Statistica Sinica **31** (2021), 1779-1805
doi:https://doi.org/10.5705/ss.202018.0063

A UNIFIED THEORY FOR ROBUST BAYESIAN PREDICTION UNDER A GENERAL CLASS OF REGRET LOSS FUNCTIONS

Nader Nematollahi, Mohammad Jafari Jozani and Razieh Jafaraghaie

Allameh Tabataba'i University, University of Manitoba and Islamic Azad University, Shoushtar Branch

Abstract: We study robust Bayesian prediction problems using the posterior regret Γ-minimax (PRGM) approach. We provide a unified theory for PRGM prediction under a very general class of regret loss functions that includes the squared error, linear-exponential, entropy and many other loss functions as special cases. We apply our results to the problem of predicting unknown parameters for finite populations under different superpopulation models (normal and non-normal, with or without auxiliary variables) and several classes of prior distributions, including the commonly used ϵ-contaminated class of priors. Our results are augmented with real-world applications and simulation studies.

Key words and phrases: Bayes predictor, finite population, posterior regret Γ-minimax, robust Bayesian analysis.

1. Introduction

The Bayesian approach provides an attractive methodology for inferences about population parameters, and allows for prior information about the underlying problem to be incorporated in the analysis through the prior distribution. Perhaps the main barrier in using the Bayesian approach is the subjectivity involved in choosing a single and completely specified prior distribution for the parameter of interest. In other words, a practitioner who produces subjective Bayesian estimates might be vulnerable to criticism, as in the case of the sampler who uses a purposive sampling plan (Little (2004)). In general, there is no single method for choosing a prior distribution. Thus, different users may produce different priors, and therefore arrive at different posteriors and conclusions. In some situations, one might choose a family of prior distributions that depends on some unknown hyperparameters. However, as shown in Ghosh and Kim (1993), even in simple examples, failing to specify the correct

Corresponding author: Nader Nematollahi, Department of Statistics, Allameh Tabataba'i University, Tehran, Iran. E-mail: nematollahi@atu.ac.ir.

values of one or more of hyperparameters might have serious consequences, from a Bayesian viewpoint. To address this issue, one solution is to use a robust Bayesian approach by choosing a class Γ of plausible prior distributions for the parameter of interest, obtaining Bayesian solutions that are relatively insensitive to the uncertainty in determining the prior distribution. This is a common practice when the underlying problem must be solved by two or more decision-makers (sources) who do not necessarily agree on which prior distribution to use. In addition, because any selected prior distribution is just an approximation of the true and unknown prior distribution, sometimes it is better to choose a wide class of prior distributions, rather than just one (Berger (1984)). A robust Bayesian analysis yields a range of Bayes estimators, and one needs to specify which of these is most appropriate. In other words, it is interesting not only to investigate the range of estimators, but also to construct optimal procedures. Several methods can be used for this purpose, including the Γ-minimax (GM) (e.g., Berger (1984)), conditional Γ-minimax (CGM) (e.g., Betro and Ruggeri (1992)), most stable (MS) (e.g., Meczarski and Zielinski (1991)), and posterior regret Γ-minimax (PRGM) (e.g., Insua, Ruggeri and Vidakovic (1995); Boratynska (2006); Jafari Jozani and Parsian (2008)) approaches, the last of which minimizes the maximal posterior regret in specifying the optimum robust Bayes estimator/predictor.

We consider the problem of a robust Bayesian prediction using the PRGM approach under a very general class of loss functions and different classes of prior distributions. We gear our methodology toward predicting unknown parameters of finite population. However, our results are equally valid for other problems, such as PRGM estimations in parametric inferences based on infinite populations. To this end, we consider a finite population and adopt a model-based approach that views the finite population as a sample from a superpopulation (parametric) model, which depends on some unknown parameters. A sample of size n is selected from this population, and the goal is to predict characteristics of the underlying population, such as the population mean, total, variance, and so on. Bayesian predictions of the finite population total and mean within the class of unbiased and linear unbiased predictors have been studied by Godambe and Joshi (1965) and Godambe (1955), respectively. More information about Bayesian inferences for finite populations can be found in Hill (1968), Ericson (1969), Bolfarine (1990), Hamner, Seaman and Young (2001), Liu and Rong (2007), Kim and Saleh (2008), Ghosh (2008), Pefeffermann and Rao (2009), Chen, Elliott and Little (2012), and Si, Pillai, and Gelman (2015), and the references there in. The need for a robust Bayesian analysis in survey sampling has been recognized by many authors, including Godambe and Thompson (1971), Ghosh and Kim (1993), Ghosh (2008), and Zangeneh and Little (2015). In the context of a robust Bayesian approach in a finite population, Ghosh and Kim (1993) use ML-II priors to obtain a robust Bayes predictor of the finite population mean based on the posterior risk and robustness procedure suggested by Berger (1984), and study its performance over a class of prior distributions under the squared error (SE) loss function. To the best of our knowledge, no previous works have examined PRGM predictions of the finite population parameters.

The reminder of the paper is organized as follows. In Section 2, we give some definitions and preliminary results, and define the general class of loss functions that will be used throughout the paper. In Section 3, we obtain our main result on the PRGM prediction of unknown parametric functions under general classes of loss functions and prior distributions. In Section 4, we apply our method to obtain PRGM predictors of the finite population mean and/or variance under the SE and linear exponential (LINEX) loss functions and some classes of prior distributions for normal and non-normal superpopulation models. In Section 5, we provide real-world examples to predict finite population parameters such as the mean and variance under normal and non-normal models. In addition, we compare the estimated risk and bias of the PRGM and Bayes predictors under the SE loss function using simulations. Lastly, Section 6 concludes the paper. Some proofs and details of the derivations in the examples, as well as numerical results for two real-data applications, are presented in the Supplementary Material.

2. Preliminaries

In this section, we give some notation and preliminary results, which are used throughout the paper. Consider a finite population of N units, denoted by an index set $\mathcal{U} = \{1, 2, \ldots, N\}$, and suppose $\mathbf{y} = (y_1, \ldots, y_N)^\top$ is the vector of unknown values associated with a characteristic of interest y. From $\mathcal{U}$, a sample $s = \{i_1, \ldots, i_n\}$ of size $n(s) = n$ is selected using the simple random sampling without replacement (SRSWOR) method. A typical sample point is then the set of labels of units contained in the observed sample, along with $\mathbf{y}(s) = (s, \mathbf{y}_s) = \big(s, (y_{i_1}, \ldots, y_{i_n})^\top\big)$, where y_{i_j} is the observed value of the characteristic of interest for unit i_j selected in the sample. Using $\mathbf{y}_s$, we are interested in making inferences about some unknown finite population quantities, $\gamma(\mathbf{y})$, such as the population mean $\gamma_1(\mathbf{y}) = (1/N)\sum_{i=1}^N y_i = \bar{Y}$ and population variance $\gamma_2(\mathbf{y}) = (1/N)\sum_{i=1}^N (y_i - \bar{Y})^2$.

We adopt a model-based approach (e.g., Little (2004)), where the finite population $\mathbf{y}$ is assumed to be a realization from a superpopulation model $f(\mathbf{y}|\theta)$

that depends on some unknown parameters θ Thus, making inferences about $\gamma(\mathbf{y})$ reduces to the problem of predicting the outcomes of the non-sampled units $\mathbf{y}_{\bar{s}} = \{y_i : i \notin s\}$. We follow a robust Bayesian methodology, where one specifies a class Γ of prior distributions for the superpopulation parameters θ. The posterior predictive density, $h(\gamma(\mathbf{y})|\mathbf{y_s})$, the posterior distribution of $\gamma(\mathbf{y})$ given the observed data $\mathbf{y}_s$, is the basis for the inference about $\gamma(\mathbf{y})$. This is a general setting that covers many situations not necessarily restricted to a finite-population context. For example, to predict a future observation in its usual parametric setting, we have a sequence of random variables $y_1, y_2, \ldots, y_n$, and we want to predict the future random variable y_{n+1}. Here, $\mathbf{y}_s = (y_1, \ldots, y_n)^\top$, $\mathbf{y}_{\bar{s}} = y_{n+1}$ and $\gamma(\mathbf{y}) = \mathbf{y}_{\bar{s}} = y_{n+1}$. Therefore, in each case, our goal is to predict some functions $\gamma(\mathbf{y})$ using the predictors $\delta(\mathbf{y}_s)$.

Let $L(\gamma(\mathbf{y}), \delta(\mathbf{y}_s))$ be the loss function for predicting $\gamma(\mathbf{y})$ using $\delta(\mathbf{y}_s)$. To obtain a robust Bayesian predictor of $\gamma(\mathbf{y})$, one needs to calculate the posterior risk of $\delta = \delta(\mathbf{y}_s)$ for a given prior $\pi \in \Gamma$, as follows:

$$\rho(\pi, \delta) = E[L(\gamma(\mathbf{y}), \delta(\mathbf{y}_s))|\mathbf{y}_s]. \tag{2.1}$$

Then, the posterior risk (2.1) is minimized over the class of all possible predictors $\mathcal{D}$ when the prior distribution is also changing in Γ. One may also attempt to determine an optimal predictor by minimizing measures such as the maximal posterior regret, defined below (e.g., Berger (1990)).

