BAYESIAN INFERENCE OF A FINITE POPULATION UNDER SELECTION BIAS

By

Zhiqing Xu

A Thesis

Submitted to the Faculty

Of

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

In

Applied Statistics

May 1, 2014

APPROVED:

_____

Professor Balgobin Nandram, Major Thesis Advisor

Acknowledgments

I would like to express my deep sense of gratitude to my advisor, Dr. Balgobin Nandram for the continuous support of my study, for his patience, motivation and immence knowledge. His guidance helped me through all the difficulties in the process of research and writing this thesis.

My sincere thanks also to Dr. Joseph D. Petrucelli for his insightful comments, humor and hard questions during the last two years.

I would also like to thank Dr. Sungsu Kim who provided valuable advises for my thesis writing.

Last but not the least, thanks to my family and friends for their unconditional support.

ABSTRACT

Length-biased sampling method gives the samples from a weighted distribution. With the underlying distribution of the population, one can estimate the attributes of the population by converting the weighted samples.

In this thesis, generalized gamma distribution is considered as the underlying distribution of the population and the inference of the weighted distribution is made. Both the models with known and unknown finite population size are considered.

In the modes with known finite population size, maximum likelihood estimation and bootstrapping methods are attempted to derive the distributions of the parameters and population mean. For the sake of comparison, both the models with and without the selection bias are built. The computer simulation results show the model with selection bias gives better prediction for the population mean.

In the model with unknown finite population size, the distributions of the population size as well as the sample complements are derived. Bayesian analysis is performed using numerical methods. Both the Gibbs sampler and random sampling method are employed to generate the parameters from their joint posterior distribution. The fitness of the size-biased samples are checked by utilizing conditional predictive ordinate.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The goal of sampling methods is to provide information on a population by studying only a subset of it, called a sample. Sampling is the process of selecting units from a population so that the sample allows estimating unknown quantities of the population. This paper is an attempt to present several modeling methods for the case when the samples are selected with probability proportional to size (size biased sampling). In this chapter, both length-biased distribution and line intercept sampling are introduced, and generalized gamma distribution is considered.

1.1 Length-Biased Distribution

Unequal probability sampling method was first suggested by Hansen and Hurwitz (1943). They demonstrated that the use of unequal selection probabilities frequently allowed more efficient estimators of the population total than did equal probability sampling.

The sampling procedure of Hansen and Hurwitz (1943) was size-biased sampling. It occurs when the sample selection probabilities are correlated with the values of the study variable. For example, consider the problem of estimating the mean size of the shrubs in a region of area A. Assume the shrub is selected with probability proportion to its area X. Then the probability of selecting the i^{th} shrub is

$$w(X_i) = \frac{X_i}{A}, i = 1, 2, \cdots, N$$

where N is the total number of shrubs.

Size-biased distributions are a special case of the more general form known as weighted distributions. First introduced by Fisher (1934) to model ascertainment bias, weighted distributions were later formalized in a unifying theory by Rao (1965). Briefly, if the random variable X has the *pdf* of $f(x)$, and the non-negative weight function is $w(x)$, then the corresponding weighted density function is

$$g(x) = \frac{w(x)f(x)}{\int w(x)f(x)dx}$$

A special case of interest arises when the weight function is of the form $w(x) = x$. Such distributions are known as length-biased distribution and are written as:

$$g(x) = \frac{xf(x)}{\mu},$$

where $\mu = \int xf(x)$ is the first raw moment of $f(x)$.

Weighted distributions in general and length-biased distributions in particular are very useful and convenient for the analysis of lifetime data. Weighted distributions occur frequently in research related to reliability, biomedicine, ecology and several other areas. Various works are done to characterize relationships between original distributions and their length-biased forms. Muttlak (1990) suggested using ranked set sampling procedure to estimate the population size and population mean. Nandram (2013) proposed using a Bayesian nonignorable selection model to accommodate the selection mechanism.

1.2 Line Intercept Sampling

Line intercept is a length-biased method of sampling particles in a study region. In general, the particles may be of any shape and size and may possess an arbitrary spatial distribution. For example, the particles may be shrubs or patches of vegetation in a field, or the projection of logs on the forest floor. It is often of interest to study certain quantitative characteristics of these particles.

The idea of line intercept sampling is using a line (transect) as a sampling unit and measuring features of the particles that crossed by it. It has found widespread application for the purpose of estimating particles density, cover and yield.

Consider a study region of area A and define the following parameters (some are unknown and to be estimated):

- n: The number of distinct shrubs intersected.
- N: The total number of distinct particles in the study region.
- X_i: The width of the i^{th} intersected shrub, i.e., the distance between tangents of the shrub that are parallel to the transect, $i = 1, 2, \cdots n$.
- V_i: The length of intersection of the i^{th} shrub along the transect, $i = 1, 2, \cdots n$.

For illustration see Fig. 1.1.

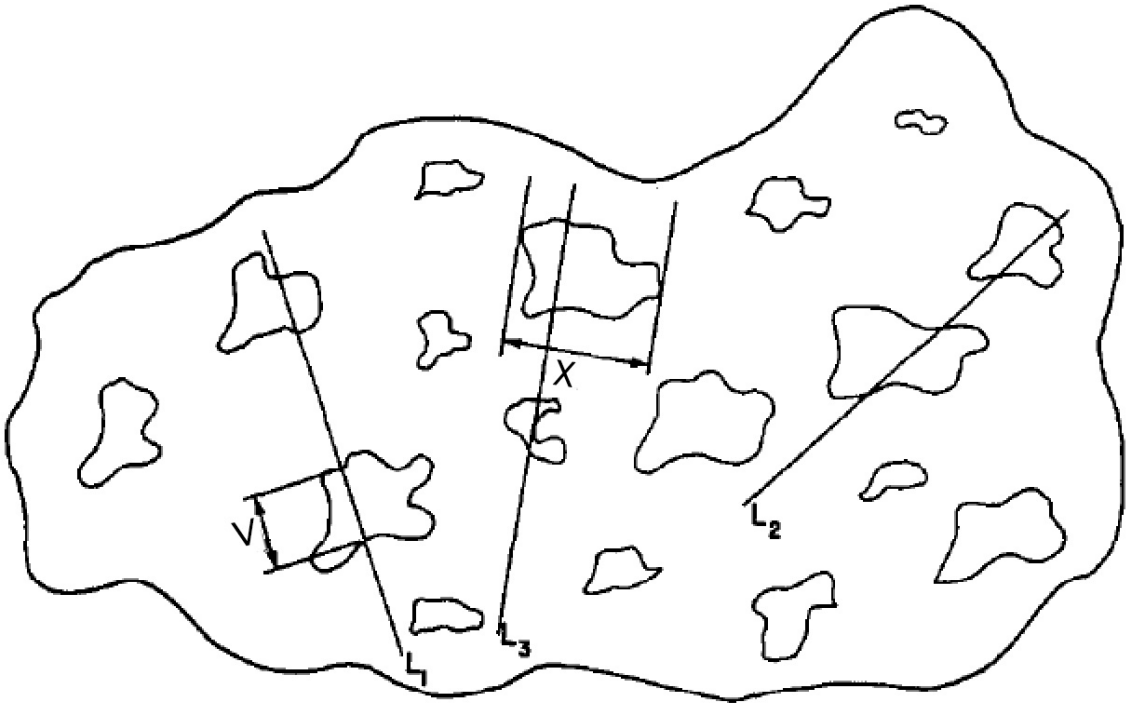

Fig. 1.1: Study region of area A with $N = 17$ distinct particles, $n = 6$ intersected particles, and $k = 3$ transects of equal length.

In general, shrubs can be collected from either randomly located or systematically located transects (Butler 1983). For example, Lucas and Seber (1977) derived unbiased estimators for density and percentage cover for any spatial distribution and randomly located transect. Eberhardt (1978) derived an unbiased estimator of density using parallel but randomly located transects. McDonald (1980) showed that the Lucas and Seber estimators for density and percentage cover are unbiased for a simple random sample of unequal length transects.

1.3 Description of the Data

The data we use was collected using the line intercept sampling method, published in Muttlak (1990). The study was conducted in a limestone quarry dominated by regrowth

of mountain mahogany. The study area was defined by the area east of the baseline and within the walls of the quarry, where the baseline was established approximately parallel to the fissures. By dividing the baseline into three equal parts, three systematically placed transects were established. To ensure uniform coverage over the study area, two independent replications, each with 3 transects were selected (Fig. 1.2).

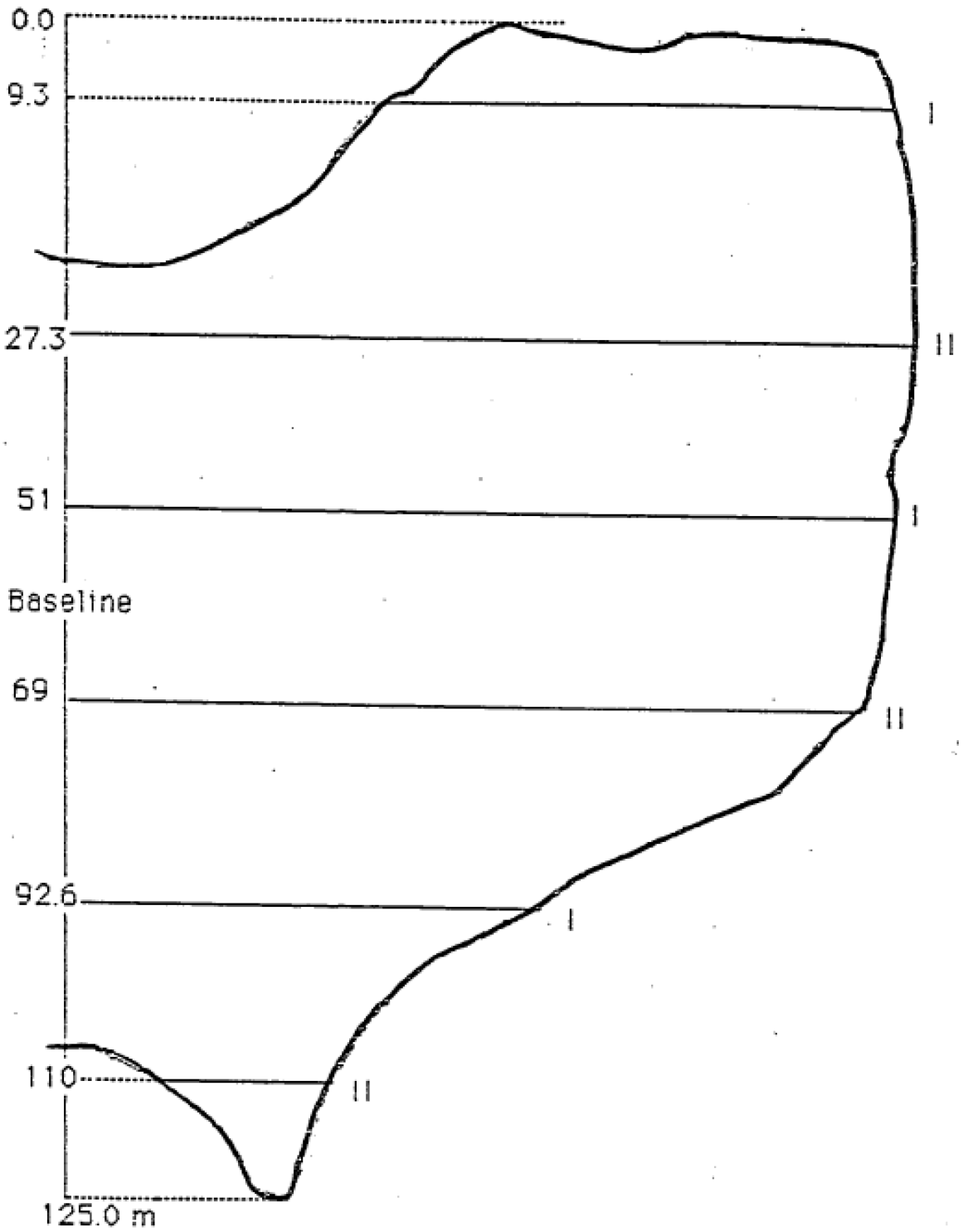

Fig. 1.2: Sketch of the study area

In our models, we are interested in estimating the mean width of the shrubs in the area.

So the variable is the width of the projection of the shrub encountered by transects onto the baseline. We use the data from both replications, as showed in Table 1.1 and 1.2.

Table 1.1: Widths of shrubs in replication 1

Transect	X_i=width
I	1.53 .87 .79 .78 1.85 1.45 .48 .52 .22 .38 .59
	.20 .42 1.02 .97 .56 .62 .42
II	1.15 .87 .57 .97 .57 1.97 .58,2.54 1.85 .35 1.24
	1.8 .78 .98 1.3 1.55 1.69 2.12 1.27 .75 1.01 1.82
III	.71,1.5,1.82,1.86,1.61,1.21

Table 1.2: Widths of shrubs in replication 2

Transect	X_i=width
I	.67 .31 .83 1.95 1.36 1.45 .72 1.15 .98 1.29 .88
	.25 .63 1.12 .34 .21 1.36 .95 1.04 .48 1.05 .88
	.16 1.08 .95 .25 .3 1.4 .58 .73 1.3 .57
II	.96 2.08 .68 1.39 .5 .72 .19 1.91 .88 .48 .12

From the box plots of these two replications (Fig. 1.3), we notice the differences of the median and the variance between these three transects in replication 1; whereas in replication 2, there are little differences. So when making inferences using replication 1, we need to regard the data from these three transects as from three different strata, and model them respectively.

