STRUCTURAL INFERENCE ON THE GENERALIZED GAMMA DISTRIBUTION BASED ON TYPE-II PROGRESSIVELY CENSORED SAMPLE

M. S. MASWADAH

Communicated by Timothy C. Brown

Abstract

This paper is concerned with the three-parameter generalized gamma distribution (g.g.d.) which is widely employed as a model in life testing. The structural probability distributions of the parameters and a number of structural prediction densities for specific future measurements have been derived based on type-II progressively censored sample.

1980 *Mathematics subject classification (Amer. Math. Soc.)* (1985 *Revision*): primary 62 F 25; secondary 62 F 15, 62 P 20.

1. Introduction

In many cases in life tests, the experiment may be discontinued after the first r failures have been observed from a sample of n items. If experimentation is discontinued at a pre-determined time so that the number r is a random variable, the data obtained are said to be time censored, or type-I censored. Unfortunately, in this case structural distributions of parameters can not be obtained. This has been pointed out by Whitney and Minder [9]. If experimentation is discontinued at a pre-determined number, r, of observations, the data obtained are said to be type-II censored. In many life testing situations, the initial censoring results in withdrawal of only a portion of the survivors, with some remaining on test and therefore continuing

under observation until ultimate failure or until a subsequent stage of censoring is performed. For sufficiently large samples, censoring may be progressive through several stages.

Suppose a random sample of n items is placed under observation at time $T = 0$. At time T_1 when the first failure occurs, K_1 of the survivors are censored and the sample is reduced to $n_2 = n - K_1 - 1$. At time T_2 when the second failure occurs, k_2 of the survivors are censored, and the sample of survivors is again reduced to size $n_3 = n_2 - K_2 - 1$, etc. Suppose that censoring occurs progressively in m stages and at time T_j, where $j = 1, 2, 3, \ldots, m$. In this situation, at the time T_j, K_j survivors are randomly censored from the sample. If r items actually fail from the sample, then the size of the initial sample is $n = r + \sum_{j=1}^{m} K_j$.

The purpose of this paper is to apply the structural method to derive a posterior distribution of the parameters and predictive densities for specific future measurements based on type-II progressively censored sample from a three-parameter g.g. population whose probability density function is given by

$$(1.1) \quad g(x; \alpha, \beta, k) = \begin{cases} \dfrac{1}{\Gamma(k)} \left(\dfrac{\beta}{\alpha}\right) \left(\dfrac{x}{\alpha}\right)^{\beta k - 1} \cdot \exp\left(-\left(\dfrac{x}{\alpha}\right)^{\beta}\right), & x > 0, \\ 0, & x \leq 0, \end{cases}$$

where $\alpha > 0$, $\beta > 0$, $k > 0$. The corresponding cumulative distribution function (c.d.f.) is given by

$$(1.2) \qquad\qquad G(x; \alpha, \beta, k) = I[k, (x/\alpha)^{\beta}],$$

where $I[k, x] = (1/\Gamma(k)) \int_0^x u^{k-1} e^{-u} \, du$ is the incomplete gamma function.

It is clear that this model includes many well-known distributions used in life time as special cases: Weibull distribution $(k = 1)$, gamma distribution $(\beta = 1)$, exponential distribution $(\beta = k = 1)$ and log-normal distribution as $k \to \infty$, where α is the scale parameter, k is the index parameter and β is the power parameter. Consequently, it has been suggested for use as a general life-testing model.