Definition 1. δ_Γ^{PRGM} is a PRGM predictor of $\gamma(\mathbf{y})$ if $\sup_{\pi\in\Gamma} R(\delta_\Gamma^{PRGM}, \delta^\pi) = \inf_{\delta\in\mathcal{D}} \sup_{\pi\in\Gamma} R(\delta, \delta^\pi)$, where

$$R(\delta, \delta^\pi) = \rho(\pi, \delta) - \rho(\pi, \delta^\pi) \tag{2.2}$$

is the posterior regret due to the loss of optimality caused by using δ instead of the Bayes predictor δ^π.

In order to obtain robust Bayes predictors of $\gamma(\mathbf{y})$, one needs to specify a loss function $L(\gamma(\mathbf{y}), \delta(\mathbf{y}_s))$ to measure the error made in predicting $\gamma(\mathbf{y})$ using $\delta(\mathbf{y}_s)$. We introduce a general class of loss functions $L(\gamma(\mathbf{y}), \delta)$, where L is assumed to be a strictly bowl-shaped (BS) function of both $\gamma(\mathbf{y})$ and δ, with a unique minimum at $\delta = \gamma(\mathbf{y})$, that satisfies some additional conditions. Note that $f(t)$ is called a strictly BS function on its domain if, as a function of t, it first decreases and then increases, with a unique minimum at t_0. In other words, $f'(t) < 0$ for all $t < t_0$, and $f'(t) > 0$ for $t > t_0$. Obviously, any convex loss function is a BS function.

Table 1. Examples of regret loss functions with associated PRGM predictors, where $\underline{\delta} = \inf_{\pi\in\Gamma} \delta^\pi$ and $\overline{\delta} = \sup_{\pi\in\Gamma} \delta^\pi$, with δ^π being the Bayes predictor w.r.t. the underlying loss function.

Loss function	$L(\gamma(\mathbf{y}),\delta)$	PRGM predictor
Squared Error Loss (SEL)	$(\delta-\gamma(\mathbf{y}))^2$	$\frac{\underline{\delta}+\overline{\delta}}{2}$ Zen and DasGupta (1993)
Linear Exponential Loss (LINEX)	$b\{e^{c(\delta-\gamma(\mathbf{y}))}-c(\delta-\gamma(\mathbf{y}))\}-1$ $c\neq 0,\ b>0$	$\frac{-1}{c}\log\frac{e^{-c\overline{\delta}}-e^{-c\underline{\delta}}}{-c(\overline{\delta}-\underline{\delta})}$ Boratynska and Drozdowicz (1999)
Entropy Loss (EL)	$\frac{\gamma(\mathbf{y})}{\delta}-\ln\frac{\gamma(\mathbf{y})}{\delta}-1$	$\frac{\overline{\delta}-\underline{\delta}}{\ln\overline{\delta}-\ln\underline{\delta}}$ Jafari Jozani and Parsian (2008)
Stein's Loss (SL)	$\frac{\delta}{\gamma(\mathbf{y})}-\ln\frac{\delta}{\gamma(\mathbf{y})}-1$	$\frac{\ln(1/\overline{\delta})-\ln(1/\underline{\delta})}{1/\overline{\delta}-1/\underline{\delta}}$ Jafari Jozani and Jafari Tabrizi (2013)
Square Log Error Loss (SLEL)	$(\ln\delta-\ln\gamma(\mathbf{y}))^2$	$\sqrt{\overline{\delta}\underline{\delta}}$ Kiapour and Nematollahi (2011)
h-Loss (HL)	$(h(\delta)-h(\gamma(\mathbf{y})))^2$	$h^{-1}\left(\frac{h(\underline{\delta})+h(\overline{\delta})}{2}\right)$ Jafari Jozani, Marchand and Parsian (2012)
Intrinsic Loss (IL)	$\ln\frac{\beta(\gamma(\mathbf{y}))}{\beta(\delta)}+(\delta-\gamma(\mathbf{y}))\frac{\beta'(\gamma(\mathbf{y}))}{\beta(\gamma(\mathbf{y}))}$ $\beta(.)>0$	$\frac{\overline{\delta}H(\overline{\delta})-\underline{\delta}H(\underline{\delta})-\ln(\beta(\overline{\delta})/\beta(\underline{\delta}))}{H(\overline{\delta})-H(\underline{\delta})}$ Jafari Jozani and Jafari Tabrizi (2013)

Definition 2. Consider a class Γ of prior distributions on an unknown parameter θ. Suppose $\pi\in\Gamma$, and let $R(\delta,\delta^\pi)$ be the posterior regret, as in (2.2), where δ^π is the Bayes estimator of $\gamma(\mathbf{y})$ with respect to π. Suppose $\mathcal{L}$ is a class of loss functions $L(\gamma(\mathbf{y}),\delta):\mathbb{R}^{\mathbf{2}}\longrightarrow\mathbb{R}^+$, such that L is a BS function of both $\gamma(\mathbf{y})$ and δ. We call $\mathcal{L}$ a class of regret loss functions if

$$R(\delta,\delta^\pi)=L(\delta^\pi,\delta). \tag{2.3}$$

One can easily show that many commonly used loss functions satisfy (2.3); see Table 1. In the following result, we obtain a necessary condition for a strictly BS loss function to satisfy (2.3). Assume that the predictive distribution of $\gamma(\mathbf{y})$ given $\mathbf{y}_s$, say $h(\gamma(\mathbf{y})|\mathbf{y}_s)$, is not trivial, that is, it is degenerate, because the result is always true for degenerate $h(\gamma(\mathbf{y})|\mathbf{y}_s)$. Now, we have the following result, which is proved in the Supplementary Material.

Lemma 1. *Suppose $\delta^\pi(\mathbf{y}_s)$ is the Bayes predictor of $\gamma(\mathbf{y})$ under a strictly BS loss function $L(\gamma(\mathbf{y}),\delta)$ with respect to a prior distribution π. Suppose the posterior predictive distribution $h(\gamma(\mathbf{y})|\mathbf{y}_s)$ is not degenerate. Then,* (2.3) *does not hold for any strictly BS loss function $L(\gamma(\mathbf{y}),\delta)$ that is bounded.*

3. PRGM Predictors of $\gamma(\mathbf{y})$ under the Regret Loss Functions

In this section, we obtain PRGM predictors of $\gamma(\mathbf{y})$ under regret loss functions and a general class Γ of priors. PRGM predictors are constructed to minimize the maximum posterior regret in predicting $\gamma(\mathbf{y})$. The posterior regret in a Bayesian analysis is essentially the difference between the posterior risk associated with the best predictor that could have been used if we knew the true prior distribution in Γ and the posterior risk of the predictor that was actually used. Using the minimax strategy, we choose a predictor that minimizes the maximum of this posterior regret within the class of all predictors when the prior distribution varies in Γ. This helps to protect against the effects of priors that are causing the worst posterior risk. The key result of this section is given in the following theorem, which is proved in the Supplementary Material.

Theorem 1. *Suppose $L(\gamma(\mathbf{y}),\delta)$ is a regret loss function. Let $\mathbf{y}$ be a random vector with pdf $f(\mathbf{y}|\theta)$. Suppose θ has a prior distribution with a pdf $\pi(\cdot)$ that belongs to a class Γ of priors, and $\mathcal{D}$ is the class of all predictors. Let $\underline{\delta}(\mathbf{y}_s) = \inf_{\pi\in\Gamma}\delta^{\pi}(\mathbf{y}_s)$ and $\overline{\delta}(\mathbf{y}_s) = \sup_{\pi\in\Gamma}\delta^{\pi}(\mathbf{y}_s)$ be finite, where $\delta^{\pi}(\mathbf{y}_s)$ is the Bayes predictor of $\gamma(\mathbf{y})$ with respect to $\pi\in\Gamma$ under the loss function $L(\gamma(\mathbf{y}),\delta)$. The PRGM predictor of $\gamma(\mathbf{y})$, denoted by $\delta^{p}(\mathbf{y}_s)$, is given as a solution to the following equation:*

$$L(\overline{\delta}(\mathbf{y}_s),\delta^{p}(\mathbf{y}_s)) = L(\underline{\delta}(\mathbf{y}_s),\delta^{p}(\mathbf{y}_s)),\quad \textit{for all } \mathbf{y}_s\in\mathbb{R}^n. \tag{3.1}$$

If the solution to (3.1) *is not unique, the PRGM predictor is chosen as the solution that results in the minimum $L(\overline{\delta}(\mathbf{y}_s),\delta^{p}(\mathbf{y}_s))$.*

Table 1 provides the PRGM predictors of $\gamma(\mathbf{y})$ under a general class of prior distributions Γ for some commonly used loss functions in the literature.