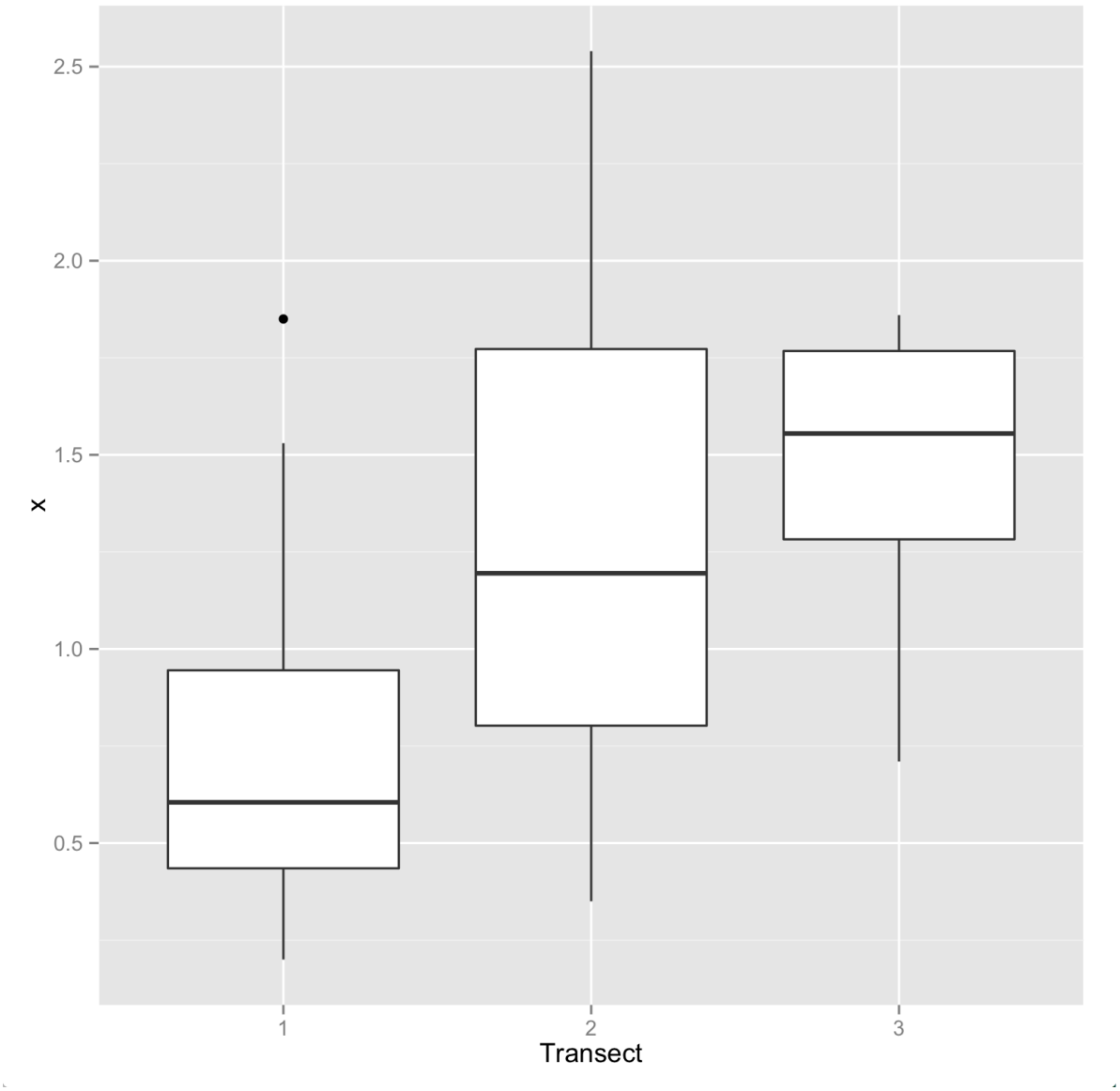

Fig. 1.3: Box plot of the first replication

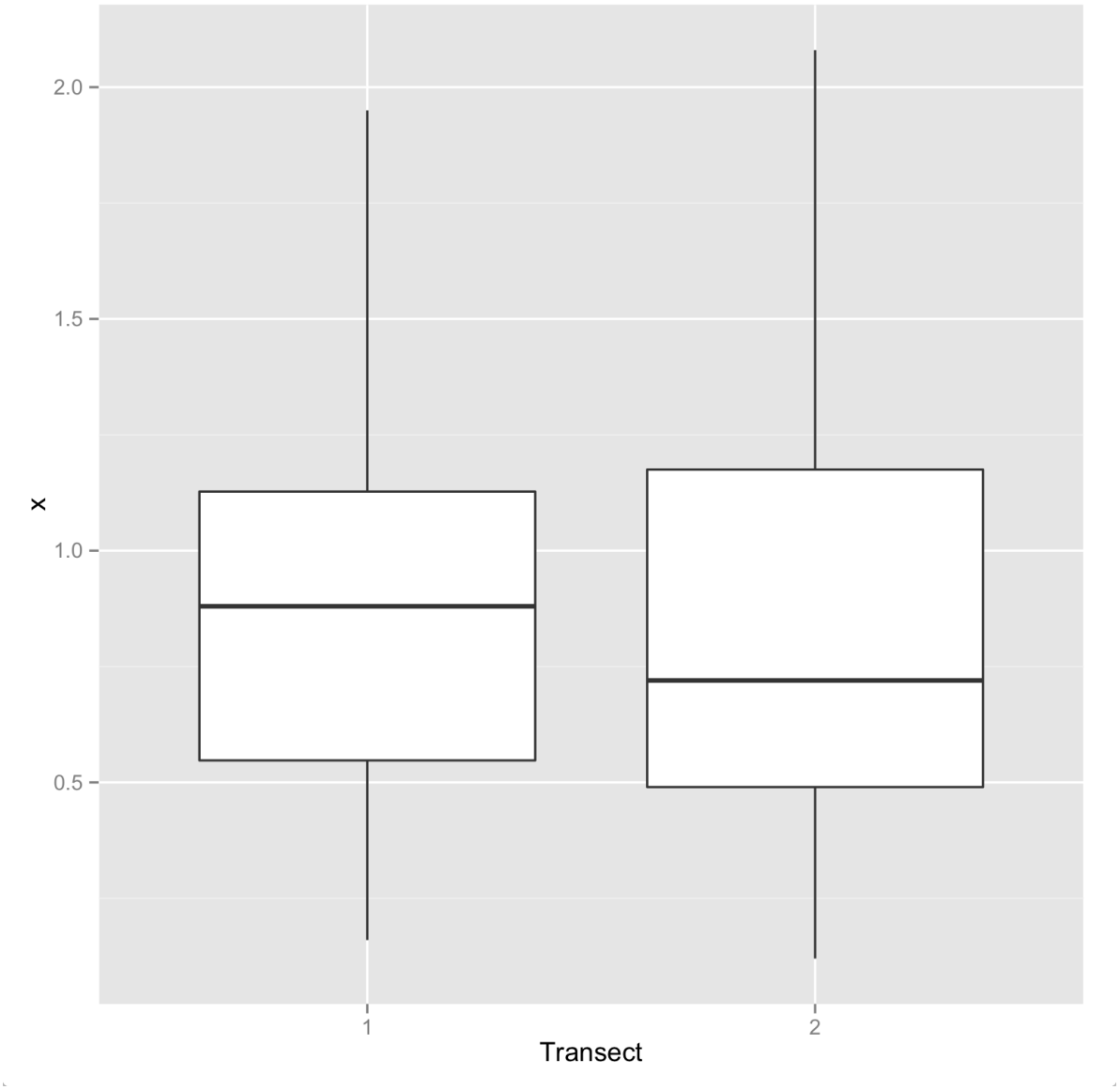

Fig. 1.4: Box plot of the second replication

1.4 Generalized Gamma Distribution

Generalized Gamma Distribution (GG) is an important distribution in statistics. It was first defined by Stacy (1962) and presented a flexible family in the varieties of shapes and hazard functions for modeling duration. It encompasses exponential, gamma, and

Weibull as subfamilies, and lognormal as a limiting distribution. Some authors have argued that the flexibility of GG makes it suitable for duration analysis, while others have advocated use of simpler models because of estimation difficulties caused by the complexity of GG parameter structure. The GG family and its properties has been remarkably presented in different papers. Prentice (1974) resolved the convergence problem using a nonlinear transformation of GG model. Hirose (1999) introduced maximum likelihood parameters estimation by model augmentation. Hwang, et al (2006) introduced a new moment estimation of parameters of the generalized gamma distribution using its characterization.

Because of the flexibility and the importance of Generalized Gamma Distribution, we use it as the underlying population distribution for our models.

The probability density of the generalized gamma distribution is given by

$$f(x|\alpha,\beta,\gamma) = \frac{\gamma x^{\gamma\alpha-1}}{\beta^{\gamma\alpha}\Gamma(\alpha)} exp\left[-\left(\frac{x}{\beta}\right)^{\gamma}\right]. \tag{1.1}$$

The mean and variance have important role,

$$E(X) = \frac{\beta\Gamma(\alpha+\frac{1}{\gamma})}{\Gamma(\alpha)}, \tag{1.2}$$

$$Var(x) = \frac{\beta^2\Gamma(\alpha+\frac{2}{\gamma})}{\Gamma(\alpha)} - \left(\frac{\beta\Gamma(\alpha+\frac{1}{\gamma})}{\Gamma(\alpha)}\right)^2. \tag{1.3}$$

Khodabin (2010) provided details of the subfamilies of the generalized gamma distribution (Table 1.3).

Table 1.3: Mean and variance for subfamilies of generalized gamma distribution

Distribution name	α	β	γ	Mean	Variance
Exponential	1	β	1	β	β^2
Gamma	α	β	1	$\alpha\beta$	$\alpha\beta^2$
Weibull	1	β	γ	$\beta\Gamma(1+\frac{1}{\gamma})$	$\beta^2\Gamma(1+\frac{2}{\gamma}) - \left(\beta\Gamma(1+\frac{1}{\gamma})\right)^2$
Generalized normal	α	β	2	$\frac{\beta\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}$	$\beta^2\alpha - \left(\frac{\beta\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}\right)^2$
Half normal	.5	$\sqrt{2\sigma^2}$	2	$\sigma\sqrt{\frac{2}{\pi}}$	$\sigma^2\left(1-\frac{2}{\pi}\right)$
Rayleigh	1	$\sqrt{2\sigma^2}$	2	$\sigma\sqrt{\frac{2}{\pi}}$	$\sigma^2\left(1-\frac{2}{\pi}\right)$
Maxwell Boltzmann	$\frac{3}{2}$	β	2	$\frac{2\beta}{\sqrt{\pi}}$	$\beta^2\left(1-\frac{4}{\pi}\right)$
Chi-square	$\frac{k}{2}$	β	2	$\frac{\beta\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}$	$\frac{\beta^2\Gamma(\frac{k+2}{2})}{\Gamma(\frac{k}{2})} - \left(\frac{\beta\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})}\right)^2$

Chapter 2

Models with Known Finite Population Size

To simplify the models, population size are not considered as random variable in this chapter. We use the maximum likelihood method to estimate the parameters of generalized gamma distribution and use bootstrap to obtain the distributions of them. For the sake of comparison, we fit both model without selection bias and with selection bias, then compare them by computer simulation trials to gain some insight.

2.1 The Model Without Selection Bias

In this Section, we do not consider the selection bias.

2.1.1 The Finite Population Size

We first consider the problem in which a random (representative) sample is taken from a finite population.

The general estimator of population size N is

$$\hat{N} = \sum_{i=1}^{n} \frac{1}{\pi_i},$$

where π_i is the probability that the i^{th} unit is selected, see Cochran(1977). Since we

regard the sample as simple random sample,

$$\pi_i = \frac{\bar{x}}{W},$$

where $W = 125$ is the length of the base line. Then,

$$\hat{N} = 125\frac{n}{\bar{x}}.$$

To avoid double-use of the data, we estimate $\hat{N}$ by replication 2. The estimated value is $\hat{N} = 6225$.

2.1.2 Parameters and Population mean

Suppose the population has the generalized gamma distribution. Then, the selection probability for each sample x_i would be

$$f(x_i|\alpha,\beta,\gamma) = \frac{\gamma x_i^{\gamma\alpha-1}}{\beta^{\gamma\alpha}\Gamma(\alpha)} exp\left[-\left(\frac{x_i}{\beta}\right)^{\gamma}\right]. \tag{2.1}$$

Remember in Section1.3, we noticed the differences between the three transects in replication 1 are so significant that three different distributions are called for. However, due to the sparseness of data, we cannot fit three generalized gamma distribution independently. A trade-off solution is regarding these three gamma distributions with different

β (scale parameter), but with the same α and γ. Then, the likelihood function is

$$\begin{aligned} Lik(\alpha, \beta_1, \beta_2, \beta_3, \gamma | x_{11}, \cdots, x_{3n_3}) &= \prod_{i=1}^{3} \prod_{j=1}^{n_i} \frac{\gamma x_{ij}^{\gamma\alpha-1}}{\beta_i^{\gamma\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x_{ij}}{\beta_i}\right)^{\gamma}\right] \\ &= \frac{\gamma^n \left(\prod_{i=1}^{3}\prod_{j=1}^{n_i} x_{ij}\right)^{\gamma\alpha-1}}{(\beta_1^{n_1}\beta_2^{n_2}\beta_3^{n_3})^{\gamma\alpha}\left[\Gamma(\alpha)\right]^n} \exp\left[-\sum_{i=1}^{3}\sum_{j=1}^{n_i}\left(\frac{x_{ij}}{\beta_i}\right)^{\gamma}\right]. \end{aligned} \tag{2.2}$$

There are two approaches to obtain the values (distributions) of these parameters. The first approach is using the maximum likelihood estimation with the restriction that every parameter should be larger than 0 to obtain the estimated parameters $\alpha, \beta_1, \beta_2, \beta_3, \gamma$ for the sample values. Then by bootstrapping the sample values and repeating the above procedure M times to get the distributions of these parameters. The second approach is generating the samples of the parameters from their joint posterior distribution directly.

In this chapter, we use the first approach, which is obtaining the estimated parameters by maximum likelihood estimation (using Nelder-Mead to perform the optimization) and the bootstrapping (using Parzen-Rosenblatt Kernel Density to smooth the data) to obtain the distributions of the parameters as well as of the population mean. The estimated values are presented in Table 2.1, and the histogram of their distributions are compared with the ones from the model with selection bias in Fig. 2.1.

Table 2.1: Estimated values of the parameters and the population mean under model without selection bias

α	β_1	β_2	β_3	γ	$\bar{x}$
1.34	0.72	1.22	1.32	1.73	1.12

2.2 The Model With Selection Bias

In this section, we added the selection bias into our model.

2.2.1 The Finite Population Size

From Chapter 1, we know that line intercept sampling gives the length biased data. Assume n units are selected, the estimator of the population size N can be denoted by

$$\hat{N} = \sum_{i=1}^{n} \frac{1}{\pi_i}$$

where π_i is the probability that the i^{th} unit is selected. It can be shown that $\hat{N}$ is an unbiased estimator of N (Cochran 1977). Since we are sampling with probability proportion to size x, then

$$\pi_i = Cx_i,$$

here C is a constant and $C = \frac{1}{W}$, where $W = 125$ is the length of the base line in line intercept sampling.

The estimated value of N under selection bias is given by

$$\hat{N} = 125 \times \sum_{i=1}^{n} \frac{1}{x_i}$$

Using the data from replication 2, we have $\hat{N} = 10061$.

2.2.2 The Sample Distribution

The samples has the *pdf* of the form

$$\begin{aligned} g(x) &= \frac{xf(x)}{\int xf(x)} \\ &= \frac{xf(x)}{E(X)}, \end{aligned}$$

where $E(X)$ is the expectation of x in the unweighted density function $f(x)$.

