

Parameter Estimation of Burr Type X Distribution for Grouped Data

K. M. Aludaat¹, M. T. Alodat² and T. T. Alodat³

^{1,2,3} Department of Statistics, Yarmouk University, Irbid, Jordan
aludaatkm@hotmail.com¹ and malodat@yu.edu.jo²

Abstract

In this paper, we obtain Bayesian and non-Bayesian estimators for the parameter of the Burr type X distribution under grouped data. We apply our results to artificial data.

Keywords: Burr type X; Bayesian estimation; Grouped data; Posterior distribution

1. Introduction

In various fields of science such as biology, engineering and medicine it is not possible to obtain the measurements of a statistical experiment exactly, but is possible to classify them into intervals, rectangles or disjoint subsets (Alodat and Al-Saleh, 2002; Heitjan, 1989; Surles and Padgett, 2001; Wu and Perloff, 2005; Phipper and Ritz, 2006). For example, in life testing experiments, we observe the failure time of a component to the nearest hour, day or month. Data for which true values are known only up to subsets of the sample space are called grouped data. In general grouped data can be formulated as follows: Let X_1, \dots, X_n be a random sample from the density $f(x; \theta)$, $x \in \chi$, $\theta \in \Theta$ and $\chi_1, \dots, \chi_{k+1}$ be a partition of the sample space χ and $N_j =$ the number of X_j 's that fall in χ_j for $j = 1, 2, \dots, k+1$. The set of pairs $\{(\chi_1, N_1), \dots, (\chi_{k+1}, N_{k+1})\}$ is called grouped data. The grouped data problem is to use these data to draw inferences about the parameter θ . Since we don't have complete information about the sample, then there will be a loss in the information due to the grouping. Schervish (1995), page 114, shows the following

$$I_{\underline{X}}(\theta) = I_{\underline{Y}}(\theta) + E_{\theta} \{I_{\underline{X}|\underline{Y}}(\theta | \underline{Y})\},$$

where $I_{\underline{X}}(\theta)$ and $I_{\underline{Y}}(\theta)$ are the Fisher's information numbers obtained from \underline{X} and \underline{Y} , respectively, and $E_{\theta}\{I_{\underline{X}|\underline{Y}}(\theta|\underline{Y})\}$ is the conditional score function. If we replace \underline{Y} by the grouped sample $\underline{n} = (N_1, \dots, N_{k+1})$, then $I_{\underline{X}}(\theta) \geq I_{\underline{Y}}(\theta)$ for all θ , which means that the information in the sample \underline{X} about θ is reduced to $I_{\underline{n}}(\theta)$ because of grouping. Kuldorff (1961) considered non-Bayesian estimation from grouped data when the data come from normal and exponential distributions. Alodat and Al-Saleh (2000) considered the Bayesian estimation from grouped data when the underlying distribution is exponential. Alodat et al. (2007) obtained Bayesian prediction intervals from grouped data when the underlying distribution is exponential. In this paper, we consider the group data problem when the density $f(x; \theta)$ is Burr type X, i.e.,

$$f(x; \theta) = 2\theta x \exp(-x^2) (1 - \exp(-x^2))^{\theta-1}, \quad x > 0; \theta > 0. \quad (1)$$

The corresponding distribution function is

$$F(x; \theta) = \begin{cases} 0 & x < 0, \\ (1 - \exp(-x^2))^{\theta} & x \geq 0. \end{cases} \quad (2)$$

The Burr type X is a member of a system of distributions introduced by Burr (1941). It is found that this distribution fits many practical data Raqab (2006). Figure 1 shows the shape of $f(x; \theta)$ for different values of θ .