Remark 1. In some cases, the posterior regret function (2.2) has the following form:

$$R(\delta(\mathbf{y}_s),\delta^{\pi}(\mathbf{y}_s)) = k(\mathbf{y}_s,\alpha)L(\delta^{\pi}(\mathbf{y}_s),\delta(\mathbf{y}_s)),$$

where $k(\mathbf{y}_s,\alpha)$ is a function of $\mathbf{y}_s$ and α, with α being a hyperparameter associated with the prior distribution π_α in the class Γ of priors. When $k(\mathbf{y}_s,\alpha)$ does not depend on α, that is, $k(\mathbf{y}_s,\alpha) = k^*(\mathbf{y}_s)$, then, similarly to the proof of Theorem 1, it can be shown that the PRGM predictor of $\gamma(\mathbf{y})$ is the solution to equation (3.1). For example, suppose the loss function is given by

$$L_w(\gamma(\mathbf{y}),\delta) = \left[\left(\frac{\delta}{\gamma(\mathbf{y})}\right)^{w/2} - \left(\frac{\gamma(\mathbf{y})}{\delta}\right)^{w/2}\right]^2 = \left(\frac{\delta}{\gamma(\mathbf{y})}\right)^{w} + \left(\frac{\gamma(\mathbf{y})}{\delta}\right)^{w} - 2.$$

Then, the Bayes predictor of $\gamma(\mathbf{y})$ is given by $\delta^{\pi}(\mathbf{y}_s)=\sqrt[2w]{E(\gamma^{w}(\mathbf{y})|\mathbf{y}_s)/E(\gamma^{-w}(\mathbf{y})|\mathbf{y}_s)}$, and one can easily show that

$$\begin{aligned} R(\delta(\mathbf{y}_s), \delta^{\pi}(\mathbf{y}_s)) &= \sqrt{E(\gamma^{w}(\mathbf{y})|\mathbf{y}_s)E(\gamma^{-w}(\mathbf{y})|\mathbf{y}_s)}\ L_w(\delta^{\pi}(\mathbf{y}_s), \delta(\mathbf{y}_s)) \\ &= K(\mathbf{y}_s, \pi_\alpha)\ L_w(\delta^{\pi}(\mathbf{y}_s), \delta(\mathbf{y}_s)). \end{aligned}$$

Now, if $K(\mathbf{y}_s, \pi_\alpha)$ does not depend on the hyperparameter α, then the PRGM predictor of $\gamma(\mathbf{y})$ is the solution to equation (3.1), and is given by $\delta^{p}(\mathbf{y}_s) = \sqrt{\overline{\delta}(\mathbf{y}_s)\ \underline{\delta}(\mathbf{y}_s)}$.

4. PRGM Prediction under Various Superpopulation Models

In this section, we use Theorem 1 to find PRGM predictors of characteristics of finite populations, such as the mean and the variance, under some regret loss functions and various normal and non-normal superpopulation models, with or without using auxiliary variables. First, we consider the prediction of the finite-population mean and variance when the underlying population is assumed to be generated from a normally distributed superpopulation. We then study the problem when auxiliary variables are also used in the prediction process. Finally, we provide results for a PRGM prediction for a non-normal superpopulation model. Note that results can be obtained for any population parameters. However, we only present those related to predicting the population mean and variance, because these are the two main parameters of interest in many finite-population studies.

4.1. PRGM prediction of the mean (normal model without auxiliary variables)

Consider the superpopulation model (Ghosh and Kim (1993))

$$y_i = \theta + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \ldots, N, \tag{4.1}$$

where ε_i, for $i = 1, \ldots, N$, is i.i.d. Suppose σ^2 is known and θ is distributed according to a $N(\mu, \tau^2)$ distribution. We find the PRGM predictors of the mean $\gamma_1(\mathbf{y})$ under the SE and LINEX loss functions and three classes of prior distributions. Let $M = \sigma^2/\tau^2$, $B = M/(M+n)$, $\overline{y}_s = n^{-1}\sum_{i=1}^{n} y_i$, and $\mathbf{y}_{\overline{s}} = \{y_i : i \notin s\}$. It is easy to see that

$$\mathbf{y}_{\overline{s}}|\mathbf{y}_s \sim MN\Big(\{\overline{y}_s - B(\overline{y}_s - \mu)\}\mathbf{1}_{N-n}, \sigma^2(\mathbf{I}_{N-n} + (M+n)^{-1}\mathbf{J}_{N-n})\Big), \tag{4.2}$$

where MN denotes a multivariate normal distribution, $\mathbf{1}_{N-n}$ is a vector of ones, $\mathbf{I}_{N-n}$ is the identity matrix, and $\mathbf{J}_{N-n} = \mathbf{1}_{N-n}\mathbf{1}'_{N-n}$. Using (4.2), the conditional distribution of $\gamma_1(\mathbf{y})$ given $\mathbf{y}_s$ is given by

$$\gamma_1(\mathbf{y})|\mathbf{y}_s \sim N\left(\overline{y}_s - (1-f)B(\overline{y}_s - \mu), \frac{\sigma^2(1-f)}{n}(1-(1-f)B)\right), \quad (4.3)$$

where $f = n/N$. In general, choosing an appropriate class Γ of priors for posterior robustness is important. The ultimate goal is to choose a large Γ to ensure that nothing is left out, while also ensuring that Γ is not so large that posterior robustness is not achievable. As suggested by Berger (1990), the process of a robust Bayesian analysis should be regarded as a data-interactive process in which we start with a perhaps very large Γ, check the robustness, and progressively refine Γ (if needed) until robustness is achieved. In this section, we consider the following classes of prior distributions for θ in (4.1):

$$\begin{aligned}
\Gamma_\mu &= \{\pi : \pi \text{ is } N(\mu, \tau_0^2); \mu_1 \le \mu \le \mu_2,\ \tau_0^2 \text{ is a known constant}\},\\
\Gamma_{\tau^2} &= \{\pi : \pi \text{ is } N(\mu_0, \tau^2); \tau_1^2 \le \tau^2 \le \tau_2^2,\ \mu_0 \text{ is a known constant}\},\\
\Gamma_\epsilon &= \{\pi : \pi = (1-\epsilon)\pi_0 + \epsilon\, q, q \in Q\}.
\end{aligned}$$

Classes Γ_μ and Γ_{τ^2} are appealing because they are very easy to work with (cf., Goldstein (1980); Berger (1985)). However, they often fail to include many priors that are plausible. For example, they do not admit much variation in the prior tails, and hence may provide the illusion that robustness is obtained. In such situations, one might consider Γ_ϵ as a very rich and flexible alternative class of prior distributions (e.g., Berger (1985)), where π_0 is a base prior, q is a contamination, Q is a class of plausible distribution functions, and $0 \le \epsilon \le 1$ reflects the amount of contamination.

Example 1. (SE and LINEX loss functions) Under the SE loss function, using (4.3), and as we show in the Supplementary Material, the PRGM predictors of the population mean $\gamma_1(\mathbf{y}) = (1/N)\sum_{i=1}^N y_i$ under Γ_μ, Γ_{τ^2}, and Γ_ϵ are give, respectively, by

$$\delta_\mu^{PRGM}(\mathbf{y}_s) = \overline{y}_s - (1-f)B_0\left(\overline{y}_s - \frac{\mu_1+\mu_2}{2}\right), \quad (4.4)$$

$$\delta_\tau^{PRGM}(\mathbf{y}_s) = \overline{y}_s - (1-f)(\overline{y}_s - \mu_0)\left(\frac{B_1+B_2}{2}\right), \quad (4.5)$$

$$\delta_\epsilon^{PRGM}(\mathbf{y}_s) = f\overline{y}_s + \frac{(1-f)}{2}\left(\frac{a\delta^0(\mathbf{y}_s) + \theta_l f(\mathbf{y}_s|\theta_l)}{a + f(\mathbf{y}_s|\theta_l)} + \frac{a\delta^0(\mathbf{y}_s) + \theta_u f(\mathbf{y}_s|\theta_u)}{a + f(\mathbf{y}_s|\theta_u)}\right), \quad (4.6)$$

where $B_i = \sigma^2/(\sigma^2 + n\tau_i^2)$, for $i = 0, 1, 2$, and $\delta^0(\mathbf{y}_s) = E_{\pi_0}(\theta|\mathbf{y}_s)$. In addition, $a = ((1-\epsilon)/\epsilon)m(\mathbf{y}_s|\pi_0)$, with $m(\mathbf{y}_s|\pi_0)$ being the marginal (predictive) density of $\mathbf{y}_s$ under the prior distribution π_0, and $\theta_\xi = (\sigma/\sqrt{n})\nu_\xi + \overline{y}_s$, $\xi \in \{l, u\}$, where ν_l and ν_u are solutions to the following equation in ν, for some specific values of c and b defined in the Supplementary Material for Example 1:

$$e^{-\nu^2/2} - c\nu^2 - b\nu + c = 0. \tag{4.7}$$

One can see that δ_μ^{PRGM} is a Bayes predictor of $\gamma_1(\mathbf{y})$ with respect to (w.r.t.) $\pi_{\mu^*} \in \Gamma_\mu$, with a $\mu^* = ((\mu_1 + \mu_2)/2) \in [\mu_1, \mu_2]$. Similarly, δ_τ^{PRGM} is a Bayes predictor of $\gamma_1(\mathbf{y})$ w.r.t. $\pi_{\tau_*^2} \in \Gamma_{\tau^2}$, with $\tau_*^2 = ((2n\tau_1^2\tau_2^2 + \sigma^2(\tau_1^2 + \tau_2^2))/(2\sigma^2 + n\tau_1^2 + n\tau_2^2))$ $(\tau_1^2 \le \tau_*^2 \le \tau_2^2)$. Furthermore, $\delta_\epsilon^{PRGM}(\mathbf{y}_s)$ can be considered a compromise between the Bayes predictor under the prior distribution π_0 (associated with $\epsilon = 0$, corresponding to the case where one is very confident in π_0) and the predictor obtained as the mid-range of the class of Bayes predictors under the ϵ-contaminated class of priors when ϵ is close to one. Under the LINEX loss function, the PRGM predictors of the population mean $\gamma_1(\mathbf{y})$ under Γ_μ, Γ_{τ^2}, and Γ_ϵ are given, respectively, by

$$\begin{aligned}\delta_\mu^{PRGM}(\mathbf{y}_s) =& \overline{y}_s - \frac{c(1-f)\sigma^2}{2n}(1-(1-f)B_0) \\ & - \frac{1}{c}\ln\frac{e^{cB_0(1-f)(\overline{y}_s-\mu_1)} - e^{cB_0[(1-f)(\overline{y}_s-\mu_2)}}{c(\mu_2-\mu_1)(1-f)B_0}\end{aligned} \tag{4.8}$$

$$\begin{aligned}\delta_\tau^{PRGM}(\mathbf{y}_s) =& \overline{y}_s - \frac{c(1-f)\sigma^2}{2n} \\ & - \frac{1}{c}\ln\frac{e^{cB_2[(1-f)(\overline{y}_s-\mu_0)-c(1-f)^2\sigma^2/2n]} - e^{cB_1[(1-f)(\overline{y}_s-\mu_0)-c(1-f)^2\sigma^2/2n]}}{c(B_2-B_1)[(1-f)(\overline{y}_s-\mu_0)-c(1-f)^2\sigma^2/2n]},\end{aligned} \tag{4.9}$$

$$\begin{aligned}\delta_\epsilon^{PRGM}(\mathbf{y}_s) =& f\overline{y}_s - \frac{c\sigma^2(1-f)^2}{2(N-n)} \\ & - \frac{1}{c}\ln\Bigg(\bigg(\frac{a_0 + e^{-c(1-f)\theta_u}f(\mathbf{y}_s|\theta_u)}{a + f(\mathbf{y}_s|\theta_u)} - \frac{a_0 + e^{-c(1-f)\theta_l}f(\mathbf{y}_s|\theta_l)}{a + f(\mathbf{y}_s|\theta_l)}\bigg) \\ & \bigg/\bigg(\ln\frac{a_0 + e^{-c(1-f)\theta_u}f(\mathbf{y}_s|\theta_u)}{a + f(\mathbf{y}_s|\theta_u)} - \ln\frac{a_0 + e^{-c(1-f)\theta_l}f(\mathbf{y}_s|\theta_l)}{a + f(\mathbf{y}_s|\theta_l)}\bigg)\Bigg),\end{aligned} \tag{4.10}$$

where ν_l and ν_u are obtained numerically as solutions to a nonlinear equation in ν; see the Supplementary Material.

4.2. PRGM prediction of the variance (normal model without auxiliary variables)

Here, we consider model (4.1), assuming that both θ and σ^2 are unknown, and obtain the PRGM predictors of the population variance $\gamma_2(\mathbf{y}) = (1/N)\sum_{i=1}^{n}(y_i - \gamma_1(\mathbf{y}))^2$ under the SE loss function and two classes of prior distributions. Let the prior distribution for σ^2 be the inverse gamma distribution $I\Gamma(\alpha,\beta)$, with known α and β, while we choose a noninformative prior for θ with density $\pi(\theta) = 1$, for $\theta \in \mathbb{R}$, resulting in $\pi(\theta,\sigma^2) \propto \sigma^{-2(\alpha+1)}e^{-\beta/\sigma^2}$. Let $A = (1/2)\sum_{i\in s}(y_i - \overline{y}_s)^2 + \beta$, $A_i = (1/2)\sum_{i\in s}(y_i - \overline{y}_s)^2 + \beta_i$, for $i = 0,1,2$, $\bar{y}_{\bar{s}} = (1/(N-n))\sum_{i\notin s} y_i$, $s^2 = (1/(n-1))\sum_{i\in s}(y_i - \bar{y}_s)^2$, and $s_{\bar{s}}^2 = (1/(N-n-1))\sum_{i\notin s}(y_i - \bar{y}_{\bar{s}})^2$. It can be shown that $E(\theta|\mathbf{y}_s) = \bar{y}_s$, $V(\theta|\mathbf{y}_s) = (2A/(n(n+2\alpha-3)))$, and $\pi(\theta,\sigma^2|\mathbf{y}_s) = c(1/\sigma^2)^{((n+2\alpha+2)/2)}\exp\{-(1/\sigma^2)(A + (n/2)(\bar{y}_s - \theta)^2)\}$, where $c = \sqrt{n/2\pi}\, A^{((n+2\alpha-1)/2)}/\Gamma(((n+2\alpha-1)/2))$. In addition,

$$\sum_{i=1}^{N}(y_i - \bar{Y})^2 = (n-1)s^2 + (N-n-1)s_{\bar{s}}^2 + (f^2 + (1-f)^2)(\bar{y}_s - \bar{y}_{\bar{s}})^2, \quad (4.11)$$

where $\bar{y}_{\bar{s}}$, s^2, and $s_{\bar{s}}^2$, given θ and σ^2, are distributed according to $N(\theta,\sigma^2/(N-n))$, $\Gamma((n-1)/2, 2\sigma^2/(n-1))$, and $\Gamma((N-n-1)/2, 2\sigma^2/(N-n-1))$, respectively. By first taking the expectation of (4.11), conditional on θ,σ^2, and $\mathbf{y}_s$, and then calculating the expectation w.r.t. the posterior distribution of (θ,σ^2) given $\mathbf{y}_s$, one can easily see that

$$E\left[\sum_{i=1}^{N}(y_i - \bar{Y})^2\middle|\mathbf{y}_s\right] = \sum_{i\in s}(y_i - \overline{y}_s)^2 + \frac{2AN}{n+2\alpha-3}N_0,$$

with $N_0 = (N-n-1)/N + (f^2 + (1-f)^2)/(n(N-n))$. Now, the Bayes predictor of $\gamma_2(\mathbf{y})$ under the SE loss function is

$$\delta^{\pi}(\mathbf{y}_s) = \frac{1}{N}\sum_{i\in s}(y_i - \overline{y}_s)^2 + \frac{2A}{n+2\alpha-3}N_0. \quad (4.12)$$

We consider the following classes of prior distributions for σ^2:

$$I\Gamma_\alpha = \{\pi : \pi \text{ is } I\Gamma(\alpha,\beta_0); \alpha_1 \le \alpha \le \alpha_2,\ \beta_0 \text{ is a known constant}\},$$
$$I\Gamma_\beta = \{\pi : \pi \text{ is } I\Gamma(\alpha_0,\beta); \beta_1 \le \beta \le \beta_2,\ \alpha_0 \text{ is a known constant}\}.$$

Under $I\Gamma_\alpha$, the Bayes predictor of $\gamma_2(\mathbf{y})$ is given by (4.12) when A is replaced with A_0. From Table 1, the PRGM predictor of $\gamma_2(\mathbf{y})$ under the SEL loss function is obtained as follows:

$$\delta_\alpha^{PRGM}(\mathbf{y}_s) = \frac{1}{N}\sum_{i\in s}(y_i - \overline{y}_s)^2 + \frac{2A_0}{n+2\alpha^*-3}N_0, \tag{4.13}$$

where $\alpha^* = (2\alpha_1\alpha_2 + (n-3)(\alpha_1+\alpha_2)/2)/((n-3)+\alpha_1+\alpha_2)$. Note that $\delta_\alpha^{PRGM}(\mathbf{y}_s)$ is a Bayes predictor of $\gamma_2(\mathbf{y})$ w.r.t. $\pi_{\alpha^*} \in I\Gamma_\alpha$, with $\alpha^* \in [\alpha_1, \alpha_2]$. Under $I\Gamma_\beta$, the Bayes predictor of $\gamma_2(\mathbf{y})$ under the SE loss function is given by (4.12), where α is replaced with α_0. Using Table 1, one can easily show that the PRGM predictor of $\gamma_2(\mathbf{y})$ is

$$\delta_\beta^{PRGM}(\mathbf{y}_s) = \frac{1}{N}\sum_{i\in s}(y_i - \overline{y}_s)^2 + \frac{2A^*}{n+2\alpha_0-3}N_0,$$

where $A^* = (1/2)\sum_{i\in s}(y_i - \overline{y}_s)^2 + \beta^*$ and $\beta^* = (\beta_1+\beta_2)/2$. Here again, δ_β^{PRGM} is a Bayes predictor of $\gamma_2(\mathbf{y})$ w.r.t. $\pi_{\beta^*} \in I\Gamma_{\beta^*}$, $(\beta_1 \le \beta^* \le \beta_2)$.