Let $I_i = 1$ if $i \in s$ and $I_i = 0$ if $i \notin s$, the length biased sample pdf $g(x)$ is defined as,

$$\begin{aligned} g(x_i|\alpha,\beta,\gamma) &= f(x_i|I_i = 1) \\ &= \frac{f(I_i = 1|x_i)f(x_i)}{\int f(I_i = 1|x_i)f(x_i)dx_i} \\ &= \frac{\pi_i f(x_i)}{\int \pi_i f(x_i)dx_i} \qquad (\pi_i = Pr(i \in s) = Cx_i) \\ &= \frac{x_i f(x_i)}{E(X)} \\ &= \frac{\frac{\gamma x_i^{\gamma\alpha}}{\beta^{\gamma\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x_i}{\beta}\right)^{\gamma}\right]}{\beta\frac{\Gamma(\alpha+\frac{1}{\gamma})}{\Gamma(\alpha)}} \\ &= \frac{\gamma x_i^{\gamma\alpha}}{\beta^{\gamma\alpha+1}\Gamma(\alpha+\frac{1}{\gamma})} \exp\left[-\left(\frac{x_i}{\beta}\right)^{\gamma}\right]. \end{aligned} \tag{2.3}$$

Note that $g(x)$ is also from generalized gamma distribution with parameters $\alpha_g = \alpha_f + \frac{1}{\gamma_f}, \beta_g = \beta_f$, and $\gamma_g = \gamma_f$. Similar to the model without selection bias, we consider the three distributions with different βs, but the same α and γ.

2.2.3 Parameters and Population mean

The likelihood function is

$$Lik(\alpha, \beta_1, \beta_2, \beta_3, \gamma | x_{11}, \cdots, x_{3n_3}) = \prod_{i=1}^{3} \prod_{j=1}^{n_i} \frac{\gamma x_{ij}^{\gamma\alpha}}{\beta_i^{\gamma\alpha+1}\Gamma(\alpha + \frac{1}{\gamma})} \exp\left[-\left(\frac{x_{ij}}{\beta_i}\right)^{\gamma}\right]$$
$$= \frac{\gamma^n \left(\prod_{i=1}^{3} \prod_{j=1}^{n_i} x_{ij}\right)^{\gamma\alpha}}{(\beta_1^{n_1}\beta_2^{n_2}\beta_3^{n_3})^{\gamma\alpha+1}\left[\Gamma(\alpha + \frac{1}{\gamma})\right]^n} \exp\left[-\sum_{i=1}^{3}\sum_{j=1}^{n_i}\left(\frac{x_{ij}}{\beta_i}\right)^{\gamma}\right], \tag{2.4}$$

where n_1, n_2, n_3 are the number of shrubs in each transect respectively.

By using the Maximum Likelihood Estimation, we presented the estimated values (Table 2.2) and the compared distributions (Fig. 2.1).

Table 2.2: Estimated values of the parameters and the population mean under model with selection bias