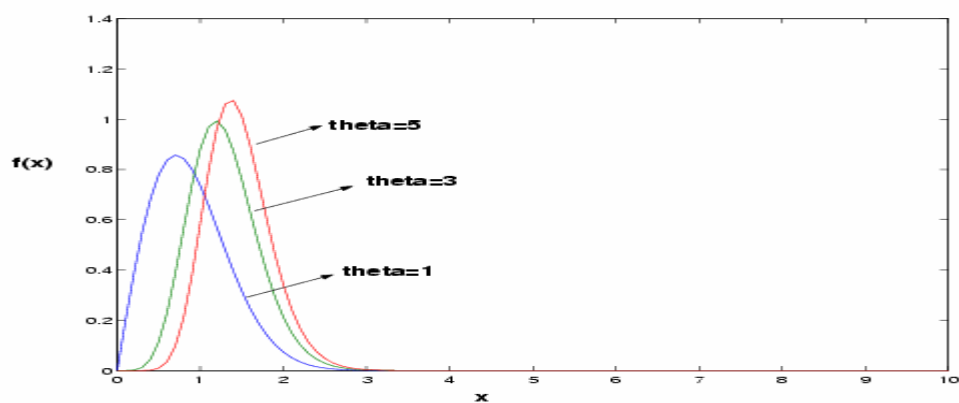


Figure 1: pdf of Burr(θ) for different values of θ .

This paper is organized as follows. In Section 2, we derive the likelihood density for the grouped data. In Section 3, we find the MLE (Maximum Likelihood Estimator) for θ from grouped data. In Section 4, we derive the Fisher's information number based on the grouped data and we use it to construct a large sample confidence interval for θ . In Section 5, we obtain a Bayes estimator for θ using grouped data. In Section 6, we discuss the estimation of the hyperparameters. In Section 7, we obtain a HPDCR (High posterior Density Credible region) for θ . In Section 8, we conducted a simulation study to compare the estimators. Our conclusion is stated in Section 9.

2. Likelihood of grouped data

Let X_1, \dots, X_n be a random sample from $\text{Burr}(\theta)$. Assume that the sample space of $f(x; \theta)$ is partitioned into $k + 1$ equally-spaced intervals as follows. Let $I_j = ((j - 1)\delta, j\delta)$, $j = 1, \dots, k$ and $I_{k+1} = [k\delta, \infty)$, $\delta > 0$. If N_j denotes the number of X_j 's that fall in I_j , $j = 1, 2, \dots, k + 1$, then $n = N_1 + \dots + N_{k+1}$. Let

$$P_j = P_j(\theta) = P(X \in I_j) = P((j - 1)\delta < X < j\delta),$$

$$= (1 - \exp(-j^2 \delta^2))^\theta - (1 - \exp(-(j - 1)^2 \delta^2))^\theta,$$

for $j = 1, \dots, k$ and $P_{k+1} = P_{k+1}(\theta) = P(X > k\delta) = 1 - (1 - \exp(-k^2 \delta^2))^\theta$. If we let $A_j = \log(1 - \exp(-(j - 1)^2 \delta^2))$, then $P_j = \exp(\theta A_{j+1}) - \exp(\theta A_j)$, for $j = 1, 2, \dots, k$ and $P_{k+1} = 1 - \exp(\theta A_{k+1})$. So the density of $\underline{n} = (N_1, \dots, N_{k+1})$ is given by the multinomial distribution as follows:

$$f(\underline{n}; \theta) = \frac{n!}{n_1! \dots n_{k+1}!} P_1^{n_1} \dots P_{k+1}^{n_{k+1}},$$

$$= C [1 - \exp(\theta A_{k+1})]^{n_{k+1}} \prod_{j=1}^k [\exp(\theta A_{j+1}) - \exp(\theta A_j)]^{n_j}, \quad (3)$$

where C is a normalizing constant.