4.3. PRGM prediction of the mean (normal model with auxiliary variables)

In many populations, particularly those that have been previously sampled or surveyed, a frame of units is available, along with some auxiliary data on each unit. In other cases, a full frame of all units is not available, but can be constructed by sampling in stages and assembling a partial frame at each stage. In both single and multi-stage sampling designs, auxiliary data may be used to construct efficient estimators of population parameters such as the population total and mean. A superpopulation model is often used to formalize the relationship between a target variable and the auxiliary data. To this end, quantities of interest are modeled as realizations of random variables with a particular joint probability distribution. A frequently used model in survey sampling is

$$y_i = \beta x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0,\sigma^2), \quad i = 1,\ldots,N, \tag{4.14}$$

where x_i is the value of the auxiliary variable and ε_i is i.i.d. To perform a Bayesian analysis, one can use prior information that often exists in survey sampling in the form of auxiliary variables through administrative records. The Bayesian approach uses this auxiliary information explicitly through prior distributions for finite-population parameters, that is, distributions that relate these parameters and the auxiliary variables. It is common to assume that in model (4.14), x_i (and hence $\sum_{i=1}^N x_i$) and σ^2 are known and β follows a $N(\mu,\tau^2)$ distribution. Let $\mathbf{x}_s = \{x_i : i \in s\}$, $\mathbf{x}_{\bar{s}} = \{x_i : i \notin s\}$, $b_s = \sum_{i\in s} x_i y_i$, $d_s = \sum_{i\in s} x_i^2$,

$c_s = \sum_{i \notin s} x_i^2$, and $a_s = \sum_{i \notin s} x_i = \sum_{i=1}^N x_i - \sum_{i \in s} x_i$. Then, one can easily show that the posterior distribution of $\mathbf{y}_{\bar{s}}$, given $\mathbf{y}_s$ and $\mathbf{x}_s$, is a multivariate normal distribution

$$MN\left(\left((1-B_s)\mu + \frac{b_s}{d_s}B_s\right)\mathbf{x}_{\bar{s}}, \sigma^2\left(I_{N-n} + \frac{B_s}{d_s}\mathbf{x}_{\bar{s}}\mathbf{x}_{\bar{s}}^\top\right)\right), \tag{4.15}$$

where $B_s = \tau^2 d_s/(\sigma^2 + \tau^2 d_s)$. Using (4.15), we can show that

$$\gamma_1(\mathbf{y})|\mathbf{y}_s, \mathbf{x}_s \sim N\left(f\overline{y}_s + \left((1-B_s)\mu + \frac{b_s}{d_s}B_s\right)\frac{a_s}{N}, \frac{\sigma^2}{N^2}\left((N-n) + \frac{B_s}{d_s}a_s^2\right)\right). \tag{4.16}$$

Thus,

$$\begin{aligned} E(\gamma_1(\mathbf{y})|\mathbf{y}_s, \mathbf{x}_s) &= f\overline{y}_s + \left((1-B_s)\mu + \frac{b_s}{d_s}B_s\right)\frac{a_s}{N}, \\ V(\gamma_1(\mathbf{y})|\mathbf{y}_s, \mathbf{x}_s) &= \frac{\sigma^2}{N^2}\left((N-n) + \frac{B_s}{d_s}a_s^2\right), \end{aligned} \tag{4.17}$$

and

$$\begin{aligned} -\frac{1}{c}\ln E(e^{-c\gamma_1(\mathbf{y})}|\mathbf{y}_s, \mathbf{x}_s) = f\overline{y}_s &+ \left((1-B_s)\mu + \frac{b_s}{d_s}B_s\right)\frac{a_s}{N} \\ &-\frac{1}{2}c\frac{\sigma^2}{N^2}\left((N-n) + \frac{B_s}{d_s}a_s^2\right). \end{aligned} \tag{4.18}$$

We consider the Γ_μ, Γ_{τ^2}, and Γ_ϵ classes of prior distributions for β in (4.14). In the following examples, we obtain the PRGM predictors of $\gamma_1(\mathbf{y}) = (1/N)\sum_{i=1}^N y_i$ under the SE and LINEX loss functions.

Example 2. (SE and LINEX loss functions) Under the SE loss function, as shown in the Supplementary Material, the PRGM predictors of the population mean under the Γ_μ, Γ_{τ^2}, and Γ_ϵ classes of priors are obtained, respectively, by

$$\delta_\mu^{PRGM}(\mathbf{y}_s) = f\overline{y}_s + \left((1-B_{s0})\left(\frac{\mu_1+\mu_2}{2}\right) + \frac{b_s}{d_s}B_{s0}\right)\frac{a_s}{N}, \tag{4.19}$$

$$\delta_\tau^{PRGM}(\mathbf{y}_s) = f\overline{y}_s + \frac{a_s}{N}\mu_0 + \frac{a_s}{N}\left(\frac{b_s}{d_s} - \mu_0\right)\left(\frac{B_{s1}+B_{s2}}{2}\right), \tag{4.20}$$

$$\delta_\epsilon^{PRGM}(\mathbf{y}_s) = f\overline{y}_s + \frac{(1-f)\bar{x}_{\bar{s}}}{2}\left(\frac{a\delta^0(\mathbf{y}_s) + \beta_l f(\mathbf{y}_s|\beta_l)}{a + f(\mathbf{y}_s|\beta_l)} + \frac{a\delta^0(\mathbf{y}_s) + \beta_u f(\mathbf{y}_s|\beta_u)}{a + f(\mathbf{y}_s|\beta_u)}\right), \tag{4.21}$$

where a is defined as in Example 1, $B_{si} = \tau_i^2 d_s/(\sigma^2 + \tau_i^2 d_s)$, for $i = 0, 1, 2$, and $\delta^0(\mathbf{y}_s) = E_{\pi_0}(\beta|\mathbf{y}_s)$. Furthermore, $\beta_\xi = (\sigma/\sqrt{d_s})\nu_\xi + b_s/d_s$, $\xi \in \{l, u\}$, where ν_l and ν_u are obtained as solutions to a nonlinear equation in ν; see the Supplementary Material for Example 2. Here again, we show that $\delta_{\mu}^{PRGM}(\mathbf{y}_s)$ and $\delta_{\tau}^{PRGM}(\mathbf{y}_s)$ are Bayes estimators with specific choices of prior distributions within Γ_μ and Γ_{τ^2}, respectively. Under the LINEX loss function, the PRGM predictors of the population mean under the Γ_μ, Γ_{τ^2}, and Γ_ϵ classes of priors are obtained, respectively, by

$$\begin{aligned}\delta_{\mu}^{PRGM} &= f\bar{y}_s + a_s b_s \frac{B_{s0}}{N d_s} - \frac{c\sigma^2}{2N^2}\left((N-n) + \frac{B_{s0}}{d_s} a_s^2\right) \\ &\quad -\frac{1}{c}\ln \frac{e^{-(ca_s/N)(1-B_{s0})\mu_2} - e^{-(ca_s/N)(1-B_{s0})\mu_1}}{-c(1/N)a_s(1-B_{s0})(\mu_2-\mu_1)}, \end{aligned} \tag{4.22}$$

$$\begin{aligned}\delta_{\tau}^{PRGM} &= f\bar{y}_s + \frac{a_s}{N}\mu_0 - \frac{c\sigma^2}{2N}(1-f) \\ -\frac{1}{c}\ln & \frac{e^{-cB_{s1}(a_s/N(b_s/d_s-\mu_0)-(c\sigma^2/2N^2)(a_s^2/d_s))} - e^{-cB_{s2}(a_s/N(b_s/d_s-\mu_0)-(c\sigma^2/2N^2)(a_s^2/d_s))}}{-c(B_{s1}-B_{s2})(a_s/N(b_s/d_s-\mu_0)-(c\sigma^2/2N^2)(a_s^2/d_s))}, \end{aligned} \tag{4.23}$$

$$\begin{aligned}\delta_{\epsilon}^{PRGM} &= f\bar{y}_s - \frac{c\sigma^2(1-f)^2}{2(N-n)} \\ &\quad -\frac{1}{c}\ln\left(\left(\frac{b_0 + e^{-c(1-f)\beta_u \bar{x}_{\bar{s}}} f(\mathbf{y}_s|\beta_u)}{a + f(\mathbf{y}_s|\beta_u)} - \frac{b_0 + e^{-c(1-f)\beta_l \bar{x}_{\bar{s}}} f(\mathbf{y}_s|\beta_l)}{a + f(\mathbf{y}_s|\beta_l)}\right)\right. \\ &\quad \left. \Big/ \left(\ln \frac{b_0 + e^{-c(1-f)\beta_u \bar{x}_{\bar{s}}} f(\mathbf{y}_s|\beta_u)}{a + f(\mathbf{y}_s|\beta_u)} - \ln \frac{b_0 + e^{-c(1-f)\beta_l \bar{x}_{\bar{s}}} f(\mathbf{y}_s|\beta_l)}{a + f(\mathbf{y}_s|\beta_l)}\right)\right), \end{aligned} \tag{4.24}$$

where $t = E(e^{-c(1-f)\beta\bar{x}_{\bar{s}}}|\mathbf{y}_s)$ and $b_0 = at$. In addition, to calculate β_l and β_u, one needs to numerically solve for ν_l and ν_u as solutions to a nonlinear equation, defined in the Supplementary Material.