α	β_1	β_2	β_3	γ	$\bar{x}$
1.05	0.68	1.05	1.06	1.93	0.83

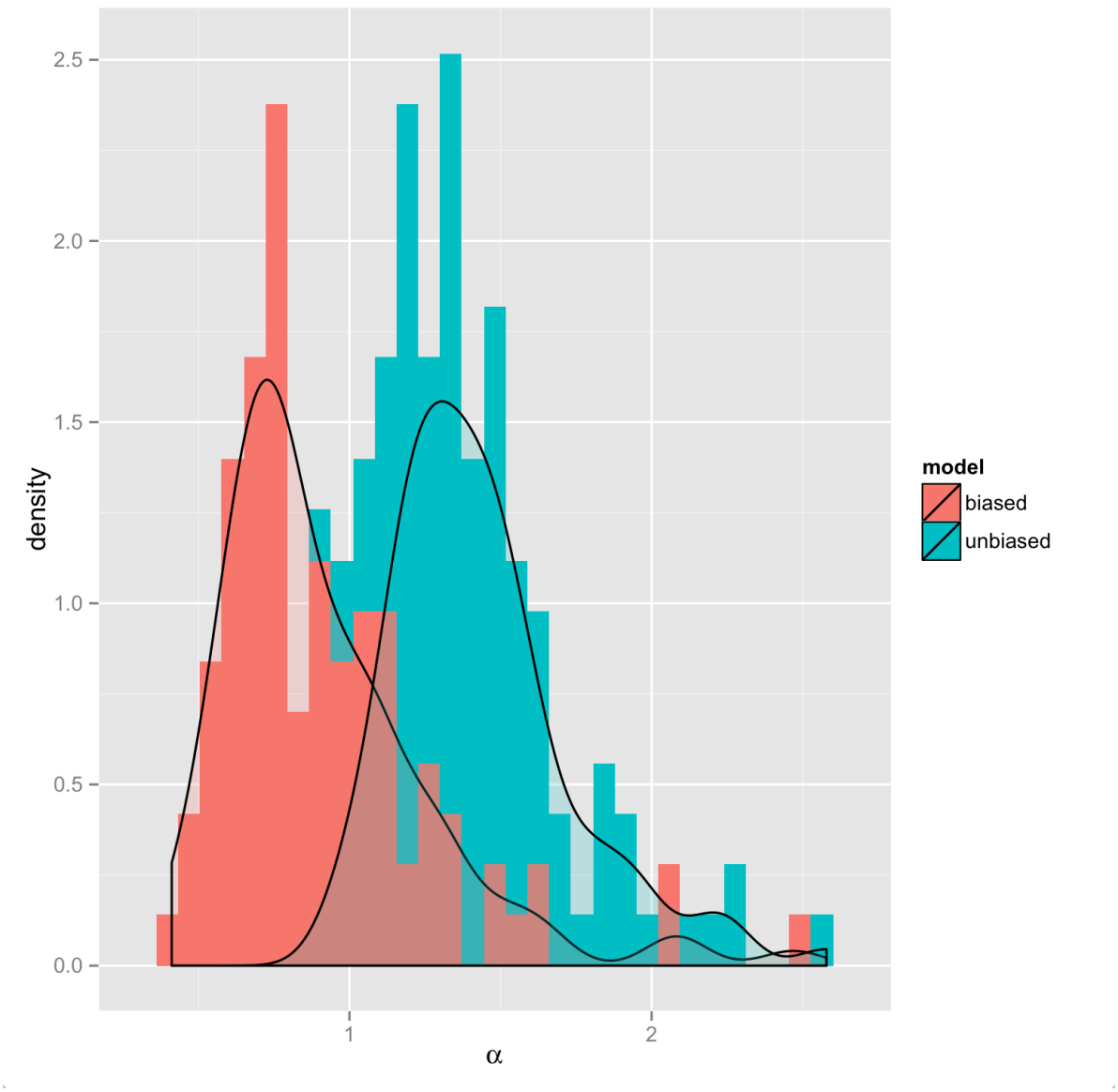

Fig. 2.1: Distribution of α estimated by MLE

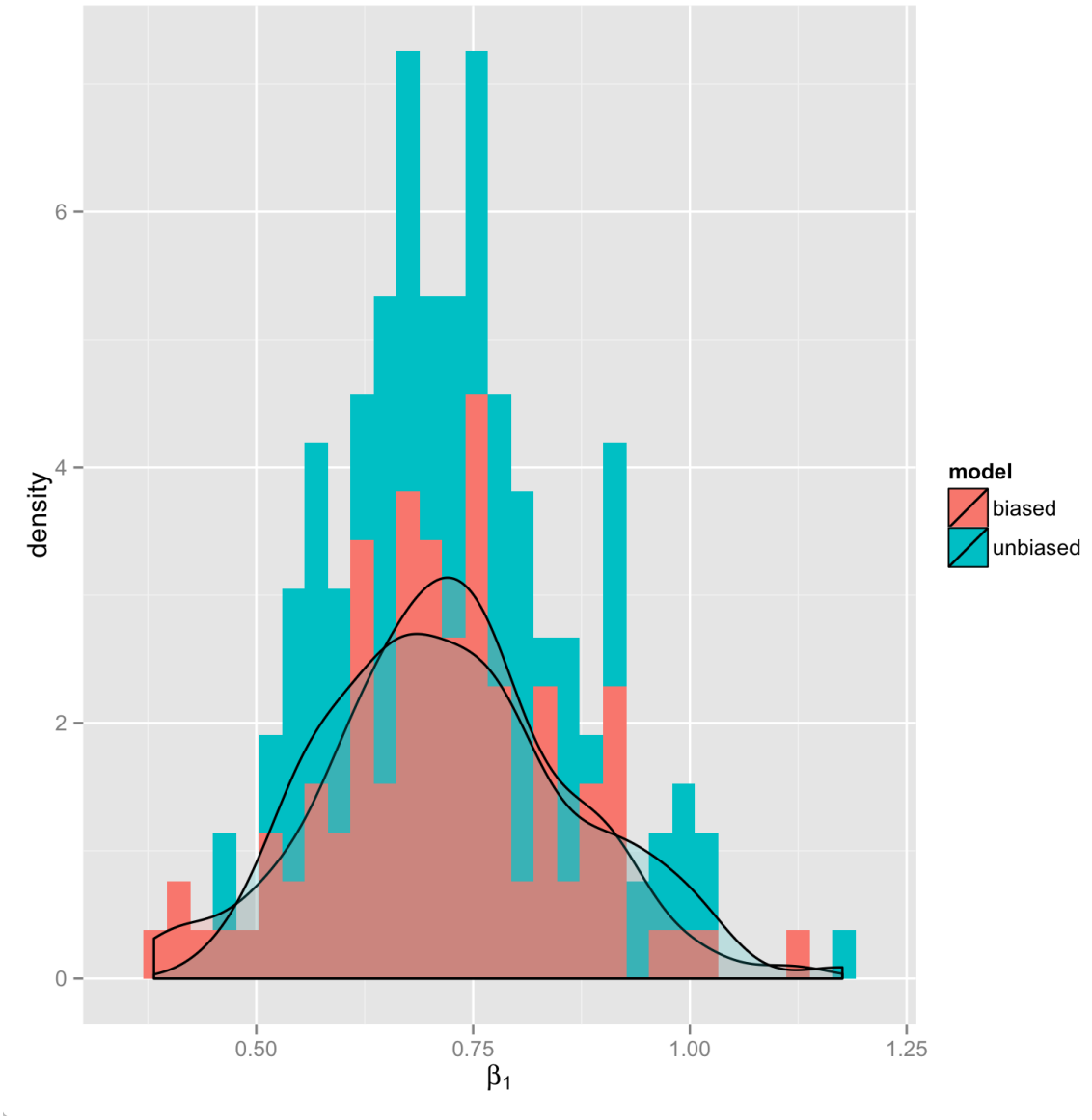

Fig. 2.2: Distribution of β_1 estimated by MLE

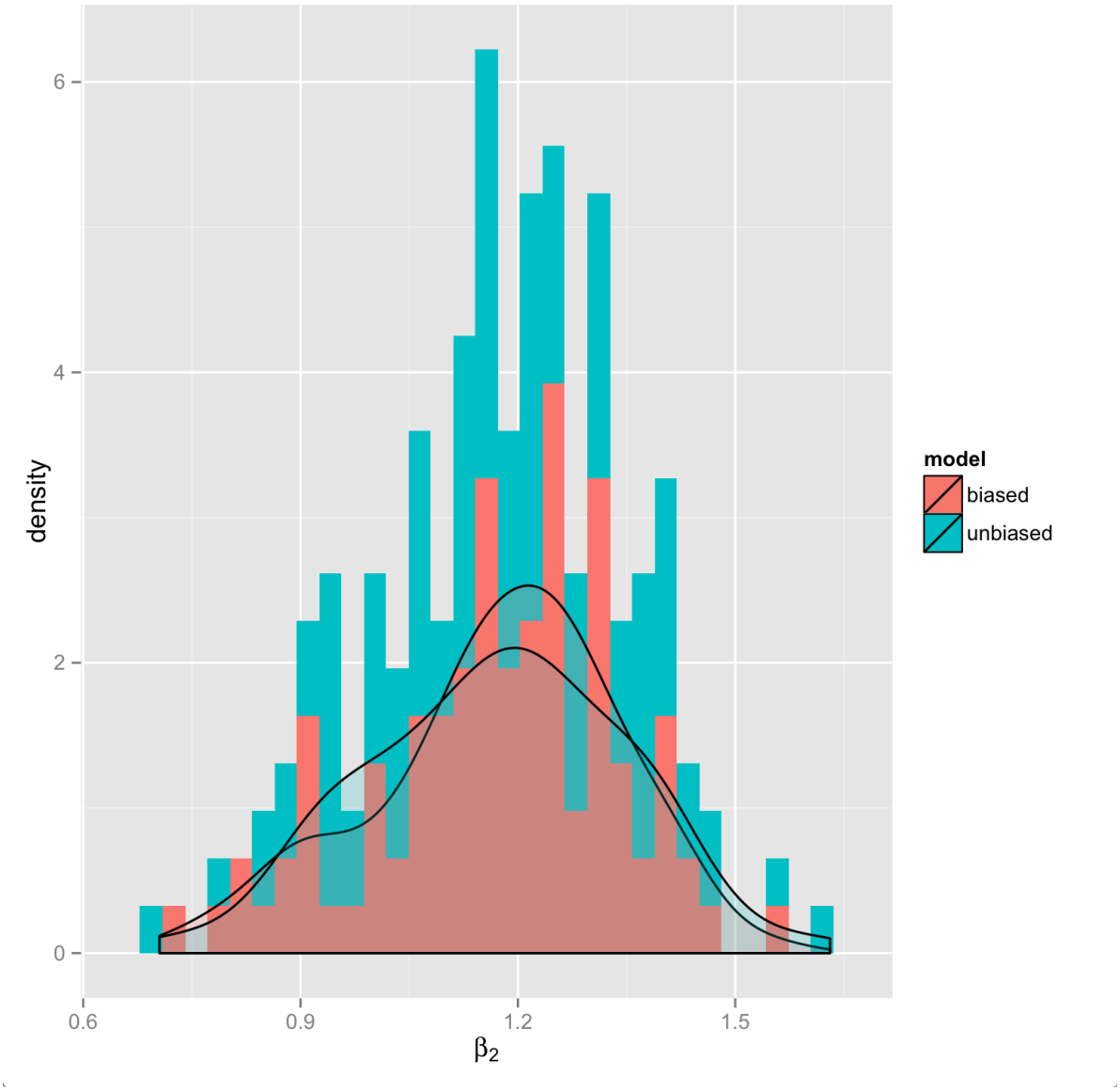

Fig. 2.3: Distribution of β_2 estimated by MLE

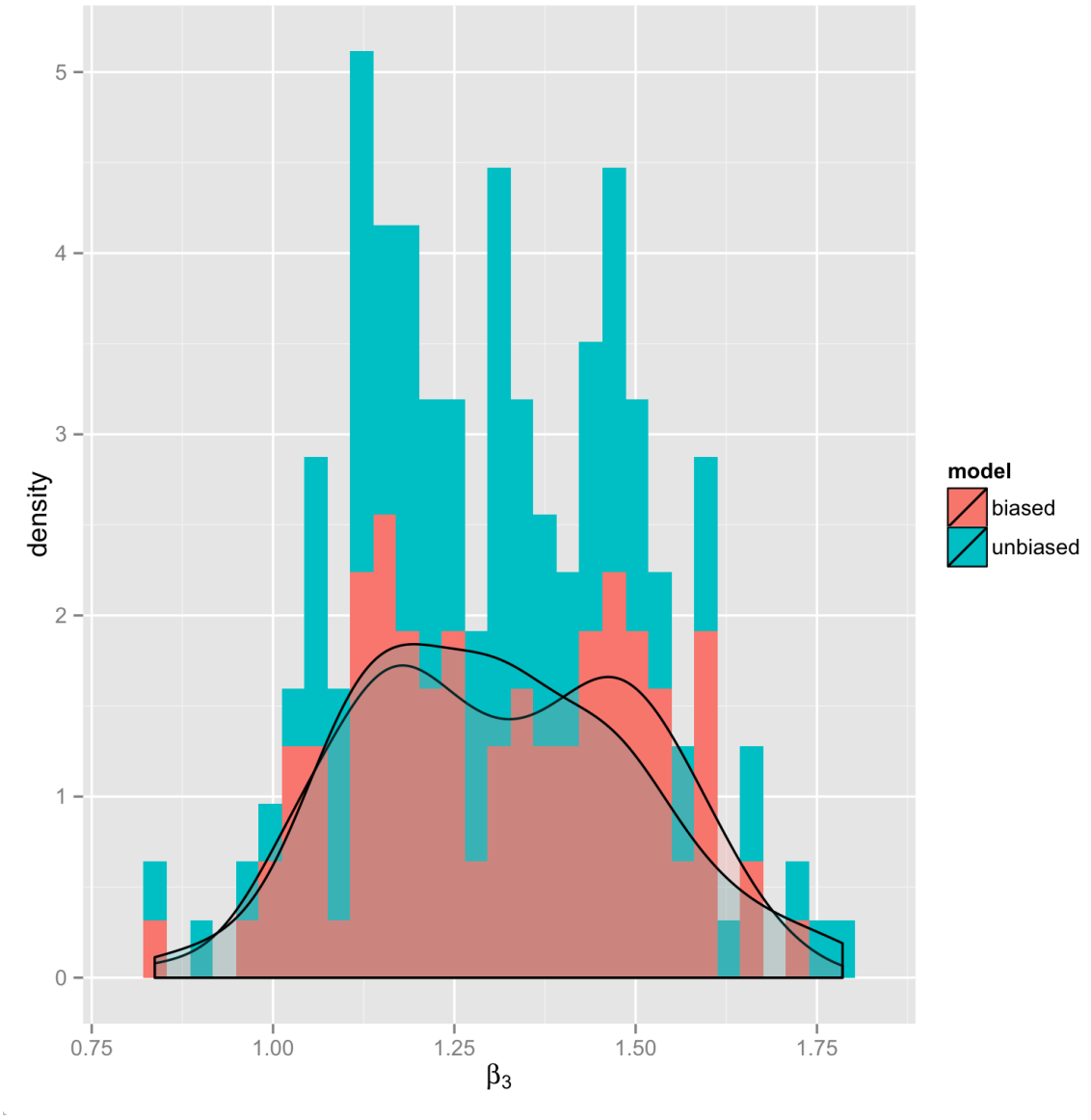

Fig. 2.4: Distribution of β_3 estimated by MLE

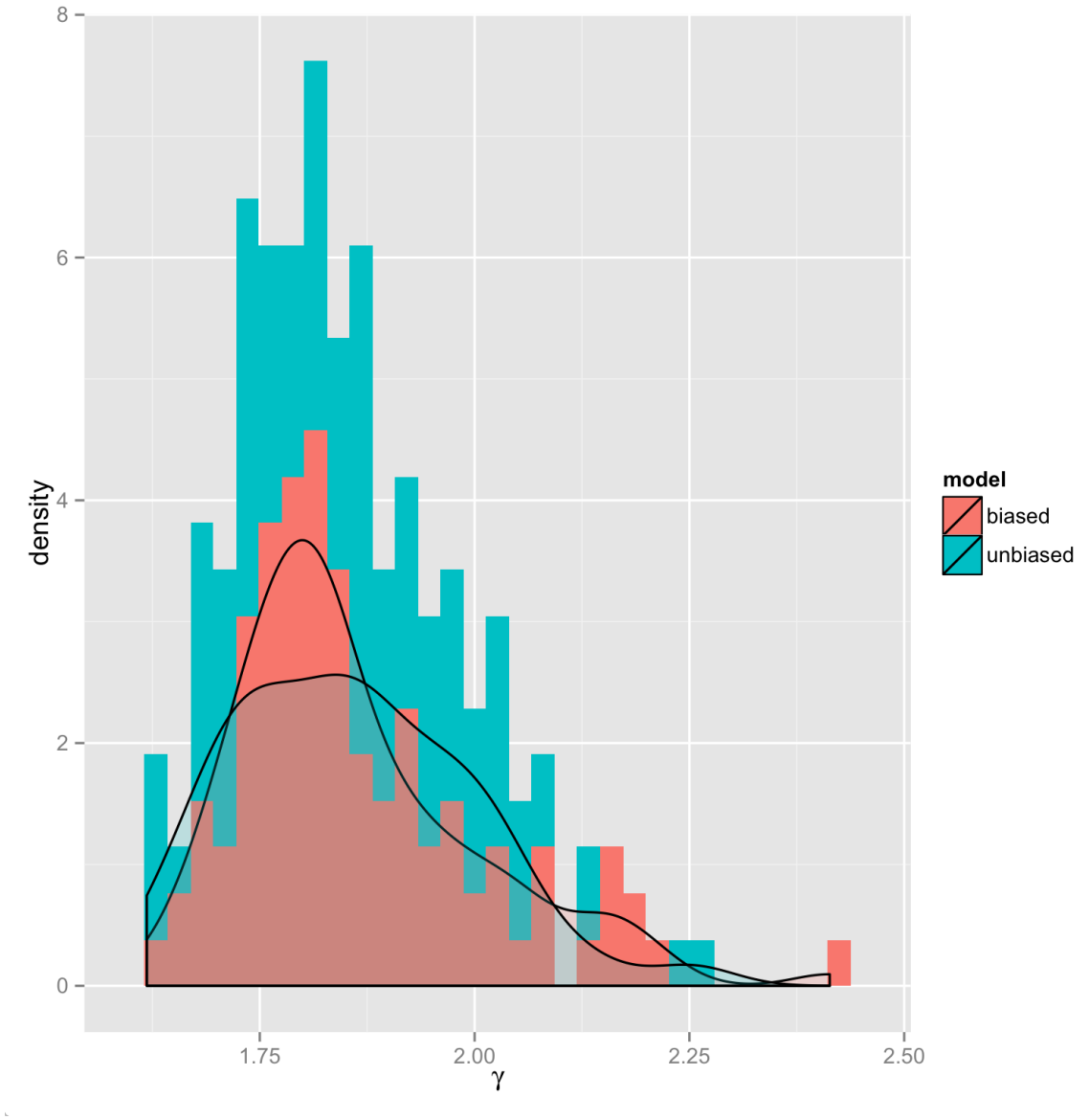

Fig. 2.5: Distribution of γ estimated by MLE

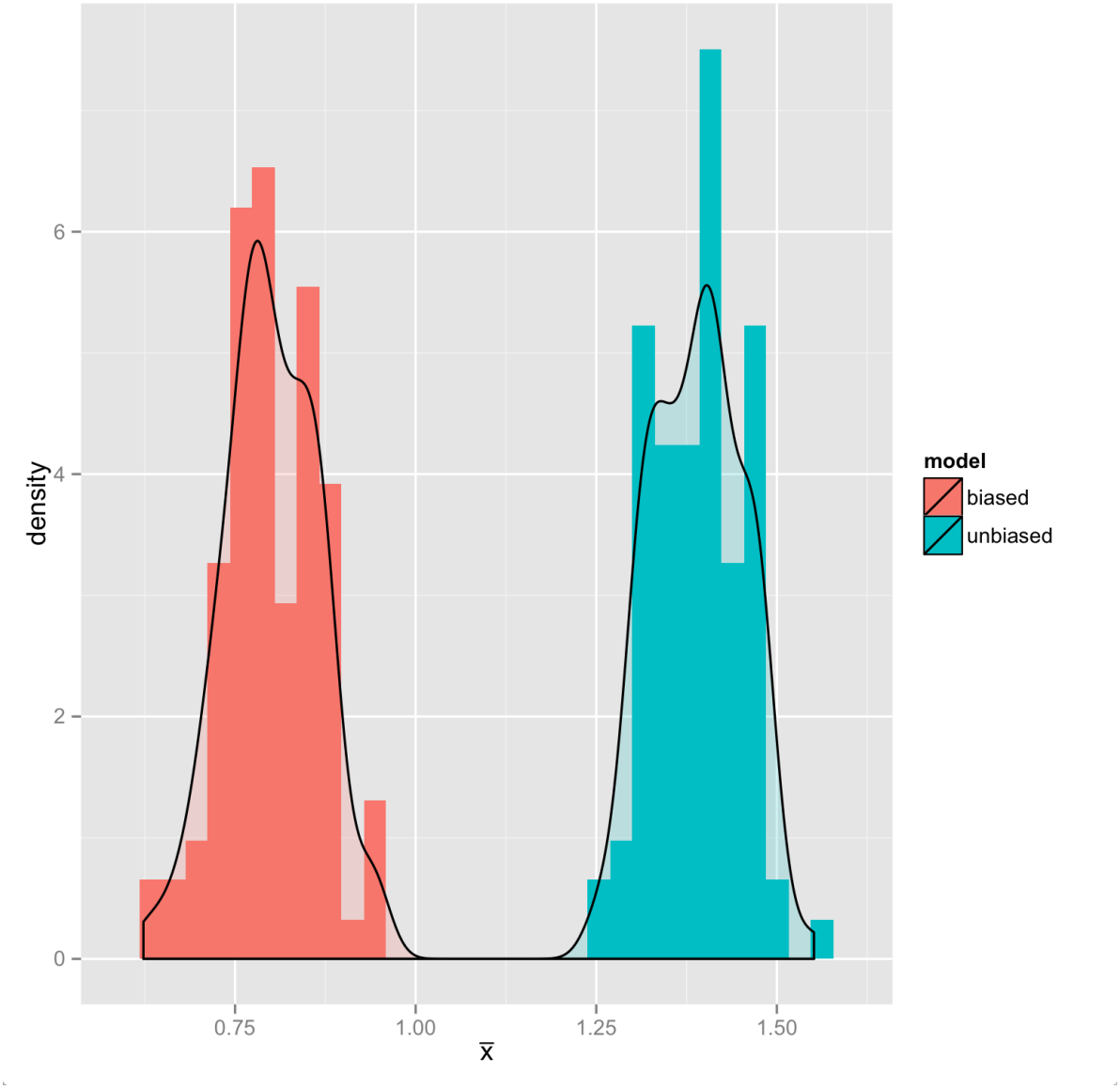

Fig. 2.6: Distribution of $\bar{X}$ estimated by MLE

2.3 Model Checking by Computer Simulation Results

From the former two sections, we see the model with selection bias gives the population mean smaller than the model without selection bias. This result is reasonable because the length biased sampling method tend to sample larger values and left smaller ones. The model with selection bias adjusts this bias in some degree. But to find out how well it adjusts the bias, we need to use computer simulation trials.

To make the trials more close to our data, we set up three strata, each has a certain generalized gamma distribution. These three generalized gamma distributions have the same α and γ, but different βs. There are 4000 random numbers in the first strata, 5000 in the second, and 1000 in the third.

We draw 50 samples out of the population, each had the chance of selection proportion to its magnitude. The parameters of the population are in Table 2.3.

The model without selection bias gives the parameters showed in Table 2.4.

The model with selection bias gives the parameters showed in Table 2.5.

Table 2.3: Population parameters and population mean

N	α	β_1	β_2	β_3	γ	$\bar{x}$
10000	1.3	0.7	1.2	1.3	1.7	1.07

Table 2.4: Estimated parameters and population mean in model without selection bias

α	β_1	β_2	β_3	γ	$\bar{x}$
1.80	0.71	1.41	1.25	1.54	1.46

95% Bootstrap Confidence Intervals of $\bar{x}$ is (1.38, 1.72).

Table 2.5: Estimated parameters and population mean in model with selection bias

α	β_1	β_2	β_3	γ	$\bar{x}$
1.17	0.71	1.41	1.24	1.7	0.93

95% Bootstrap Confidence Intervals of $\bar{x}$ is (.86, 1.11).

Chapter 3

Models with Unknown Finite Population Size

In many studies the population size N is unknown and it needs to be considered as a random variable. In this chapter, we derived the sample size distribution and the sample-complement distribution. We performed Bayesian analysis of the posterior distribution. Although analytical Bayesian inference is not possible, two numerical methods are proposed in this Chapter.

3.1 Distribution of the Population Size N

We define the sample size n is from Binomial distribution, that is

$$n|N, \mu_0 \sim Bin(N, \mu_0), n = 0, 1, 2, \cdots, N$$

and the prior distribution of N has the form of

$$\Pi(N) \propto \frac{1}{N}, N = 1, 2, \cdots$$

By Bayes' theorem, the posterior density of N is

$$\begin{aligned}\Pi(N|n,\mu_0) &\propto \frac{1}{N}\cdot\frac{N!}{n!(N-n)!}\mu_0^n(1-\mu_0)^{N-n}\\ &\propto \frac{(N-1)!}{(n-1)!(N-n)!}\mu_0^n(1-\mu_0)^{N-n}. \qquad (3.1)\end{aligned}$$

Note that $N|n,\mu_0$ is from negative binomial distribution with the expectation of N as

$$E(N)=\frac{n}{\mu_0}.$$

Substitute $E(N)$ with $\hat{N}$ derived in Section 2.2.1 , we have the estimated value of μ_0 as

$$\mu_0=\frac{n}{\hat{N}},$$

which is .0043 in our case. So far, we have obtained the posterior distribution of N.

3.2 The Sample-Complement Distribution

Next, we need to make inference about the non-sampled values.

Let $I_i=1$ if $i\in s$ and $I_i=0$ if $i\notin s$. Then

$$\begin{aligned}&I_i|x_i\sim Ber\left(\frac{x_i}{W}\right), and\ x_i\sim f(x_i)\\ &\Rightarrow P(I_i,x_i)\propto\left[\frac{x_i}{W}f(x_i)\right]^{I_i}\left[\left(1-\frac{x_i}{W}\right)f(x_i)\right]^{1-I_i}\\ &\Rightarrow P(x_i|I_i=0)=\frac{\left(1-\frac{x_i}{W}\right)f(x_i)}{\int\left(1-\frac{x_i}{W}\right)f(x_i)}.\end{aligned}$$

The density function of the non-sampled data is

$$\begin{aligned}&\Pi(x_{n+1},...,x_N|\alpha,\beta_1,\beta_2,\beta_3,\gamma,x_1,...,x_n,N)\\&=\prod_{i=n+1}^{N}\frac{\left(1-\frac{x_i}{W}\right)f(x_i)}{\int\left(1-\frac{x_i}{W}\right)f(x_i)}=\prod_{i=n+1}^{N}\left[\frac{1-\frac{x_i}{W}}{1-\frac{\mu}{W}}\right]f(x_i),\end{aligned}\tag{3.2}$$

where $f(x)$ is the density function of the population distribution, and μ is the expected value of x from that distribution. x_i is the sample from the sample-complement distribution. We can use the Sampling Importance Resampling (SIR) algorithm to do the sampling. The SIR algorithm is ideal because $\prod_{i=n+1}^{N} f(x_i)$ is a good approximation to the density function and it is easy to take draws.