3. MLE from grouped data

In this section we find the MLE of θ based on the density (3). To do this, we maximize the log-likelihood function

$$\log f(\underline{n}; \theta) = \text{const} + \sum_{j=1}^k n_j \log(\exp(\theta A_{j+1}) - \exp(\theta A_j)) + n_{k+1} \log(1 - \exp(\theta A_{k+1})).$$

The first derivative of the log-likelihood is

$$\frac{\partial \log f(\underline{n}; \theta)}{\partial \theta} = \sum_{j=1}^k n_j \frac{A_{j+1} \exp(\theta A_{j+1}) - A_j \exp(\theta A_j)}{\exp(\theta A_{j+1}) - \exp(\theta A_j)} - n_{k+1} \frac{A_{k+1} \exp(\theta A_{k+1})}{1 - \exp(\theta A_{k+1})}.$$

The MLE for θ is the solution of $\partial \log f(\underline{n}; \theta) / \partial \theta = 0$. so the MLE is $\hat{\theta}$ such that

$$\sum_{j=1}^k n_j \frac{A_{j+1} \exp(\hat{\theta} A_{j+1}) - A_j \exp(\hat{\theta} A_j)}{\exp(\hat{\theta} A_{j+1}) - \exp(\hat{\theta} A_j)} = n_{k+1} \frac{A_{k+1} \exp(\hat{\theta} A_{k+1})}{1 - \exp(\hat{\theta} A_{k+1})}.$$

We use the notation $\hat{\theta}_g$ to denote the MLE of θ obtained from the grouped data. Numerical calculations are required to find the solution of the last equation. It is easy to show that the MLE of θ based on the un-grouped sample X_1, \dots, X_n is

$$\hat{\theta} = \frac{-n}{\sum_{i=1}^n \log(1 - \exp(-X_i^2))}.$$

4. Fisher's information number

To find the Fisher's information number contained in the grouped sample about θ , we find the expectation of the second derivative of the log-likelihood. So

$$\frac{\partial^2 \log f(\underline{n} | \theta)}{\partial \theta^2} = -\sum_{j=1}^k N_j \Psi_{1j}(\theta) - N_{k+1} \Psi_2(\theta),$$

where

$$\Psi_{1j}(\theta) = \frac{(A_{j+1} - A_j)^2 \exp(\theta(A_j + A_{j+1}))}{[\exp(\theta A_{j+1}) - \exp(\theta A_j)]^2}$$

and

$$\Psi_2(\theta) = \frac{A_{k+1}^2 \exp(\theta A_{k+1})}{[1 - \exp(\theta A_{k+1})]^2}.$$

If $I_g(\theta)$ denotes the Fisher's information number from grouped data, then

$$I_g(\theta) = -E \left\{ \frac{\partial^2 \log f(\underline{n}; \theta)}{\partial \theta^2} \right\}.$$

Since $E\{N_j\} = nP_j$, then

$$\begin{aligned} I_g(\theta) &= E \left\{ \sum_{j=1}^k N_j \Psi_{1j}(\theta) + N_{k+1} \Psi_2(\theta) \right\}, \\ &= n \sum_{j=1}^k \frac{(A_{j+1} - A_j)^2 \exp(\theta(A_j + A_{j+1}))}{\exp(\theta A_{j+1}) - \exp(\theta A_j)} + n \frac{A_{k+1}^2 \exp(\theta A_{k+1})}{1 - \exp(\theta A_{k+1})}. \end{aligned} \quad (7)$$

Using $I_g(\theta)$, we can find a large sample $(1 - \alpha)100\%$ confidence interval for θ as follows:

$$\hat{\theta}_g \pm Z_{1-\alpha/2} \sqrt{I_g(\hat{\theta}_g)^{-1}}.$$

Simple calculations can show that the Fisher's information number about θ in a random sample X_1, \dots, X_n from (1) is $I(\theta) = n/\theta^2$.

5. Bayesian estimation

In classical statistics, we use the likelihood of the data, i.e., $L(\theta) = f(\underline{n}; \theta)$, to estimate the parameter θ . In Bayesian statistics, we assume that θ has a prior distribution $\pi(\theta)$, say. Then we combine the likelihood data with the prior distribution to get the posterior distribution $\pi(\theta | \underline{n})$. With respect to the squared error loss function, the posterior mean is used to estimate θ since it minimizes the posterior expected loss, i.e., $E\{(\theta - \hat{\theta})^2 | \underline{n}\}$. In this section, we use the following prior distribution for θ to derive an estimate for θ

$$\pi(\theta) = \frac{\theta^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(-\frac{\theta}{\beta}\right), \quad \theta > 0.$$