2. Derivation of the structural distribution
of the g.g. parameters

The structural probability distribution of the three-parameter g.g.d. is derived directly from considerations of the group structural of its density function (1.1) by means of invariant differentials described in Fraser's book [5, page 53] when k is assumed known, where good inference procedures

are difficult to obtain with k assumed unknown. One can easily show that the model (1.1) is related to an error model

$$(2.1) \qquad f(e; k) = \begin{cases} \dfrac{1}{\Gamma(k)} e^{k-1} \exp(-e), & e > 0, \\ 0, & e \leq 0, \end{cases}$$

by the equation $X = \alpha e^{1/\beta}$, $0 < \alpha$, $\beta < \infty$, which can be written as $X = [\alpha, \beta]e$. Here $[\alpha, \beta]$ is a member of a group G of transformations of the sample space of $X = (x_1, x_2, \ldots, x_n)$ or of $e = (e_1, e_2, \ldots, e_n)$ onto itself.

The group $G = \{[a, b], 0 < a, b < \infty\}$ is defined by the operations $[a, b][c, d] = [ac^{1/b}, bd]$, $[a, b]^{-1} = [a^{-b}, b^{-1}]$ and $[1, 1]$ is the identity of G. Thus the structural probability density function (p.d.f.) of the parameters given a complete sample $(x_1, x_2, \ldots, x_n)$ is given by
(2.2)

$$g^*(\alpha, \beta | X, k) = C[\Gamma(k)]^{-n} \beta^{n-1} \alpha^{-nk\beta-1} \prod_{i=1}^{n} x_i^{\beta k} \cdot \exp\left(-\sum_{i=1}^{n} \left(\frac{x_i}{\alpha}\right)^{\beta}\right).$$

Equation (2.2) has been derived in Maswadah [6]. The approach in [6] has been followed by Tan [7] for Weibull distribution, and by Bury and Bernholtz for different life-time models which are special cases of the g.g. model (see Bury and Bernholtz [2, 3, 4]).

Given a set of type-II progressively censored data from the g.g. population (1.1) specified by

$$X = (x_1 < x_2 < \cdots < x_r), \qquad 1 \leq r < n,$$

we can derive the structural p.d.f. of the parameters from the modified structural model as follows.

Let $e = (e_1 < e_2 < \cdots < e_r)$ have the p.d.f.

$$(2.3) \qquad h(e; k) = \prod_{i=1}^{r} n_i' f(e_i; k)[1 - F(e_i; k)]^{K_i},$$

where $n_i' = n - \sum_{j=1}^{i-1} K_j - i + 1$, $f(e_i; k)$ is given by (2.1), and $F(e_i; k) = I(k; e_i)$ is the c.d.f. of $f(e_i; k)$. The structural model in this case has two parts: an error distribution (2.3) which describes the multiple operation of the internal sources of variation in the experiment, and the structural equation

$$X = [\alpha, \beta]e = ([\alpha, \beta]e_1, \ldots, [\alpha, \beta]e_r),$$

in which a realized vector e from the error distribution provides the link between the observation vector $X = (x_1, x_2, \ldots, x_r)$ and the physical quantities α and β.

Now a direct application of Fraser's structural method, using the invariant differential as described in Maswadah [6] yields the following structural p.d.f. of t_1 and t_2 given D and k:

$$(2.4) \qquad g(t_1, t_2|D; k) = C \frac{J_N(t_1, t_2)}{J_M^*([t_1, t_2])} \prod_{i=1}^r h([t_1, t_2]d_i, k),$$

where

$$[t_1(\mathbf{e}), t_2(\mathbf{e})] = [\alpha, \beta]^{-1}[t_1(X), t_2(X)] = \left[\left(\frac{t_1(X)}{\alpha}\right)^\beta, \frac{t_2(X)}{\beta}\right]$$

is the position variable, $D = D(X) = (d_1(X), d_2(X), \ldots, d_r(X))$ is the reference point which index the orbit GX, $J_N(\Theta)$ is the Jacobian $|\partial[\Theta]Y/\partial Y|_{Y=D}$ $= t_1^r t_2^{-r} \prod_{i=1}^r d_i^{1/t_2-1}$, $J_M^*([\Theta])$ is the Jacobian $|\partial[\Theta]g/\partial g|_{g=i} = t_1$, where i is the identity of G, $[\Theta] \in G$ and $h(\cdot)$ is the p.d.f. as defined in (2.3).