We define the importance function as

$$\Pi_a(x_{n+1},\cdots,x_N|N)=\frac{\prod_{i=n+1}^{N}f(x_i)}{\int\prod_{i=n+1}^{N}f(x_i)dx_{n+1}\cdots dx_N}.\tag{3.3}$$

Then, the importance ratios are

$$\frac{\Pi(x_{n+1},\cdots,x_N|N)}{\Pi_a(x_{n+1},\cdots,x_N|N)}\propto\prod_{i=n+1}^{N}\frac{1-\frac{x_i}{W}}{1-\frac{\mu}{W}}.\tag{3.4}$$

A random sample can now be obtained by resampling with probability proportional to the ratios.

3.3 The Model Algorithm

The objective is to obtain the distribution of population mean.

- Step 1. Obtain M sets of $(\alpha, \beta_1, \beta_2, \beta_3, \gamma,)$ using the sampling methods described in the next Section.

- Step 2. Obtain a sample of N from formula (3.1).

- Step 3. For each set of parameters, generate the vector $\tilde{x_j}$ where $x_{ij}, i = n_j + 1, \cdots, (N_j - n_j),\ j = 1, 2, 3$ from the corresponding generalized gamma distribution.

- Step 4. Computing the population mean and the importance ratio w.

- Step 5. Repeat the step 2 to step 4 $M - 1$ times.

- Step 6. Draw αM values of the population mean with probability proportional to $\tilde{w}$. We choose $\alpha = .1$

3.4 The Population Distribution

In the last chapter, we use the maximum likelihood method to estimate the parameters of the population distribution. But estimation is troublesome for generalized gamma distributions. Parr and Webster (1965), Hager and Bain (1971), and Lawless (1980) have considered maximum likelihood estimation in the three-parameter generalized gamma distribution. They reported problems with iterative solution of the nonlinear equations implied by the maximum likelihood method. They remarked that maximum likelihood estimators might not exist unless the sample size exceeds 400. In this section, we perform Bayesian analysis of generalized gamma distributions. Two numerical methods are proposed, and the second one is shown to perform well.

3.4.1 Gibbs Sampler

We applied the Gibbs sampler at first because its idea is straightforward. Gibbs sampling is applicable when the joint distribution is difficult to sample from directly, but the conditional distribution of each variable is easier to sample from. The Gibbs sampling algorithm generates an instance from the distribution of each variable in turn, conditional on the current values of the other variables.

To use a Bayes approach, we assume the shrinkage prior for each parameter, that is

$$\Pi(\theta) = \frac{1}{(1+\theta)^2},$$

where $\theta = \alpha, \beta_1, \beta_2, \beta_3,$ andγ respectively.

The joint posterior density of $\alpha, \beta_1, \beta_2, \beta_3,$ andγ given the vector of size biased samples $x_{11}, \cdots, x_{3n_3}$ is

$$\begin{aligned}
&\Pi(\alpha, \beta_1, \beta_2, \beta_3, \gamma | x_{11}, \cdots, x_{3n_3}) \\
&= \frac{\gamma^n \left(\prod_{i=1}^{3} \prod_{j=1}^{n_i} x_{ij} \right)^{\gamma\alpha}}{(\beta_1^{n_1} \beta_2^{n_2} \beta_3^{n_3})^{\gamma\alpha+1} \left[\Gamma(\alpha + \frac{1}{\gamma}) \right]^n} \exp\left[- \sum_{i=1}^{3} \sum_{j=1}^{n_i} \left(\frac{x_{ij}}{\beta_i} \right)^{\gamma} \right] \\
&\cdot \frac{1}{(1+\alpha)^2} \frac{1}{(1+\beta_1)^2} \frac{1}{(1+\beta_2)^2} \frac{1}{(1+\beta_3)^2} \frac{1}{(1+\gamma)^2},
\end{aligned}$$

where n_1, n_2, n_3 are the number of shrubs in each transect respectively.

The conditional distributions of α is

$$\Pi(\alpha|\beta_1,\beta_2,\beta_3,\gamma,x_1,\cdots,x_n) \propto \frac{\left(\prod_{i=1}^{3}\prod_{j=1}^{n_i} x_{ij}\right)^{\gamma\alpha}}{(\beta_1^{n_1}\beta_2^{n_2}\beta_3^{n_3})^{\gamma\alpha}\left[\Gamma(\alpha+\frac{1}{\gamma})\right]^n} \cdot \frac{1}{(1+\alpha)^2},$$

$\alpha \in (0,\infty)$.

Since the α cannot be sampled from this form directly, we use grid method to do the sampling. Grid method is based on defining a grid in the parameter space and sampling the grid points with their corresponding probabilities. To apply grid method effectively, we need to transform the parameter space of α into a bounded space. The transformation $\alpha' = \frac{\alpha}{1+\alpha}$ would do the trick.

$$\Pi(\alpha'|\beta_1,\beta_2,\beta_3,\gamma,x_1,\cdots,x_n) \propto \left\{ \frac{\left(\prod_{i=1}^{3}\prod_{j=1}^{n_i} x_{ij}\right)^{\gamma\alpha}}{(\beta_1^{n_1}\beta_2^{n_2}\beta_3^{n_3})^{\gamma\alpha}\left[\Gamma(\alpha+\frac{1}{\gamma})\right]^n} \right\}_{\alpha=\frac{\alpha'}{1-\alpha'}}, \tag{3.5}$$

$\alpha' \in (0,1)$.

Apply the same strategy, we have the conditional distributions of $\beta_1', \beta_2', \beta_3'$, and γ' respectively.

$$\Pi(\beta_1'|\alpha,\beta_2,\beta_3,\gamma,x_1,\cdots,x_n) \propto \left\{ \frac{1}{\beta_1^{n_1(\gamma\alpha+1)}} \cdot \exp\left[-\sum_{j=1}^{n_1}\left(\frac{x_{1j}}{\beta_1}\right)^{\gamma}\right] \right\}_{\beta_1=\frac{\beta_1'}{1-\beta_1'}}, \tag{3.6}$$

$\beta_1' \in (0,1)$.

$$\Pi(\beta_2'|\alpha,\beta_1,\beta_3,\gamma,x_1,\cdots,x_n) \propto \left\{ \frac{1}{\beta_2^{n_2(\gamma\alpha+1)}} \cdot \exp\left[-\sum_{j=1}^{n_2} \left(\frac{x_{2j}}{\beta_2} \right)^{\gamma} \right] \right\}_{\beta_2=\frac{\beta_2'}{1-\beta_2'}}, \qquad (3.7)$$

$\beta_2' \in (0,1)$.

$$\Pi(\beta_3'|\alpha,\beta_1,\beta_2,\gamma,x_1,\cdots,x_n) \propto \left\{ \frac{1}{\beta_3^{n_3(\gamma\alpha+1)}} \cdot \exp\left[-\sum_{j=1}^{n_3} \left(\frac{x_{3j}}{\beta_3} \right)^{\gamma} \right] \right\}_{\beta_3=\frac{\beta_3'}{1-\beta_3'}}, \qquad (3.8)$$

$\beta_3' \in (0,1)$.

$$\Pi(\gamma'|\alpha,\beta_1,\beta_2,\beta_3,x_1,\cdots,x_n) \propto \left\{ \frac{\gamma^n \left(\prod_{i=1}^{3} \prod_{j=1}^{n_i} x_{ij} \right)^{\gamma\alpha}}{(\beta_1^{n_1}\beta_2^{n_2}\beta_3^{n_3})^{\gamma\alpha} \left[\Gamma(\alpha+\frac{1}{\gamma}) \right]^n} \exp\left[-\sum_{i=1}^{3} \sum_{j=1}^{n_i} \left(\frac{x_{ij}}{\beta_i} \right)^{\gamma} \right] \right\}_{\gamma=\frac{\gamma'}{1-\gamma'}}, \qquad (3.9)$$

$\gamma' \in (0,1)$.

Gibbs sampling is a Markovian updating scheme that proceeds as follows.

Given an arbitrary starting set of values $\beta_1^{(0)}, \beta_2^{(0)}, \beta_3^{(0)}, \gamma^{(0)}$, we draw $\alpha'^{(1)}$ from formula 3.5, then transform the value of $\alpha'^{(1)}$ to $\alpha^{(1)}$ by $\alpha = \frac{\alpha'}{1-\alpha'}$. Then draw $\beta_1'^{(1)}$ from 3.6 and transform to $\beta_1^{(1)}$, and so on up to $\gamma^{(1)}$ to complete one iteration of the scheme. After t such iterations we would arrive at a joint sample $\alpha^{(k)}, \beta_1^{(k)}, \beta_2^{(k)}, \beta_3^{(k)}, \gamma^{(k)}$. Then ignore some number of samples at the beginning (the so-called burn-in period), and

consider only every *mth* sample. The reason is that it may take a while for stationary distribution to be reached and successive samples are not independent of each other but form a Markov chain with some amount of correlation.

3.4.2 Computation Results using Gibbs Sampler

A total of 1000 sets of $(\alpha, \beta_1, \beta_2, \beta_3, \gamma)$ were sampled. After burn-in, selecting every 15th sample and resampling, 64 of them were left. The summary of the parameters and population mean were shown in Table 3.1.

The histogram of the population mean was in Fig. 3.1

We see most of estimated values of population mean are larger than what we expected. The problem occurs when some of the parameters having a wide range, for example, α ranging from 0.25 to 6.78 and β_3 ranging from 0.25 to 4.83. One of the reasons is that the high correlation between these parameters makes the Gibbs sampler inefficient in the sense it may take a very large number of iterations to converge in distribution.

Table 3.1: Summary of the Parameters and Population Mean

Name	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
α	0.25	0.28	1.62	2.86	5.36	6.78
β_1	0.25	0.25	0.40	0.55	0.84	1.25
β_2	0.25	0.25	0.62	0.82	1.33	1.91
β_3	0.25	0.25	0.53	0.92	1.33	4.83
γ	0.48	0.58	0.83	1.06	1.35	2.50
$\bar{x}$	0.066	0.088	3.44	13.92	26.01	58.86

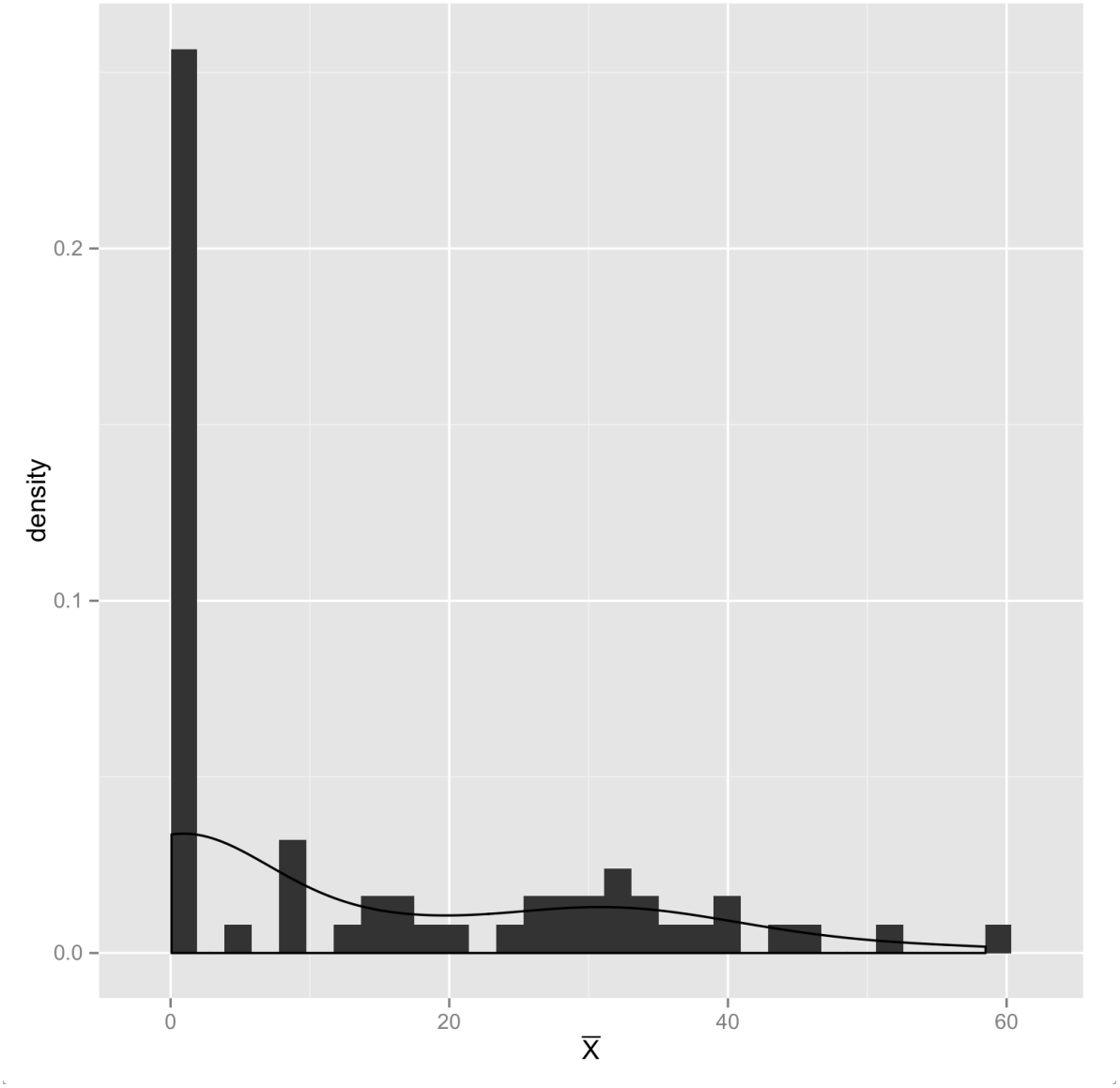

Fig. 3.1: Distribution of $\bar{x}$ using Gibbs Sampler

3.4.3 Random Sampler

In this method, we used the same shrinkage priors for α and γ, but employed another non-informative prior for $\beta_i, i = 1, 2, 3$.

$$\Pi(\beta_i) = \frac{1}{\beta_i}, i = 1, 2, 3.$$

To make the sampling procedure more efficient, we transformed the $\beta_i, i = 1, 2, 3$ by $\phi_i = \beta_i^{-\gamma}, i = 1, 2, 3$. Then the joint posterior distribution is given by

$$\begin{aligned}
&\Pi(\alpha, \phi_1, \phi_2, \phi_3, \gamma | x_{11} \cdots, x_{3n_3}) \\
&= \frac{\gamma^n \left(\prod_{i=1}^{3} \prod_{j=1}^{n_i} x_{ij} \right)^{\gamma\alpha}}{(\phi_1^{n_1} \phi_2^{n_2} \phi_3^{n_3})^{-(\alpha + \frac{1}{\gamma})} \left[\Gamma(\alpha + \frac{1}{\gamma}) \right]^n} \exp \left[- \sum_{i=1}^{3} \sum_{j=1}^{n_i} \phi_i x_{ij}^{\gamma} \right] \frac{1}{(1+\alpha)^2} \frac{1}{\phi_1} \frac{1}{\phi_2} \frac{1}{\phi_3} \frac{1}{\gamma^3 (1+\gamma)^2},
\end{aligned}$$

where n_1, n_2, n_3 are the number of shrubs in each transect respectively.

