 12, 69-90.

David, H. A. (1970). Order Statistics. Wiley, New York.

Fisher, R. A. (1934). The E®ects of Methods of Ascertainment upon the Estimation of Frequencies. Ann. Eugenics, 6, 13-25.

Hedges, L. V. (1992). Modeling Publication Selection E®ects in Meta-Analysis. Statist. Sci., 7, 246-255.

Iyengar, S. (1986). On a Lower Bound for the Multivariate Normal Mills' Ratio. Ann. Probab., 14, 1399-1403.

Iyengar, S. and Greenhouse, J. B. (1988). Selection Models and the File-Drawer Problem. Statist. Sci., 3, 109-135.

Kagan, A. M., Linnik, Yu. V., and Rao, C. R. (1973) Characterization Problems in Mathematical Statistics. Wiley, New York.

Lehmann, E. L. (1983). Theory of Point Estimation. Wiley, New York.

Little, R. J. A., and Rubin, D. B. (1987) Statistical Analysis with Missing Data. Wiley, New York.

Park, S. (1996). Fisher Information in Order Statistics. J. Amer. Statist. Assoc., 91, 385-389.

Patil, G. P. and Taillie, C. (1987). Weighted Distributions and the E®ects of Weight Functions on Fisher Information. Technical Report, Centre for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University.

Rao, C. R. (1965). On Discrete Distributions Arising out of Methods of Ascertainment. Classical and Contagious Discrete Distributions, edited by G. P. Patil. Pergamon Press and Statistical Publishing Society, Calcutta, 320-332.

Rao, C. R. (1985). Weighted Distributions Arising Out of Methods of Ascertainment: What Population Does a Sample Represent? A Celebration of Statistics: The ISI Centenary Volume, edited by A. C. Atkinson and S. E. Fienberg. Springer-Verlag, New York, 543-569.

Shaked, M. and Shantikumar, J. G. (1994). Stochastic Orders and their Applications. Academic Press, New York.