From (2.4) we can transform the variables from (t_1, t_2) into (α, β) and using the values of the above Jacobians we obtain the conditional structural probability density for α and β as follows:

(2.5)

$$g^*(\alpha, \beta|X; k) = C_1[\Gamma(k)]^{-r}\beta^{r-1}\alpha^{-rk\beta-1}\prod_{i=1}^r x_i^{\beta k}$$

$$\cdot \exp\left(-\sum_{i=1}^r \left(\frac{x_i}{\alpha}\right)^\beta\right)\left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^\beta\right]\right)^{\sum_{j=1}^m K_j},$$

where C_1 is the normalizing constant and is defined by

$$(2.6) \; C_1^{-1} = [\Gamma(k)]^{-r} \int_{\alpha=0}^\infty \int_{\beta=0}^\infty \beta^{r-1}\alpha^{-rk\beta-1}\prod_{i=1}^r x_i^{\beta k}$$

$$\cdot \exp\left(-\sum_{i=1}^r \left(\frac{x_i}{\alpha}\right)^\beta\right)\left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^\beta\right]\right)^{\sum_{j=1}^m K_j} d\beta\, d\alpha.$$

Hence for censored samples, (2.5) is the basis for inference, just as (2.2) is the basis for inference in the case of complete samples.

In fact, a simple method of obtaining (2.5) by using Bayes concepts has been discussed by Bury and Bernholtz [2] for deriving the corresponding structural p.d.f. in case of Weibull distribution, which is a special case from (2.5) at $k = 1$. They used as prior distribution for α and β the structural p.d.f. of α and β given $X = (x_1 < x_2 < \cdots < x_r)$, and k as a complete ordered sample of size r which is given by (2.2) when $n = r$, and combined it with the likelihood modulation of K_j withdrawals at time

T_j $(j = 1, 2, \ldots, m)$. Thus an application of Bayes theorem yields the posterior p.d.f. of α and β which is defined by (2.5).

Another simple method of obtaining (2.5) using the right invariant differential which is defined in Fraser's book [5, page 64]. This method yields

$$(2.7) \qquad g^*(\alpha, \beta | X, k) \, d\alpha \, d\beta = C g(x; \alpha, \beta, k) \, dv([\alpha, \beta]),$$

where the right invariant differential under the group G for the g.g.d. is

$$(2.8) \qquad dv([\alpha, \beta]) = d\mu([\alpha, \beta]^{-1}) = \alpha^{-1} \beta^{-1} \, d\alpha \, d\beta,$$

and

$$(2.9) \qquad g(x; \alpha, \beta, k) = \prod_{i=1}^{r} n'_i g(x_i; \alpha, \beta, k)[1 - G(x_i; \alpha, \beta, k)]^{K_i},$$

where $n'_i = n - \sum_{j=1}^{i-1} K_j - i + 1$. Substituting (2.8) and (2.9) in (2.7) yields immediately (2.5).

From (2.5), the marginal structural densities of α and β can be obtained to construct conditional confidence intervals for α and β respectively, for any desired probability level $0 < \gamma < 1$, given X and k.

3. Structural prediction densities

A number of structural prediction densities have been derived, based on N independent future measurements Z_s from a three-parameter g.g. population with type-II progressively censored data.