The posterior conditional distribution of $\phi_i, i = 1, 2, 3$, has the simple form

$$\begin{aligned}
&\Pi(\phi_i | \alpha, \phi_k, \gamma, x_{11}, \cdots, x_{3n_3}) \\
&\propto \phi_i^{n_i(\alpha + \frac{1}{\gamma}) - 1} \exp \left[-\phi_i \sum_{j=1}^{n_i} x_{ij}^{\gamma} \right] \\
&\sim \text{Gamma} \left(n_i(\alpha + \frac{1}{\gamma}), \sum_{j=1}^{n_i} x_{ij}^{\gamma} \right), i = 1, 2, 3; k \neq i. \qquad (3.10)
\end{aligned}$$

Formal integration with respect to $\beta_i, i = 1, 2, 3$ yields the marginal posterior distribution of α and γ .

$$\begin{aligned}
&\Pi(\alpha, \gamma | x_{11}, \cdots, x_{3n_3}) = \int_{\phi_1} \int_{\phi_2} \int_{\phi_3} \Pi(\alpha, \phi_1, \phi_2, \phi_3, \gamma | x_{11} \cdots, x_{3n_3}) d\phi_1 d\phi_2 d\phi_3 \\
&= \frac{\gamma^n \left(\prod_{i=1}^{3} \prod_{j=1}^{n_i} x_{ij} \right)^{\gamma\alpha}}{\left[\Gamma(\alpha + \frac{1}{\gamma}) \right]^n} \frac{\Gamma\left(n_1(\alpha + \frac{1}{\gamma}) \right) \Gamma\left(n_2(\alpha + \frac{1}{\gamma}) \right) \Gamma\left(n_3(\alpha + \frac{1}{\gamma}) \right)}{\left(\sum x_{1j}^{\gamma} \right)^{n_1(\alpha + \frac{1}{\gamma})} \left(\sum x_{2j}^{\gamma} \right)^{n_2(\alpha + \frac{1}{\gamma})} \left(\sum x_{3j}^{\gamma} \right)^{n_3(\alpha + \frac{1}{\gamma})}} \frac{1}{(1+\alpha)^2} \frac{1}{\gamma^3 (1+\gamma)^2},
\end{aligned}$$

$\alpha \in (0, \infty), \gamma \in (0, \infty)$.

α and γ cannot be sampled directly from their unbounded parameters space. The transformation $\alpha' = \frac{\alpha}{1+\alpha}$ and $\gamma' = \frac{\gamma}{1+\gamma}$ are required.

$$\Pi(\alpha',\gamma'|x_{11},\cdots,x_{3n_3}) = \int_{\phi_1}\int_{\phi_2}\int_{\phi_3}\Pi(\alpha,\phi_1,\phi_2,\phi_3,\gamma|x_{11}\cdots,x_{3n_3})d\phi_1 d\phi_2 d\phi_3$$

$$= \left\{ \frac{\gamma^n\left(\prod_{i=1}^{3}\prod_{j=1}^{n_i}x_{ij}\right)^{\gamma\alpha}}{\left[\Gamma(\alpha+\frac{1}{\gamma})\right]^n} \frac{\Gamma\left(n_1(\alpha+\frac{1}{\gamma})\right)\Gamma\left(n_2(\alpha+\frac{1}{\gamma})\right)\Gamma\left(n_3(\alpha+\frac{1}{\gamma})\right)}{\left(\sum x_{1j}^{\gamma}\right)^{n_1(\alpha+\frac{1}{\gamma})}\left(\sum x_{2j}^{\gamma}\right)^{n_2(\alpha+\frac{1}{\gamma})}\left(\sum x_{3j}^{\gamma}\right)^{n_3(\alpha+\frac{1}{\gamma})}\cdot\gamma^3} \right\}_{\alpha=\frac{\alpha'}{1-\alpha'},\gamma=\frac{\gamma'}{1-\gamma'}},$$

$\alpha' \in (0,1), \gamma' \in (0,1)$.

Two-dimensional grid method can be applied to draw α' and γ' from their joint distribution. But grid method is computationally intensive in more than one dimension. We used the Bayes' rule to draw samples of α' and γ'.

$$\Pi(\alpha',\gamma'|x_{11}\cdots,x_{3n_3}) = \Pi(\alpha'|\gamma',x_{11}\cdots,x_{3n_3})\Pi(\gamma'|x_{11}\cdots,x_{3n_3}). \tag{3.11}$$

To apply this rule, we fist generated a sample of $\gamma'^{(1)}$ from $\Pi(\gamma'|x_{11}\cdots,x_{3n_3})$, then generated a sample of $\alpha'^{(1)}$ from $\Pi(\alpha'|\gamma'^{(1)},x_{11}\cdots,x_{3n_3})$. Repeating this procedure M times to obtain M sets of $\alpha'(\alpha)$ and $\gamma'(\gamma)$. The corresponding $\phi(\beta)$ can also be obtained by sampling from $\Pi(\phi_i|\alpha,\phi_k,\gamma,x_{11},\cdots,x_{3n_3})$.

The first term of (3.12), $\Pi(\alpha'|\gamma',x_{11}\cdots,x_{3n_3})$ is easy to derive.

$$\Pi(\alpha'|\gamma', x_{11}\cdots, x_{3n_3})$$
$$\propto \left\{ \frac{\gamma^n \left(\prod_{i=1}^{3}\prod_{j=1}^{n_i} x_{ij}\right)^{\gamma\alpha}}{\left[\Gamma(\alpha+\frac{1}{\gamma})\right]^n} \frac{\Gamma\left(n_1(\alpha+\frac{1}{\gamma})\right)\Gamma\left(n_2(\alpha+\frac{1}{\gamma})\right)\Gamma\left(n_3(\alpha+\frac{1}{\gamma})\right)}{\left(\sum x_{1j}^{\gamma}\right)^{n_1(\alpha+\frac{1}{\gamma})}\left(\sum x_{2j}^{\gamma}\right)^{n_2(\alpha+\frac{1}{\gamma})}\left(\sum x_{3j}^{\gamma}\right)^{n_3(\alpha+\frac{1}{\gamma})}} \right\}_{\alpha=\frac{\alpha'}{1-\alpha'}},$$
$$\alpha' \in (0,1).$$

The second term $\Pi(\gamma'|x_{11}\cdots, x_{3n_3})$ can be derived by integrating $\Pi(\alpha', \gamma'|x_{11}\cdots, x_{3n_3})$ with respect to α'. Unfortunately, it is not possible to integrate $\Pi(\alpha', \gamma'|x_{11}\cdots, x_{3n_3})$ by analytical techniques. For this reason, numerical methods have to be used. We use the 20-point Gaussian quadrature to approximate $\Pi(\gamma'|x_{11}\cdots, x_{3n_3})$.

$$\begin{aligned}\Pi(\gamma'|x_{11}\cdots, x_{3n_3}) &= \int_0^1 \Pi(\alpha', \gamma'|x_{11}\cdots, x_{3n_3})d\alpha' \\ &= \frac{1}{2}\int_{-1}^{1} \Pi\left(\frac{1}{2}+\frac{1}{2}\alpha', \gamma'\right)d\alpha' \\ &\approx \frac{1}{2}\sum_{i=1}^{20}\omega_i \Pi\left(\frac{1}{2}+\frac{1}{2}x_i, \gamma'\right),\end{aligned}$$

where $x_i, i = 1, \cdots, 20$ are the roots of orthogonal polynomials $P_{20}(x)$ for $[-1, 1]$ and $\omega_i, i = 1, \cdots, 20$ are the corresponding Gauss-Legendra weights.

The Laguerre quadrature rules can be created by R package gaussquad.

3.4.4 Computation Results using Random Sampler

The summary for each parameters and the population mean are shown in Table 3.2. Their distributions are in Fig. 3.2 to Fig. 3.7. We see the population mean has the

range of $(.45, .87)$, the median is .74, which is reasonable.

In the next section, we will perform the model checking by Conditional Predictive Ordinate.

Table 3.2: Summary of the Parameters and Population Mean

Name	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
α	0.25	0.77	1.37	1.34	1.93	2.33
β_1	0.07	0.29	0.46	0.52	0.71	1.50
β_2	0.17	0.49	0.75	0.83	1.12	2.19
β_3	0.13	0.56	0.87	0.96	1.29	2.98
γ	0.64	1.05	1.36	1.43	1.68	3.54
$\bar{X}$	0.31	0.67	0.75	0.74	0.81	1.01