4.4. PRGM prediction of the mean (non-normal population)

In many applications, such as business surveys dealing with income data and some medical research, the underlying variable of interest is a positive and continuous random variable with a right-skewed distribution. In such cases, using normal models is not appropriate, and one might decide to use other densities, such as the gamma model (e.g., Engelhardt and Bain (1977), Glaser (1973), and Gross and Clark (1975)). In this section, we assume that our sample is taken from a superpopulation that is distributed according to a gamma model. In other words, given θ, suppose $y_1, \ldots, y_N$ are conditionally independent, with

$$y_i|\theta \sim \Gamma(\alpha, \theta), \quad i = 1, \ldots, N, \tag{4.25}$$

where α is assumed to be known. Therefore, we have

$$f(y_i|\theta) = \frac{\theta^\alpha y_i^{\alpha-1} e^{-\theta y_i}}{\Gamma(\alpha)}, \quad i = 1, \ldots, N,$$

and $T = \sum_{i \notin s} y_i$, given θ, has a $\Gamma((N-n)\alpha, \theta)$ distribution. We obtain the PRGM predictors of the population mean $\gamma_1(\mathbf{y}) = (1/N)\sum_{i=1}^N y_i$ under the SE loss function and two classes of prior distributions for θ. Let $\theta \sim \Gamma(a, b)$, with known a and b, $C = \sum_{i \in s} y_i + b$, and $C_i = \sum_{j \in s} y_j + b_i$, for $i = 0, 1, 2$. It is easily shown that

$$f(T = t|\mathbf{y}_s) = \frac{\Gamma((N-n)\alpha + n\alpha + a)}{\Gamma((N-n)\alpha)\Gamma(n\alpha + a)} \frac{(C/t)^{n\alpha+a+1}}{C(1 + C/t)^{(N-n)\alpha+n\alpha+a}}.$$

Using the change of variable $U = (1 + T/C)^{-1}$, we obtain the distribution of U given $\mathbf{y}_s$ as $Beta(n\alpha + a, (N-n)\alpha)$; thus, $E(T|\mathbf{y}_s) = CE((1-U)/U|\mathbf{y}_s) = C(N-n)\alpha/(n\alpha + a - 1)$. Hence, the Bayes predictor of $\gamma_1(\mathbf{y})$ under the SE loss function is as given as follows:

$$\delta^\pi(\mathbf{y}_s) = \frac{1}{N}\sum_{i \in s} y_i + \frac{1}{N} C \frac{(N-n)\alpha}{n\alpha + a - 1}. \tag{4.26}$$

We consider the following classes of prior distributions for θ in (4.25):

$$\Gamma_a = \{\pi : \pi \text{ is } \Gamma(a, b_0); a_1 \leq a \leq a_2,\ b_0 \text{ is a known constant}\},$$
$$\Gamma_b = \{\pi : \pi \text{ is } \Gamma(a_0, b); b_1 \leq b \leq b_2,\ a_0 \text{ is a known constant}\}.$$

Over Γ_a, the Bayes predictor of $\gamma_1(\mathbf{y})$ under the SE loss function is given by (4.26) when C is replaced with $C_0 = \sum_{j \in s} y_j + b_0$. The PRGM predictor of $\gamma_1(\mathbf{y})$ is then obtained as follows:

$$\delta_a^{PRGM}(\mathbf{y}_s) = \frac{1}{N}\sum_{i \in s} y_i + \frac{1}{N} C_0 \frac{(N-n)\alpha}{n\alpha + a^* - 1}, \tag{4.27}$$

where $a^* = (a_1 a_2 + n\alpha(a_1 + a_2)/2 - (a_1 + a_2)/2)/(n\alpha + (a_1 + a_2)/2 - 1)$. Note that $\delta_a^{PRGM}(\mathbf{y}_s)$ is a Bayes predictor of $\gamma_1(\mathbf{y})$ w.r.t. $\pi_{\alpha^*} \in \Gamma_a$. Under the class Γ_b of priors, the Bayes predictor of $\gamma_1(\mathbf{y})$ is given by (4.26), with a replaced with a_0. In addition, the PRGM predictor of $\gamma_1(\mathbf{y})$ is given by $\delta_b^{PRGM}(\mathbf{y}_s) =$

$(1/N)\sum_{i\in s} y_i + (1/N)C^*(N-n)\alpha/(n\alpha + a_0 - 1)$, where $C^* = \sum_{i\in s} y_i + b^*$ and $b^* = (b_1 + b_2)/2$. Here again, δ^{PRGM} is a Bayes predictor of $\gamma_1(\mathbf{y})$ w.r.t. $\pi_{b^*} \in \Gamma_{b^*}$.

5. Real-Data Applications and Simulation Studies

In this section, we study the performance of the PRGM predictors of the population mean and/or variance with respect to several classes of prior distributions compared with their corresponding Bayes predictors under the commonly used SE loss function and different superpopulation models. To this end, we consider three data sets for the prediction:

(1) the average and the variance of the weight loss of 579 participants in a special diet program in a clinical study in Iran;

(2) the average weight of 224 seven-month-old sheep at the Research Farm of Ataturk University, Erzurum, Turkey (Ozturk, Bilgin and Wolfe (2005); Jafaraghaie and Nematollahi (2018)); and

(3) the average remission time (in months) of 128 patients with bladder cancer from a study conducted by the American Cancer Society (Lee and Wang (2003); Lemonte and Cordeiro (2013)).

The first study deals with model (4.1), associated with a normal superpopulation model without auxiliary information, while the second study considers model (4.14), using an auxiliary variable. Finally, our third study deals with model (4.25), based on a gamma distribution as an example of a non-normal superpopulation model. We present the results for the first application in this paper; details of the second and third applications are presented in the Supplementary Material. In our first application, we study a finite population consisting of the weight loss measurements of 579 participants enrolled in a special diet program in a clinical study in Isfahan city of Iran in 2006. The weight loss is computed as the difference between the weight of each person before starting the program and after finishing it. Because the normality assumption for the weight loss was not rejected using the Kolmogorov-Smirnov test with a p-value = 0.514, we assume that our data set is a realization of a normal superpopulation model $N(\theta, \sigma^2)$, with $\theta = 3.82295$ and $\sigma^2 = 6.499586$, which are obtained using the maximum likelihood approach.

We would like to predict the average and variance of the weight loss due to the special diet. The doctors who are involved with this study have previous

Table 2. The PRGM predicted values of the finite-population mean over Γ_μ, Γ_{τ^2}, and Γ_ϵ under the SE loss function. Corresponding Bayes predictions are obtained under a $N(6, 0.5)$ prior distribution.

$\delta^{\pi_{\mu_0,\tau_0^2}}$	$\delta_{\Gamma_\mu}^{PRGM}$	$\delta_{\Gamma_{\tau^2}}^{PRGM}$	$\delta_{\Gamma_\epsilon}^{\pi}$	$\delta_{\Gamma_\epsilon}^{PRGM}$
4.753405	4.564885	4.68362	4.239414	3.972537

information from similar research conducted by the clinic, and they want to incorporate this information in the prediction process. The Bayesian methodology can be used to express a doctor's previous experience as suitable classes of prior distributions in the analysis. We do this by incorporating their information in the prediction process using Bayesian and robust Bayesian approaches. For the PRGM prediction of the population mean and variance, we choose reasonable classes of prior distributions for θ and σ^2, instead of working with completely determined prior distributions, and obtain the PRGM predictors of the mean and the variance of the weight loss. We also obtain the bias and variance associated with each prediction using simulation studies.