(i) Consider the problem of predicting an arbitrary vector $Z = (z_1, \ldots, z_2, \ldots, z_s, \ldots, z_N)$ of N future measurements z_s. The marginal density of Z is found as

$$g_1(z; \alpha, \beta, k, N) = \prod_{s=1}^{N} g(z_s; \alpha, \beta, k, N)$$

$$= [\Gamma(k)]^{-N} \beta^N \alpha^{-kN\beta} \prod_{s=1}^{N} (z_s)^{k\beta-1} \exp\left(-\sum_{s=1}^{N} \left(\frac{z_s}{\alpha}\right)^{\beta}\right).$$

We can now obtain the conditional structural density of the prediction vector Z by taking the expectation of $g_1(z; \alpha, \beta, k, N)$ with respect to the structural distribution of (α, β) given X and k which is given in (2.5). Thus

we get

(3.1)

$$\phi_1(z|X, k, N) = \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} g_1(z; \alpha, \beta, k, N) g^*(\alpha, \beta|X, k) \, d\beta \, d\alpha$$

$$= C_1[\Gamma(k)]^{-(N+r)} \int_{\beta=0}^{\infty} \int_{\alpha=0}^{\infty} \beta^{N+r-1} \prod_{i=1}^{r} x_i^{k\beta}$$

$$\cdot \prod_{s=1}^{N} (z_s')^{k\beta-1} \exp\left[-\alpha^{-\beta}\left(\sum_{s=1}^{N} z_s^{\beta} + \sum_{i=1}^{r} x_i^{\beta}\right)\right]$$

$$\cdot \alpha^{-rk\beta-N\beta k-1} \left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^{\beta}\right]\right)^{\sum_{j=1}^{m} K_j} d\beta \, d\alpha.$$

For a single future observation Z, its prediction density can be obtained from (3.1) when $N = 1$.

(ii) Consider the problem of predicting the smallest future observation $Z_{(1)}$ of the N future measurements Z_s. From order statistics, the density function of $Z_{(1)}$ is given by

$$f_{Z_{(1)}}(z; \alpha, \beta, k, N) = N[1 - G(z; \alpha, \beta, k)]^{N-1} g(z; \alpha, \beta),$$

where $G(z; \alpha, \beta, k)$ is the corresponding distribution function of the three parameter g.g.d. Thus the structural prediction density of $Z_{(1)}$ is given by

(3.2)

$$\phi_{Z_{(1)}}(z|X, k, N) = \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} f_{Z_{(1)}}(z; \alpha, \beta, k, N) g^*(\alpha, \beta|X, k) \, d\beta \, d\alpha$$

$$= C_1 N[\Gamma(k)]^{-(r+1)} \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} \prod_{i=1}^{r} (X_i)^{k\beta} \beta^r z^{k\beta-1}$$

$$\cdot \left(1 - I\left[k, \left(\frac{z}{\alpha}\right)^{\beta}\right]\right)^{N-1} \exp\left(-\left(\frac{z}{\alpha}\right)^{\beta} - \sum_{i=1}^{r}\left(\frac{x_i}{\alpha}\right)^{\beta}\right)$$

$$\cdot \alpha^{-rk\beta-\beta k-1} \left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^{\beta}\right]\right)^{\sum_{j=1}^{m} K_j} d\beta \, d\alpha.$$

Similarly we can find the structural prediction density of $Z_{(N)}$, the largest

of N future measurements Z_s, as

(3.3) $\phi_{Z_{(N)}}(z|X, k, N) = C_1 N[\Gamma(k)]^{-(r+1)}$

$$\cdot \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} \prod_{i=1}^{r} x_i^{k\beta} \beta^r z^{k\beta-1} \alpha^{-\beta k(r+1)-1}$$

$$\cdot \left(I\left[k, \left(\frac{z}{\alpha}\right)^{\beta}\right]\right)^{N-1} \exp\left(-\left(\frac{z}{\alpha}\right)^{\beta} - \sum_{i=1}^{r}\left(\frac{x_i}{\alpha}\right)^{\beta}\right)$$

$$\cdot \left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^{\beta}\right]\right)^{\sum_{j=1}^{m} K_j} d\beta \, d\alpha.$$