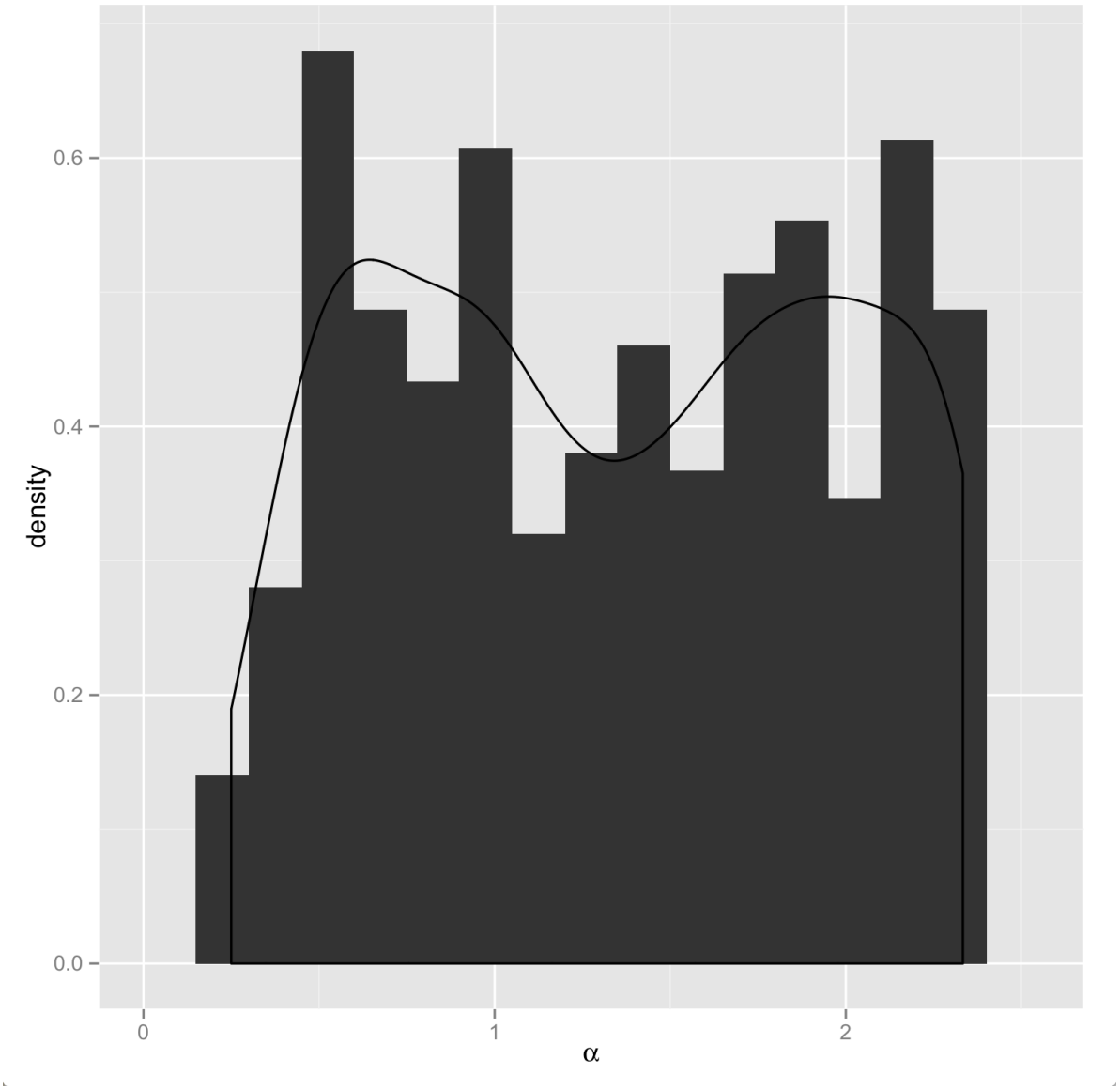

Fig. 3.2: Distribution of α using Random Sampling

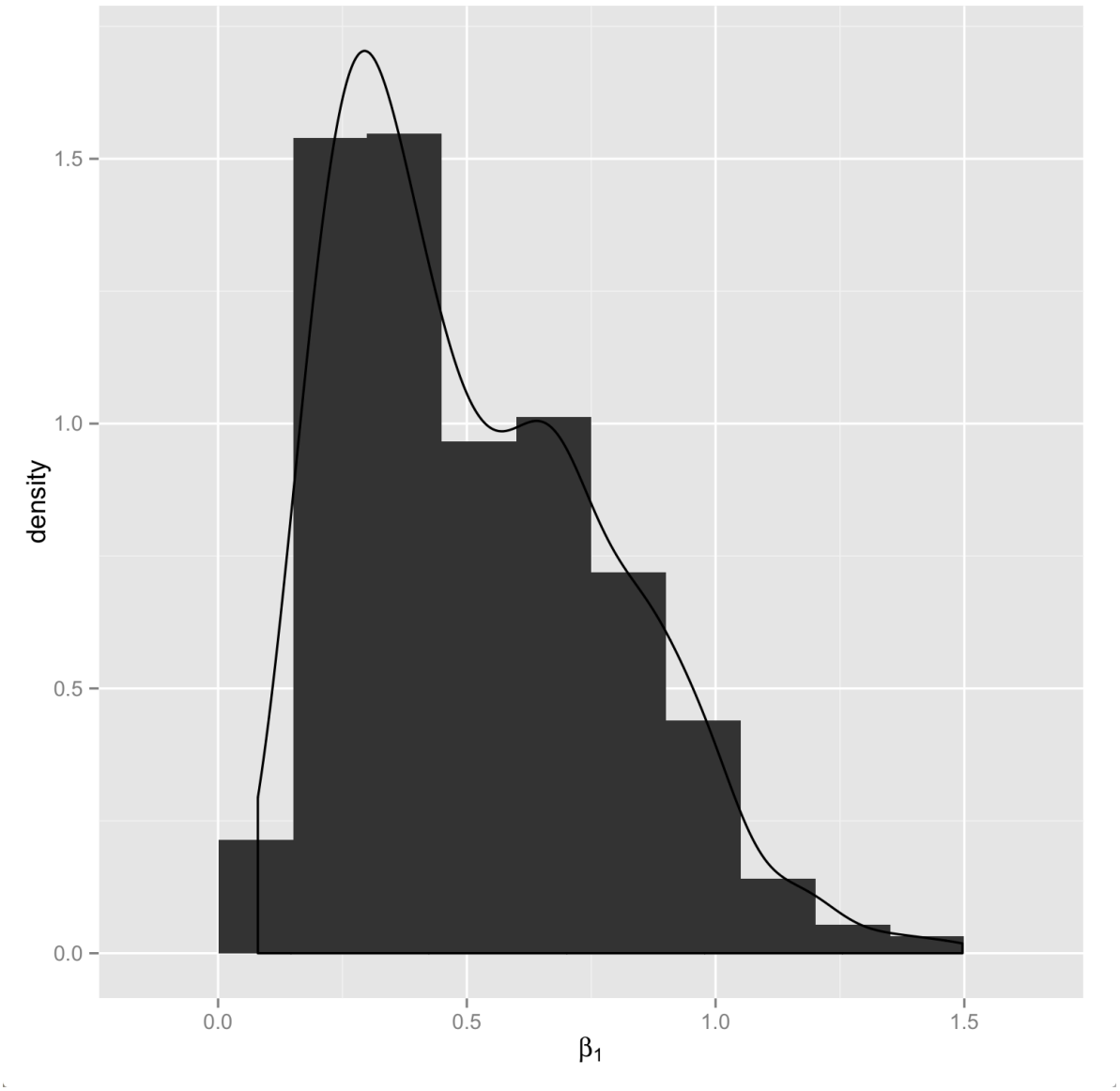

Fig. 3.3: Distribution of β_1 using Random Sampling

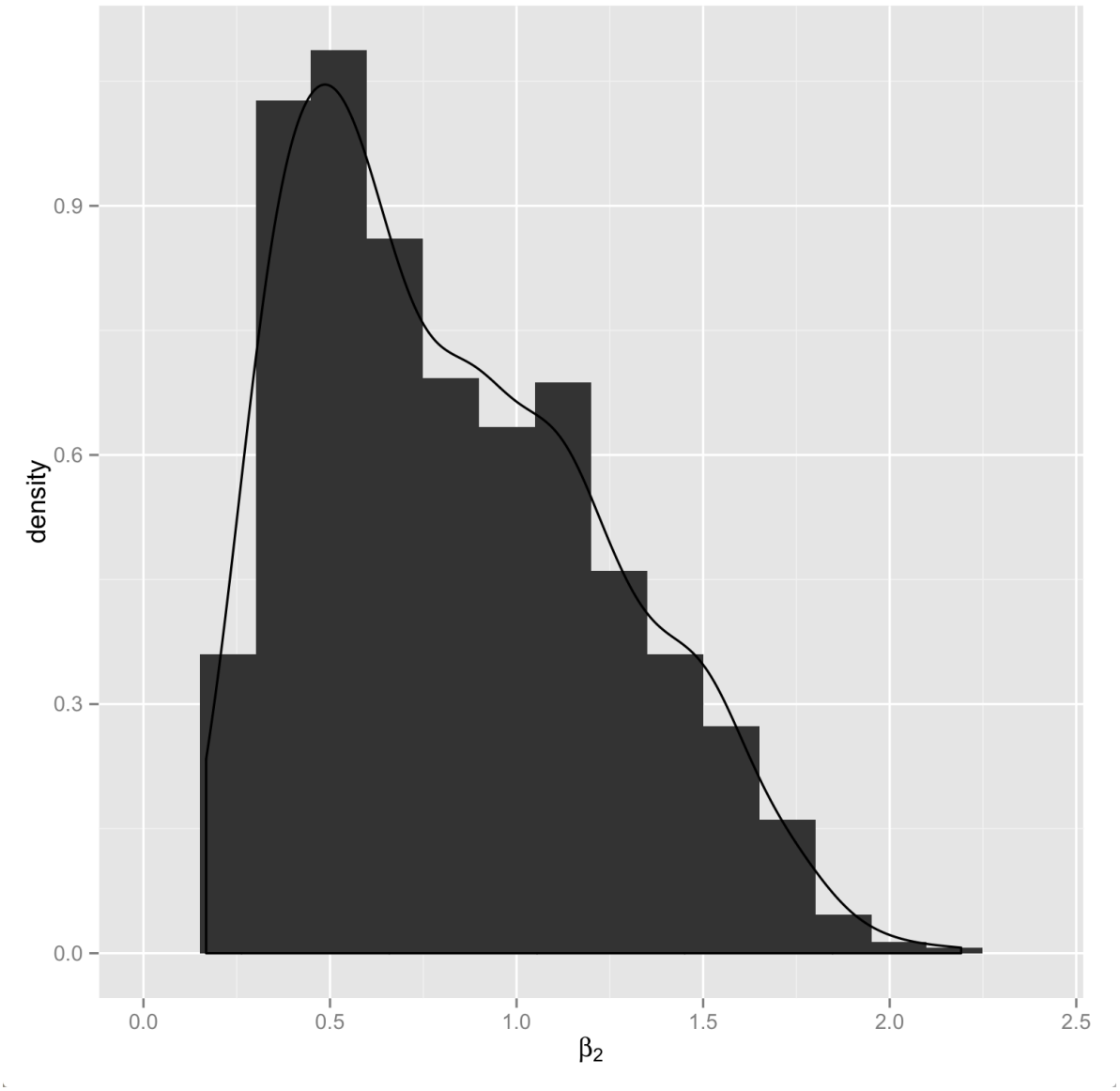

Fig. 3.4: Distribution of β_2 using Random Sampling

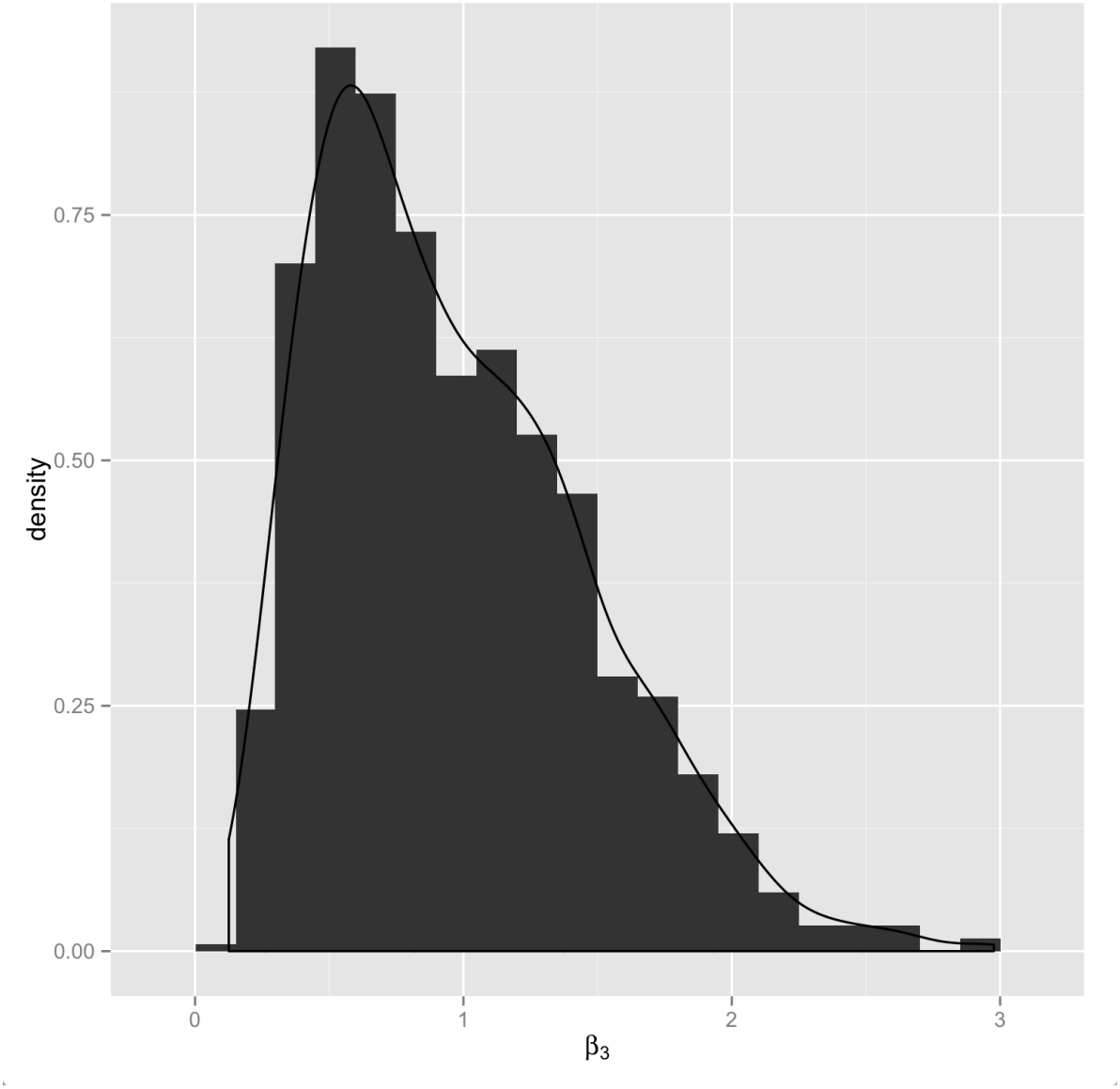

Fig. 3.5: Distribution of β_3 using Random Sampling

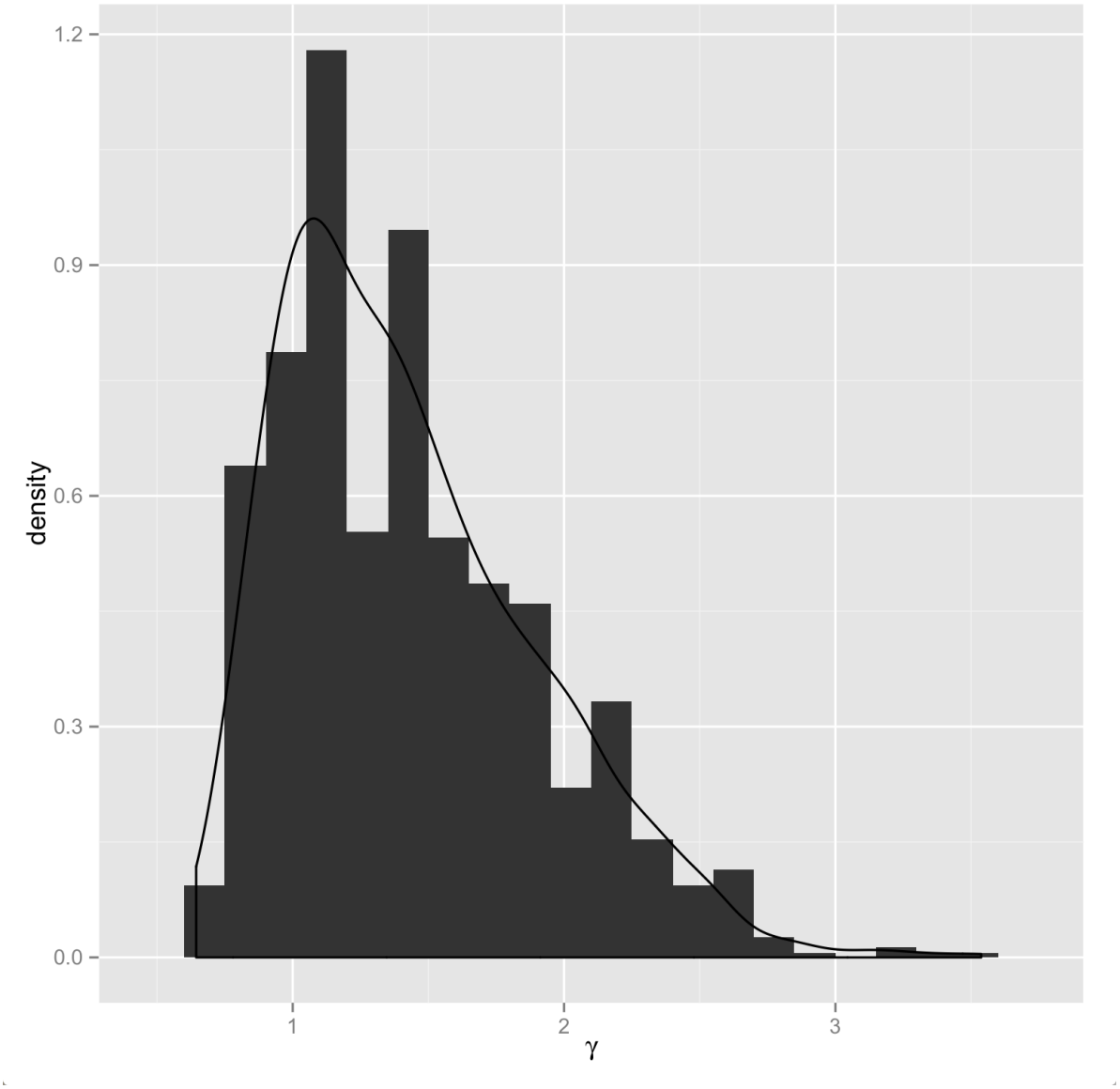

Fig. 3.6: Distribution of γ using Random Sampling

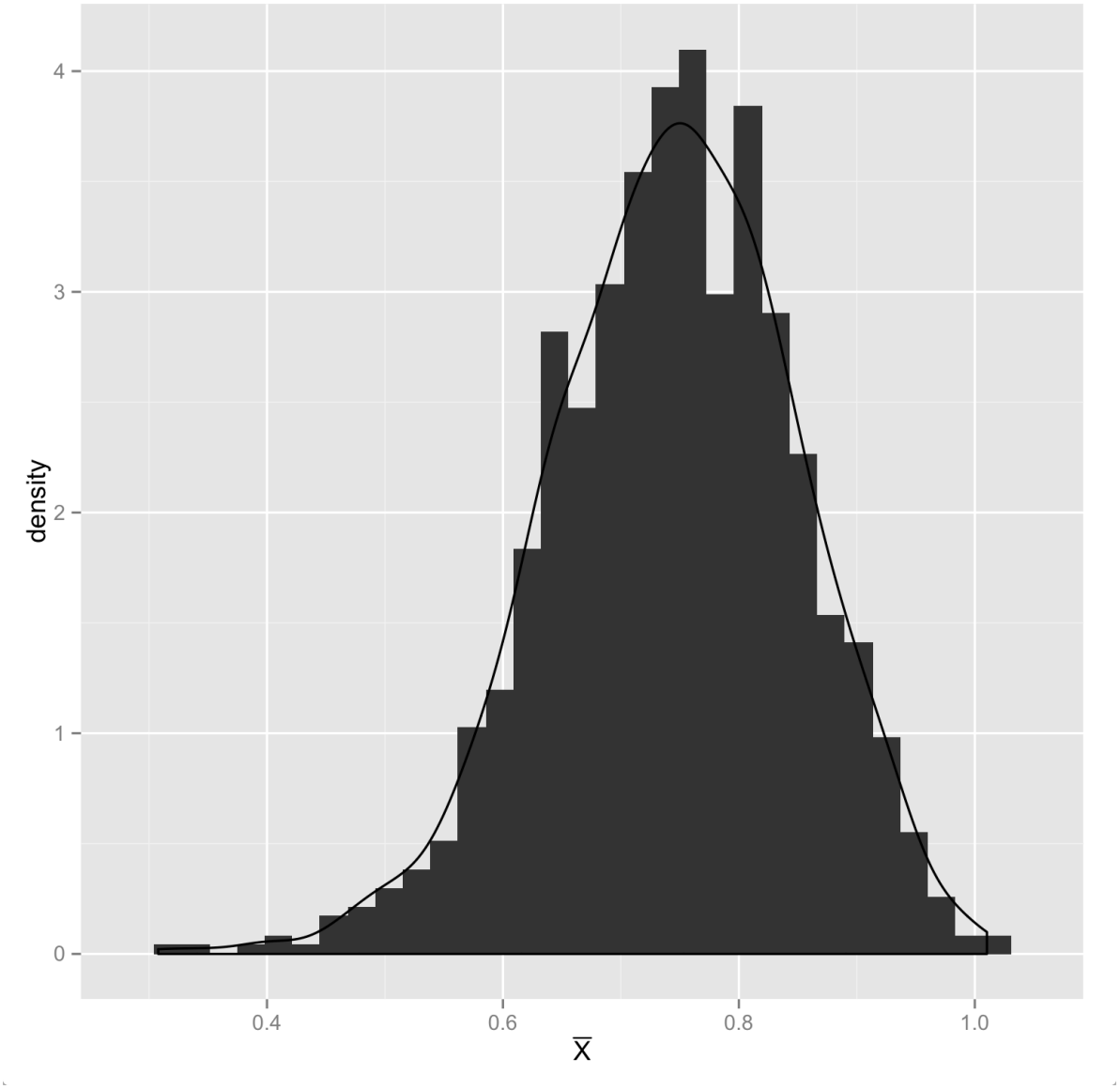

Fig. 3.7: Distribution of $\bar{X}$ using Random Sampling

3.5 Model Checking by Conditional Predictive Ordinate

Comparing the predictive distribution to the observed data is generally termed a "posterior predictive check". This type of check includes the uncertainty associated with the