5.1. Predicting the average weight loss

To predict the finite-population mean, we consider a single prior distribution $N(\mu_0 = 6, \tau_0^2 = 0.5)$, as well as three classes of prior distributions for θ, denoted by $\Gamma_\mu = \{N(\mu, \tau_0^2) : \mu \in [2, 8]\}$, $\Gamma_{\tau^2} = \{N(\mu_0, \tau^2) : \tau^2 \in [0.1, 0.7] \subseteq \mathbb{R}^+\}$, and $\Gamma_\epsilon = \{\pi = (1-\epsilon)\pi_0 + \epsilon\, q : \pi_0 \sim N(6, 0.5), q \sim N(8, 0.3)\}$, $\epsilon = 0.5\}$. We obtain the usual Bayes predictor ($\delta^{\pi_{\mu_0,\tau_0^2}}$ with $\mu_0 = 6$ and $\tau_0^2 = 0.5$), the PRGM predictor over the class Γ_μ ($\delta_{\Gamma_\mu}^{PRGM}$), the PRGM predictor over the class Γ_{τ^2} ($\delta_{\Gamma_{\tau^2}}^{PRGM}$), and the Bayes and PRGM predictors over the class Γ_ϵ ($\delta_{\Gamma_\epsilon}^{\pi}, \delta_{\Gamma_\epsilon}^{PRGM}$). Table 2 summarizes the predicted values under the SE loss function for a fixed sample size $n = 50$. As shown the PRGM predicted values are closer to the true mean weight loss, that is, 3.82295, than are their corresponding Bayes predictors. To obtain the bias and precision associated with each prediction, we perform simulation studies. To this end, by considering the weight loss data as a realization of a superpopulation model, we extract samples from the underlying model in order to compute the predictors, MSEs, and biases. We repeat this process 10,000 times, and calculate the estimated MSE (EMSE) and absolute bias (EAB) of each predictor. To study the effect of the sample size, we repeat our study with different sample sizes, $n = 20, 30$, and 50, for comparison. In addition, we study the effect of μ_0 and τ_0^2 on the performance of the Bayes predictors compared with their corresponding PRGM predictors under the Γ_μ, Γ_{τ^2}, and Γ_ϵ

classes of priors. To this end, we consider $\pi_0 \sim N(\mu_0, \tau_0^2)$, with $\mu_0 = 2, 4, 6$, and 8 and $\tau_0^2 = 0.1, 0.3, 0.5$, and 0.7. Furthermore, we use the following four ϵ-contaminated classes of prior distributions corresponding to different choices of π_0 and q, and study the PRGM prediction with different values of contamination $\epsilon \in \{0, 0.2, \ldots, 0.8, 1\}$:

1. $\Gamma_\epsilon^1 = \{\pi \ : \ \pi = (1-\epsilon)N(6, 0.3) + \epsilon N(5, 0.2)\}$,
2. $\Gamma_\epsilon^2 = \{\pi \ : \ \pi = (1-\epsilon)N(6, 0.3) + \epsilon N(8, 0.3)\}$,
3. $\Gamma_\epsilon^3 = \{\pi \ : \ \pi = (1-\epsilon)N(6, 0.5) + \epsilon N(5, 0.2)\}$,
4. $\Gamma_\epsilon^4 = \{\pi \ : \ \pi = (1-\epsilon)N(6, 0.5) + \epsilon N(8, 0.3)\}$,

To calculate the PRGM predictors, we use the necessary expressions developed in Example 1, where Γ_ϵ^i, $i = 1, \ldots, 4$, and the necessary values of ν_l and ν_u are obtained numerically as solutions to the corresponding nonlinear equations (4.7). We perform simulation studies using the following steps in order to calculate the precision and bias associated with each prediction:

1. Generate $\epsilon_1^*, \epsilon_2^*, \ldots, \epsilon_n^*$ from a $N(0, 6.499586)$ distribution.
2. Create $y_i^* = 3.822954 + \epsilon_i^*$ and consider y_i^*, for $i = 1, \ldots, n$, as samples generated from the underlying superpopulation model.
3. Calculate the Bayes and PRGM predictors.
4. Repeat steps 1–3 $b = 10^4$ times, and calculate the values of EMSE and EAB for each predictor using the following formula:

$$\text{EMSE} = \frac{1}{b}\sum_{i=1}^{b}(\hat{\delta}_i^k - \bar{Y})^2, \quad \text{EAB} = |\frac{1}{b}\sum_{i=1}^{b}(\hat{\delta}_i^k - \bar{Y})|, \quad k = Bayes, PRGM,$$

   where $\hat{\delta}_i^k$ is the predictor in the ith repetition of the sampling, and $\bar{Y}$ is the population mean.

Tables 3 and 4 present the EMSEs and EABs of the Bayes and the corresponding PRGM predictors under Γ_μ and Γ_{τ^2}, respectively. From Table 3, we observe that for small and large values of μ_0 ($\mu_0 = 2, 6, 8$) [moderate values of μ_0 ($\mu_0 = 4$)] and all values of τ_0^2, the PRGM predictors [the Bayes predictors] perform reasonably well compared with their corresponding Bayes predictors [the PRGM predictors] in terms of the EMSE and EAB. This is a useful observation, because for the Bayesian prediction, one needs to specify a prior distribution, which might be difficult to do in practice. However, the PRGM predictor works with a class of plausible priors, and still performs as well as the Bayesian approach. From Table 4, we observe that for all values of τ_0^2 and large values of μ_0 ($\mu_0 = 6, 8$), and for large values of τ_0^2 and moderate values of μ_0 ($\tau_0^2 = 0.5, 0.7$ and $\mu_0 = 4$), the PRGM predictors perform reasonably well compared with the Bayes predictors in terms of the EMSE and absolute bias. We have the opposite result for other values of τ_0^2 and μ_0. Note that the estimated values of the MSE and bias decrease as the sample size increases. Tables 5 and 6 present the EMSEs and EABs of the Bayes and PRGM predictors under the class Γ_ϵ^i, for $i = 1, \ldots, 4$, of priors for different values of ϵ. According to these tables, the PRGM predictors perform reasonably well compared with their corresponding Bayes predictors in terms of the EMSE and EAB under the Γ_ϵ^2 and Γ_ϵ^4 classes of priors. In addition, under Γ_ϵ^1 and Γ_ϵ^3, for large values of ϵ, the PRGM predictors are better than the Bayes predictors, and for small values of ϵ, the Bayes predictors are better than the PRGM predictors. Thus, when the contaminated distribution is far from the π_0 distribution (Γ_ϵ^2 and Γ_ϵ^4), the PRGM predictor performs reasonably well compared with the Bayes predictor, and when these two distributions are close (Γ_ϵ^1 and Γ_ϵ^3), the PRGM predictor (Bayes predictor) is preferred for large (small) values of ϵ. Note that in these tables, for $\epsilon = 0$, the PRGM predictor is equal to the Bayes predictor.

5.2. Predicting the variance of weight loss

For the prediction of the finite-population variance, we consider a single prior $I\Gamma(\alpha_0 = 10, \beta_0 = 3)$, as well as two classes of prior distributions, $I\Gamma_\beta = \{I\Gamma(\alpha_0, \beta) : \beta \in [1, 7] \subseteq \mathbb{R}^+\}$ and $I\Gamma_\alpha = \{I\Gamma(\alpha, \beta_0) : \alpha \in [4, 10] \subseteq \mathbb{R}^+\}$ for σ^2. We obtain the Bayes predictor ($\delta^{\pi_{\alpha_0,\beta_0}}$ with $\alpha_0 = 10$ and $\beta_0 = 3$), the PRGM predictor over the class $I\Gamma_\alpha$ ($\delta_{I\Gamma_\alpha}^{PRGM}$), and the PRGM predictor over the class $\Gamma_{I\Gamma_\beta}$ ($\delta_{I\Gamma_\beta}^{PRGM}$). Table 7 summarizes the predicted values under the SE loss function for a fixed sample size $n = 50$. As shown the PRGM predicted values are closer to the variance of the weight loss, that is, 6.499586, than are their corresponding Bayes predictions. To evaluate the performance of the Bayes and PRGM predictors of the population variance, we performed a simulation study similar to that presented for the mean (using the variance instead of the mean). Then we calculated the EMSE and EAB of each PRGM predictor for different sample sizes $n = 20, 30$, and 50 over the $I\Gamma_\alpha$ and $I\Gamma_\beta$ classes of priors, with $\alpha_0 = 4, 6, 8, 10$, and $\beta_0 = 1, 3, 5, 7$. We compare the performance of the PRGM predictors of the population variance with their corresponding Bayes predictors with respect to their associated inverse gamma prior distributions, with $\alpha_0 =$

Table 3. Simulated MSE and absolute bias for the Bayes and PRGM predictors for $\mu_0 = 2,4,6,8$, $\tau_0^2 = 0.1, 0.3, 0.5, 0.7$, and $\mu \in [2,8]$ over Γ_μ (losing weight data).