Using the Binomial theorem, we rewrite the likelihood function of the grouped data as follows

$$\begin{aligned} f(\underline{n}; \theta) &= C \prod_{r=1}^k \sum_{j_r=0}^{n_r} \binom{n_r}{j_r} (-1)^{j_r} (\exp(\theta A_{j_r+1}))^{n_r-j_r} (\exp(\theta A_{j_r+1}))^{j_r} \times \\ &\quad \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_{k+1}}{j_{k+1}} (-1)^{j_{k+1}} \exp(\theta j_{k+1} A_{k+1}), \quad \text{where} \\ &= C \sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \exp(\theta H(j_1, \dots, j_{k+1})), \\ H(j_1, \dots, j_{k+1}) &= j_{k+1} A_{k+1} + \sum_{r=1}^k (n_r - j_r) A_{j_r+1} + \sum_{r=1}^k j_r A_{j_r}. \end{aligned}$$

Combining the likelihood information with the prior information yields the posterior distribution of θ given \underline{n} , i.e.

$$\pi(\theta | \underline{n}) = \frac{f(\underline{n}; \theta)\pi(\theta)}{\int_0^{\infty} f(\underline{n}; \theta)\pi(\theta)d\theta},$$

$$\propto \sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \theta^{\alpha-1} \times \quad \text{So we get}$$

$$\exp\left(-\theta\left(\frac{1}{\beta} - H(j_1, \dots, j_{k+1})\right)\right).$$

$$\pi(\theta | \underline{n}) = \frac{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \theta^{\alpha-1} \exp\left(-\theta\left(\frac{1}{\beta} - H\right)\right)}{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \left(\frac{1}{\beta} - H\right)^{-\alpha}}, \text{ where}$$

$$H = H(j_1, \dots, j_{k+1}).$$

The Bayesian estimate of θ with respect to the squared error loss function is the posterior mean, i.e.,

$$\hat{\theta}_b = E(\theta | \underline{n}),$$

$$= \frac{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \left(\frac{1}{\beta} - H\right)^{-\alpha-1}}{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \left(\frac{1}{\beta} - H\right)^{-\alpha}}.$$

The variance of the posterior distribution is used as a measure of precision for the Bayes estimator $\hat{\theta}_b$. The posterior variance of $\hat{\theta}_b$ is

$$\text{Var}(\hat{\theta}_b | \underline{n}) = E\{\theta^2 | \underline{n}\} - \hat{\theta}_b^2,$$

where

$$E(\theta^2 | \underline{n}) = \frac{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \left(\frac{1}{\beta} - H\right)^{-\alpha-2}}{\sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \left(\frac{1}{\beta} - H\right)^{-\alpha}}$$

6. Estimation of α and β

In general the parameters α and β are unknown. Empirical Bayesian methods use the marginal density of the data, i.e., $f(\underline{n}) = f(\underline{n}; \alpha; \beta)$ to estimate them. The density $f(\underline{n})$ can be found as follows:

$$\begin{aligned} f(\underline{n}) &= \int_0^\infty f(\underline{n}; \theta) \pi(\theta) d\theta. \\ f(\underline{n}) &= C \sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} \\ &\quad \times \int_0^\infty \frac{\theta^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(-\theta\left(\frac{1}{\beta} - H\right)\right), \\ &= C \sum_{j_1=0}^{n_1} \dots \sum_{j_{k+1}=0}^{n_{k+1}} \binom{n_1}{j_1} \dots \binom{n_{k+1}}{j_{k+1}} (-1)^{j_1+\dots+j_{k+1}} (1 - \beta H)^{-\alpha}, \end{aligned}$$

where C is a constant does not depend on α and β . Since the last density depends on α and β , we may use this marginal density to estimate the parameters α and β via the standard statistical methods.