(iii) Consider now the problem of predicting the average $\overline{Z}$ of N future observations Z_s from the model (1.1). By the addition theorem of gamma variables, with common parameters, $Y = \sum_{s=1}^{N} Z_s$ is distributed as a gamma variable with parameters α, β and kN. Thus the p.d.f. of the mean $\overline{Z} = Y/N$ is given by

$$q(\overline{z}; \alpha, \beta, k, N) = \frac{N^{Nk\beta}}{\Gamma(Nk)} \left(\frac{\beta}{\alpha}\right) \left(\frac{\overline{z}}{\alpha}\right)^{Nk\beta-1} \exp\left[-\left(\frac{N\overline{z}}{\alpha}\right)^{\beta}\right].$$

As above, the structural prediction density of $\overline{Z}$ is given by
(3.4)

$$\phi_{\overline{Z}}(\overline{z}|X, k, N) = \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} q(\overline{z}; \alpha, \beta, k, N) g^*(\alpha, \beta|X, k) d\beta \, d\alpha$$

$$= C_1 N[\Gamma(k)]^{-r}[\Gamma(kN)]^{-1} \int_{\alpha=0}^{\infty} \int_{\beta=0}^{\infty} \beta^r$$

$$\cdot \prod_{i=1}^{r} x_i^{k\beta} (N\overline{z})^{Nk\beta-1} \exp\left(-\left(\frac{N\overline{z}}{\alpha}\right)^{\beta} - \sum_{i=1}^{r}\left(\frac{x_i}{\alpha}\right)^{\beta}\right)$$

$$\cdot \alpha^{-\beta k(N+r)-1} \left(1 - I\left[k, \left(\frac{T_j}{\alpha}\right)^{\beta}\right]\right)^{\sum_{j=1}^{m} K_j} d\beta \, d\alpha,$$

where C_1 in (3.1) to (3.4) is the normalizing constant which is given in (2.6). It is seen that these densities in (3.1) to (3.4) are similar to those corresponding for a complete sample which can be deduce from Maswadah [6] by following the same procedures, so that censored and complete samples can be analysed by essentially the same numerical procedure by using tables of the incomplete gamma function $I(k, x)$.

Finally, it is noteworthy that the corresponding densities derived by Bury and Bernholtz [2] for the Weibull distribution are special cases of the densities (3.1) to (3.3) when $k = 1$, as might be expected from the nature of the g.g.d.

4. Conclusion

Structural inference theory provides structural probability densities for parameters of the model, which permits us to determine prediction densities for some future measurements without making any prior assumption such as in Bayesian inference. Thus the structural inference is a consistent mathematical framework for statistical inference.

Acknowledgements

References

[1] B. Bernholtz, 'Type-I censoring and the structural approach', *Statist. Hefte* **17** (1971), 2–12.
[2] K. V. Bury and B. Bernholtz, 'On structural inference applied to the Weibull distribution', *Statist. Hefte* **12** (1971), 177–192.
[3] K. V. Bury and B. Bernholtz, 'Structural inference for a class of stochastic process', *INFOR—Canad. J. Oper. Res. Inform. Process.* **9**, No. I (1970), 23–31.
[4] K. V. Bury and B. Bernholtz, 'Life testing: Structural inference on the exponential model', *INFORM—Canad. J. Oper. Res. Inform. Process.* **9**, No. 2 (1971), 148–160.
[5] D. A. S. Fraser, *The structural of inference*, (John Wiley & Sons, New York, 1968).
[6] M. S. Maswadh, 'Two approaches based on the structural model to inference on the generalized gamma parameters', *Egyptian Statist. J.*, to appear.
[7] P. Tan, 'Four approaches to inference on Weibull parameters', *Utilitas Math.* **6** (1974), 111–120.
[8] P. Tan and B. Bernholtz, 'Applications of structural models in statistics', *Math. Sci.* **5** (1980), 31–44.
[9] J. B. Whitney and Ch. E. Minder, 'Time censoring and the structural model', *Statist. Hefte* **15** (1974), 27–35.

Department of Mathematics
Faculty of Science
Aswan
Egypt