estimated parameters of the model. Posterior predictive checks (via the predictive distribution) involve a double-use of the data, which causes predictive performance to be overestimated. To overcome this drawbacks, Geisser and Eddy (1979) has proposed the leave-one-out cross-validation predictive density. This is also known as the Conditional Predictive Ordinate or CPO (Gelfand, 1996).

The CPO is a handy posterior predictive check because it may be used to identify outliers, influential observations, and for hypothesis testing across different nonnested models.The CPO expresses the posterior probability of observing the value of x_i when the model is fitted to all data except x_i, with a larger value implying a better fit of the model to x_i, and very low CPO values suggest that x_i is an outlier and an influential observation.

A Monte Carlo estimate of the CPO is obtained without actually omitting x_i from the estimation, and is provided by the harmonic mean of the likelihood for x_i. Specifically, the CPO_i is the inverse of the posterior mean of the inverse likelihood of x_i.

The Monte Carlo estimate of the CPO is

$$\widehat{CPO_i} = \left[\frac{1}{M} \sum_{h=1}^{M} \frac{1}{f(x_i|\tilde{\theta}^{(h)})} \right]^{-1}, i = 1, 2, ..., n$$

$$\tilde{\theta}^{(h)} \overset{iid}{\sim} \Pi(\tilde{\theta}|\tilde{x}).$$

The sum of the logged CPOs can be an estimator for the logarithm of the marginal likelihood, sometimes called the log pseudo marginal likelihood (LPML)

$$LPML = \sum_{i=1}^{n} log(\widehat{CPO_i})$$

Models with larger LPMLs are better.

To compare the predictive distributions (both model with selection bias and model without selection bias) to our size-biased sample. we calculated the LPML for both models.

The likelihood of x_i under selection biased model is given by

$$f(x_i|\alpha,\beta,\gamma) = \frac{\gamma x_i^{\gamma\alpha}}{\beta^{\gamma\alpha+1}\Gamma(\alpha+\frac{1}{\gamma})} \exp\left[-\left(\frac{x_i}{\beta}\right)^{\gamma}\right],$$

where β is the corresponding parameter for the strata that x_i from.

The likelihood of x_i under no selection biased model is given by

$$f(x_i|\alpha,\beta,\gamma) = \frac{\gamma x_i^{\gamma\alpha-1}}{\beta^{\gamma\alpha}\Gamma(\alpha)} exp\left[-\left(\frac{x_i}{\beta}\right)^{\gamma}\right],$$

where β is the corresponding parameter for the strata that x_i from.

Table 3.3: $LPMLs$ for the two model

Model	$LPML$
Model With Selection Bias	-36.10
Model Without Selection Bias	-47.54

We see the $LPML$ of for model with selection bias is larger than the one for the model without selection bias, which means the model with selection bias fits our size-biased sample better.

Chapter 4

Summary

In this paper we have presented models for estimating population mean under size-biased sampling. We have used the three-parameter generalized gamma distribution to model the shrub widths. We have extended the generalized gamma distribution to accommodate length bias. Our interest is the finite population mean of shrub width in the entire quarry.

Both classical inference and Bayesian analysis have been attempted. Classical inference presents certain technical problems when the number of the parameters is large. Although analytical inference is not possible, numerical Bayesian inference can be conducted using random sampler. Posterior population distribution can be easily estimated using this method. Conditional predictive ordinate shows that the model with selection bias performs better than the model without selection bias.

An interesting topic for future research would be including covariates to study potential predictors. In Muttlak (1988), in addition to the measurement of shrub widths (Width), two more attributes of mountain mahogany, maximum height (Height), and number of stems (Stem), were measured. The data are presented in Table A.1. Both attributes are important predictors of the average shrub width of an area's vegetation. Semiparametric linear regression (Chen, 2010) or generalized linear regression can be considered to measure this association.

We can incorporate the covariates through a gamma type regression model. Let the

covariates be $\underset{\sim}{z}_{ij}, i = 1, 2, 3, j = n_1, n_2, n_3$ and $\underset{\sim}{\phi}_i, i = 1, 2, 3$. Because the mean of each stratum is linearly related to $\beta_1, \beta_2, \beta_3$ respectively, we take $\beta_i = e^{\underset{\sim}{z_i}' \underset{\sim}{\phi_i}}, i = 1, 2, 3$.

For the shrub data, our model is

$$P(\underset{\sim}{z}|\underset{\sim}{\phi}, \alpha, \gamma) = \prod_{i=1}^{3} \prod_{j=1}^{n_i} \frac{\gamma x_{ij}^{\gamma\alpha-1} \left[e^{-z_{ij}\phi_i}\right]^{\gamma\alpha}}{\Gamma(\alpha)} \exp\left\{-\left(x_{ij} e^{-z_{ij}\phi_i}\right)^{\gamma}\right\}.$$

A similar form can be easily written down for the size biased sampling. Our future plan is to fit a model to accommodate the covariats.

Appendix A

Tables

Table A.1: Data with covariates for the line intercept sampling method

Repli.	Trans	Number	Inter.	Width	Height	Stems
I	1	1	2.68	1.53	1.70	19
I	1	2	2.34	0.87	0.83	6
I	1	3	1.93	0.79	1.10	9
I	1	4	0.53	0.78	1.04	5
I	1	5	0.50	1.85	1.41	17
I	1	6	0.24	1.45	0.65	28
I	1	7	0.43	0.48	0.28	23
I	1	8	0.13	0.52	0.42	19
I	1	9	0.19	0.22	0.24	4
I	1	10	0.43	0.38	0.42	24
I	1	11	0.74	0.59	0.31	16
I	1	12	0.72	0.20	0.34	4
I	1	13	0.39	0.42	0.29	14
I	1	14	0.40	1.02	0.66	26
I	1	15	0.40	0.97	0.9	26
I	1	16	–	0.56	0.51	12
I	1	17	–	0.62	0.26	22

Continued on next page

Table A.1 – continued from previous page

Repli.	Trans	Number	Inter.	Width	Height	Stems
I	1	18	–	0.42	0.48	5
I	2	1	1.90	1.15	1.20	37
I	2	2	2.63	0.87	0.80	11
I	2	3	0.60	0.57	0.40	19
I	2	4	0.80	0.97	0.65	30
I	2	5	0.40	0.57	0.50	9
I	2	6	1.82	1.97	1.35	61
I	2	7	0.94	0.58	0.93	16
I	2	8	0.50	2.54	1.20	31
I	2	9	1.50	1.85	0.95	46
I	2	10	0.39	0.35	0.33	19
I	2	11	1.18	1.24	0.96	19
I	2	12	1.68	1.80	1.46	67
I	2	13	0.78	0.78	0.80	27
I	2	14	0.69	0.98	1.07	18
I	2	15	0.84	1.30	1.47	16
I	2	16	1.00	1.55	0.96	21
I	2	17	1.04	1.69	1.05	39
I	2	18	1.77	2.12	1.43	69
I	2	19	1.14	1.27	1.15	26
I	2	20	0.28	0.75	0.72	10
I	2	21	0.43	1.01	1.22	22
I	2	22	–	1.82	1.15	33

Continued on next page

Table A.1 – continued from previous page

Repli.	Trans	Number	Inter.	Width	Height	Stems
I	3	1	0.20	0.71	0.50	21
I	3	2	2.46	1.50	1.20	74
I	3	3	2.68	1.82	1.93	28
I	3	4	–	1.86	1.32	30
I	3	5	–	1.61	0.72	18
I	3	6	–	1.21	1.20	37
II	1	1	0.52	0.67	0.68	9
II	1	2	0.16	0.31	0.26	4
II	1	3	0.95	0.83	0.70	16
II	1	4	0.59	1.95	1.20	44
II	1	5	1.12	1.36	1.36	30
II	1	6	0.62	1.45	1.19	11
II	1	7	0.44	0.72	0.58	15
II	1	8	0.56	1.15	0.81	20
II	1	9	0.86	0.98	0.39	3
II	1	10	0.31	1.29	1.14	21
II	1	11	0.15	0.88	0.78	23
II	1	12	0.27	0.25	0.83	23
II	1	13	0.53	0.63	0.39	12
II	1	14	0.34	1.12	0.87	20
II	1	15	0.38	0.34	0.06	4
II	1	16	1.40	0.21	0.41	5
II	1	17	1.42	1.36	1.13	18

Continued on next page

Table A.1 – continued from previous page

Repli.	Trans	Number	Inter.	Width	Height	Stems
II	1	18	0.30	0.95	0.68	9
II	1	19	1.37	1.04	0.82	22
II	1	20	0.03	0.48	0.42	11
II	1	21	1.20	1.05	0.98	12
II	1	22	1.06	0.88	0.98	9
II	1	23	0.33	0.16	0.11	6
II	1	24	0.39	1.08	1.19	31
II	1	25	1.43	0.95	1.26	22
II	1	26	0.48	0.25	1.04	7
II	1	27	0.86	0.30	0.97	4
II	1	28	1.22	1.40	1.20	29
II	1	29	0.43	0.58	0.89	11
II	1	30	–	0.73	1.33	24
II	1	31	–	1.30	1.26	14
II	1	32	–	0.57	0.55	14
II	2	1	1.30	0.96	0.93	16
II	2	2	1.75	2.08	1.43	25
II	2	3	1.59	0.68	0.79	10
II	2	4	1.52	1.39	0.86	38
II	2	5	0.47	0.5	0.58	6
II	2	6	0.04	0.72	0.63	16
II	2	7	–	0.19	0.43	13
II	2	8	–	1.91	1.21	24

Continued on next page

Table A.1 – continued from previous page

Repli.	Trans	Number	Inter.	Width	Height	Stems
II	2	9	–	0.88	0.48	13
II	2	10	–	0.48	0.59	19
II	2	11	–	0.12	0.14	1
II	3	0	–	–	–	–

Bibliography

[1] S. A. Butler and L. L. McDonald, *Unbiased systematic sampling plans for the line intercept method*, Jouran of Range Management **36(4)** (1983), 463–468.

[2] Y. Q. Chen, *Semiparametric regression in size-biased sampling*, Biometrics **66** (2010), 149–158.

[3] W. G. Cochran, *Sampling techniques*, Wiley (1977).

[4] L.L. Eberhnrdt, *Transect methods for population studies*, J. Wildl. Manage **42** (1978), 1–31.

[5] R. A. Fisher, *The effects of methods of ascertainment upon the estimation of frequencies*, Annals of Eugenics 6 (1934), 13–25.

[6] S. Geisser and W. F. Eddy, *A predictive approach to model selection*, Journal of the American Statistical Association **74** (1979), 153–160.

[7] A. E. Gelfand, *Model determination using sampling-based methods*, Markov Chain Monte Carlo in Practice (1996), 145–161.

[8] H. W. Hager and L. J. Bain, *Reliability estimation for the generalized gamma distribution and robustness of the weibull model*, Technometrics **13** (1971), 547–557.

[9] M. M. Hansen and W. N. Hurwitz, *On the theory of sampling from finite populations*, Annals of Mathematical Statistics **14** (1943), 333–362.

[10] H. Hirose, *Maximum likelihood parameters estimation by model augmentation with application to the extended four-parameters generalized gamma distribution*, Department of control Engineering and Science Kyushu Institute of technology, Fukuoka (1999), 820–8502.

[11] T. Hwang and P. Huang, *On new moment estimation of parameters of the generalized gamma distribution using its characterization*, Taiwanese journal of Mathematics **10** (2006), 1083–1093.

[12] M. Khodabin and A. Ahmadabadi, *Some properties of generalized gamma distribution*, Mathematical Sciences **4** (2010), 9–28.

[13] J. F. Lawless, *Inference in the generalized gamma and log gamma distributions*, Technometrics, **22** (1980), 409–419.

[14] H. A. Lucas and G. F. Seber, *Estimating coverage and particle density using the line intercept method*, Biometrika **64** (1977), 618–622.

[15] L. L. McDonald, *Line-intercept sampling for attributes other than coverage and density*, J. Wildl. Manage **44** (1980), 530–533.

[16] H. A. Muttlak and L. L. McDonald, *Ranked set sampling with size-biased probability of selection*, Biometrics **46** (1990), 435–445.

[17] B. Nandram, *Bayesian predictive inference of a finite population proportion under selection bias*, Statistical Methodology **11** (2013), 1–21.

[18] V. B. Parr and J. T. Webster, *A method for discriminating between failure density functions used in reliability predictions*, Technometrics **7** (1965), 1–10.

[19] G. P. Patil and C. R. Rao, *Weighted distributions and size-biased sampling with applications to wildlife populations and human families*, BIOMETRICS **34** (1978), 179–189.

[20] R. L. Prentice, *A log gamma model and its maximum likelihood estimation*, Biometrika **61** (1974), 539–544.

[21] C. R. Rao, *On discretedipopulations arising out of mmethod of ascertainment*, Pergamon Press and Statistical Publishing Society (1965), 320–332.

[22] E. W. Stacy, *A generalization of the gamma distribution*, The Annals of Mathematical Statistics **33** (1962), 1187–1192.