7. HPDCR for θ

A Bayesian $(1 - \alpha)100\%$ HPDCR (High Posterior Density Credible Region), \mathfrak{R} for θ is a set defined by

$$\mathfrak{R} = \{\theta : \theta > 0 \text{ and } \pi(\theta | \underline{n}) \geq a\}$$

Subject to

$$P\{\theta \in \mathfrak{R} | \underline{n}\} = 1 - \alpha.$$

If $\pi(\theta | \underline{n})$ is unimodal, then problem of finding \mathfrak{R} is equivalent to find the interval $[u, v]$ such that

$$\pi(u | \underline{n}) = \pi(v | \underline{n}) \quad \text{and} \quad P\{u \leq \theta \leq v | \underline{n}\} = 1 - \alpha\}.$$

Finding u and v requires numerical calculations.

8. Application to artificial data

In this section, we apply our results to an artificial data generated from Burr Type X. We generated a sample of size $n = 30$ from Burr Type X with $\theta = 1$. So we get the following sample:

0.982546, 1.09843, 0.31402, 1.77377, 1.51687, 0.570998, 0.826924, 0.56066, 0.807545, 0.461524, 0.418278, 0.670564, 1.01147, 0.538708, 1.21488, 0.319864, 0.656511, 1.83214, 0.602107, 0.271986, 0.778316, 1.08244, 0.558287, 0.53136, 1.34272, 2.13202, 0.437059, 0.584021, 1.65493, 1.1144.

Then we put this sample into six intervals ($k = 5$) with $\delta = 1$. Using $\alpha = 2$ and $b = 0.5$, we see that the MLE of θ from the grouped data is $\hat{\theta}_g = 1.04626$ with Variance equal to 0.00111111. We also see that $I_g(\hat{\theta}_g) = 10.7998$, $I(\hat{\theta}) = 29.8469$ and the Bayes estimator is $\hat{\theta}_b = 1.00513$ with variance 0.0785888. A 95% confidence interval for θ is (0.449845, 1.64267).

9. Conclusion

In this paper we obtained Bayesian and non-Bayesian estimators for the parameter of the Burr type X distribution when the data are given in groups. The estimators have no closed forms, so we need numerical methods to find them. The application to the artificial data shows that the estimators work well.

Acknowledgements. This research has been supported by a grant from Yarmouk University.

References

- [1] A. Amjad and B. Ayman, Interval estimation for the scale parameter of Burr type X distribution based on grouped data. *Journal of Modern Applied Statistical Methods*, 3 (2006), 386-398.
- [2] M. T. Alodat and M. F. Al-Saleh, Bayesian estimation using grouped data with application to the exponential distribution. *Soochow J. of Mathematics*, 26(2000), 342-357.

- [3] M. T. Alodat, K. M. Aludaat and T. T. Alodat, Bayesian prediction intervals from grouped data: exponential distribution. *Abhath Al-yarmouk*, 2007 (Accepted)
- [4] I. W. Burr, Cumulative frequency functions. *Ann. of. Math. Statist.*,13 (1941), 215-232.
- [5] D. Heitjan, Inference from grouped data: a review (with discussion). *Statistical sciences*, 4(1989), 164-183.
- [6] Kuldorff, *Estimation from grouped and partially grouped samples*. John Wiley, Inc. New York, 1961.
- [7] M. Z. Raqab and D. Kundu, Burr type X distribution: revisited. *JPSS*, 8(2006), 179-198.
- [8] M. J. Schervish, *Theory of Statistics*. Springer-Verlag, New York, Inc., 1995.
- [9] J. G. Surles and W. J. Padgett, Inference for reliability and stress-length for a scaled Burr type X distribution. *Lifetime Data analysis*, 7(2001), 187-2002.
- [10] X. Wu and J. M. Perloff, "China's income distribution: 1985-2001." *Review of Econometrics and Statistics*, 87(2005), 763-775.
- [11] C. B. Pipper and C. Ritz, Cheking the grouped data version of Cox model for interval-grouped survival data. *Scandinavian Journal of Statistics*, 10(2006), 1467-9469.

Received: June 15, 2007