		δ^π				δ^{PRGM}
	n	$\mu_0 = 2$	$\mu_0 = 4$	$\mu_0 = 6$	$\mu_0 = 8$	
$\tau_0^2 = 0.1$	20	1.8300	0.0394	2.6092	9.5396	0.7793
	30	1.4227	0.0387	2.0217	7.3716	0.6093
EMSE	50	0.9089	0.0366	1.2975	4.6918	0.4004
	20	1.3446	0.1619	1.6085	3.0851	0.8702
EAB	30	1.1821	0.1592	1.4129	2.7104	0.7641
	50	0.9387	0.1529	1.1268	2.1596	0.6104
$\tau_0^2 = 0.3$	20	0.9129	0.0879	1.2792	4.4867	0.4315
	30	0.5986	0.0797	0.8255	2.8361	0.2945
EMSE	50	0.3118	0.0649	0.4285	1.4024	0.1708
	20	0.9130	0.2363	1.0953	2.0993	0.5968
EAB	30	0.7246	0.2259	0.8667	1.6618	0.4782
	50	0.5041	0.2031	0.6063	1.1577	0.3501
$\tau_0^2 = 0.5$	20	0.5999	0.1282	0.8136	2.6562	0.3263
	30	0.3765	0.1073	0.4955	1.5409	0.2192
EMSE	50	0.1933	0.0796	0.2503	0.7053	0.1294
	20	0.6979	0.2853	0.8332	1.5915	0.4824
EAB	30	0.5356	0.2620	0.6313	1.1985	0.3876
	50	0.3709	0.2248	0.4311	0.7922	0.2913
$\tau_0^2 = 0.7$	20	0.4623	0.1579	0.6030	1.7978	0.2867
	30	0.2897	0.1255	0.3630	1.0024	0.1940
EMSE	50	0.1528	0.0881	0.1872	0.4500	0.1171
	20	0.5844	0.3166	0.6832	1.2819	0.4379
EAB	30	0.4496	0.2833	0.5151	0.9382	0.3569
	50	0.3201	0.2366	0.3581	0.6066	0.2738

$4, 6, 8, 10$, $\beta_0 = 1, 3, 5, 7$. The EMSE and bias of each predictor under $I\Gamma_\beta$ and $I\Gamma_\alpha$ are presented in Tables 8 and 9 . From Table 8, we observe that the PRGM predictors perform satisfactorily compared with the Bayes predictor in terms of the EMSE and the associated bias for small values of β_0 ($\beta_0 = 1, 3$) and all values of α. However, we have the opposite result for moderate to large values of β_0 ($\beta_0 = 5, 7$). Note that the MSE and the bias decrease as the sample size increases. From Table 9, we observe that for all values of β_0 and large (small to moderate) values of α_0, the PRGM predictors (the Bayes predictors) are preferred to the Bayes predictors (the PRGM predictors) in terms of the EMSE and bias. Furthermore, the MSE and the bias decrease as the sample size increases.

Table 4. Simulated MSE and absolute bias for the Bayes and PRGM predictors for $\tau_0^2 = 0.1, 0.3, 0.5, 0.7$, $\mu_0 = 2, 4, 6, 8$, and $\tau^2 \in [0.1, 0.7]$ over Γ_{τ^2} (losing weight data).

		δ^{π}				δ^{PRGM}
μ_0	n	$\tau_0^2 = 0.1$	0.3	0.5	0.7	
2	20	1.8300	0.9129	0.5999	0.4623	1.8302
	30	1.4227	0.5986	0.3765	0.2897	1.4227
EMSE	50	0.9089	0.3118	0.1933	0.1528	0.9089
	20	1.3446	0.9130	0.698	0.5844	1.3447
EAB	30	1.1821	0.7246	0.53556	0.4496	1.1821
	50	0.9387	0.5041	0.3709	0.3201	0.9387
4	20	0.0394	0.0879	0.1282	0.1579	0.0946
	30	0.0387	0.0797	0.1073	0.1255	0.0840
EMSE	50	0.0366	0.0649	0.07961	0.0881	0.0662
	20	0.1619	0.2363	0.2853	0.3166	0.2519
EAB	30	0.1592	0.2259	0.2620	0.2833	0.2378
	50	0.1529	0.2031	0.2248	0.2366	0.2109
6	20	2.6092	1.2792	0.8136	0.6030	0.6029
	30	2.0217	0.8255	0.4955	0.3630	0.3630
EMSE	50	1.2975	0.4285	0.2503	0.1872	0.1872
	20	1.6085	1.0953	0.8332	0.6832	0.6832
EAB	30	1.4129	0.8667	0.6313	0.5151	0.5151
	50	1.1268	0.6063	0.4311	0.3581	0.3581
8	20	9.5396	4.4867	2.6562	1.7978	1.7978
	30	7.3716	2.8361	1.5409	1.0024	1.0024
EMSE	50	4.6918	1.4024	0.7053	0.4500	0.4500
	20	3.0851	2.0993	1.5915	1.2812	1.2819
EAB	30	2.7104	1.6618	1.1985	0.9382	0.9382
	50	2.1596	1.1577	0.7922	0.6066	0.6066

6. Discussion

We have examined the PRGM prediction of population parameters under general classes of loss functions and prior distributions. In particular, we studied PRGM predictions in finite populations, and developed a unified approach to calculate these predictions under a very general setting. Under two different normal superpopulation models and different classes of prior distributions on the parameter of the underlying superpopulation model, we obtained the PRGM predictors of the finite-population mean under the LINEX and SE loss functions. Furthermore, we obtained the PRGM predictor of the finite-population variance under the SE loss function in a normal superpopulation model. We also considered a non-normal model, and derived the Bayes and PRGM predictors of the finite-

Table 5. Simulated MSE and absolute bias for the Bayes and PRGM predictors over Γ_ϵ (losing weight data).

		Γ_ϵ^1		$\pi_0 \sim N(6, 0.3)$		$q \sim N(5, 0.2)$		
	n	ϵ	0	0.2	0.4	0.6	0.8	1
EMSE	20	δ^{PRGM}	1.284	0.894	0.755	0.623	0.508	0.314
		δ^π	1.284	0.741	0.645	0.592	0.567	0.545
	30	δ^{PRGM}	0.833	0.529	0.425	0.354	0.289	0.218
		δ^π	0.833	0.484	0.429	0.411	0.396	0.393
	50	δ^{PRGM}	0.432	0.244	0.195	0.175	0.149	0.13
		δ^π	0.432	0.260	0.243	0.238	0.233	0.233
EAB	20	δ^{PRGM}	1.097	0.83	0.738	0.655	0.578	0.448
		δ^π	1.097	0.809	0.758	0.728	0.714	0.703
	30	δ^{PRGM}	0.869	0.613	0.533	0.479	0.429	0.374
		δ^π	0.869	0.641	0.605	0.593	0.583	0.582
	50	δ^{PRGM}	0.605	0.401	0.353	0.332	0.306	0.287
		δ^π	0.605	0.452	0.439	0.433	0.429	0.43
		Γ_ϵ^3		$\pi_0 \sim N(6, 0.5)$		$q \sim N(5, 0.2)$		
	n	ϵ	0	0.2	0.4	0.6	0.8	1
EMSE	20	δ^{PRGM}	0.819	0.707	0.643	0.561	0.479	0.314
		δ^π	0.819	0.662	0.614	0.578	0.562	0.545
	30	δ^{PRGM}	0.504	0.439	0.384	0.337	0.285	0.218
		δ^π	0.504	0.44	0.412	0.403	0.394	0.393
	50	δ^{PRGM}	0.256	0.223	0.194	0.179	0.153	0.13
		δ^π	0.256	0.241	0.236	0.235	0.231	0.233
EAB	20	δ^{PRGM}	0.836	0.738	0.686	0.626	0.566	0.448
		δ^π	0.836	0.759	0.737	0.719	0.71	0.703
	30	δ^{PRGM}	0.634	0.565	0.515	0.474	0.432	0.374
		δ^π	0.634	0.605	0.59	0.587	0.58	0.582
	50	δ^{PRGM}	0.434	0.391	0.358	0.341	0.313	0.287
		δ^π	0.434	0.43	0.43	0.429	0.427	0.43

population mean under the SE loss function. Then, we applied the results to different real data sets to illustrate the practical utility of the Bayes and PRGM procedures. We provided real-world data to predict finite-population means and variances under normal and non-normal models. We compared the estimated risk and bias of the obtained predictors under the SE loss function using simulation studies. In some cases, the Bayes predictors have smaller risk and bias than those of the robust Bayes predictors. However, we recommended using the robust Bayes predictors, owing to a lack of confidence in δ^π under $\pi = \pi(\mu_0, \tau_0^2)$, especially if it is difficult to specify a single prior distribution for the parameters

Table 6. Simulated MSE and absolute bias for the Bayes and PRGM predictors over Γ_ϵ (losing weight data) continued.

		Γ_ϵ^2		$\pi_0 \sim N(6, 0.3)$		$q \sim N(8, 0.3)$		
	n	ϵ	0	0.2	0.4	0.6	0.8	1
EMSE	20	δ^{PRGM}	1.284	0.894	0.755	0.623	0.508	0.314
		δ^π	1.284	1.264	1.275	1.268	1.277	4.472
	30	δ^{PRGM}	0.833	0.529	0.425	0.354	0.289	0.218
		δ^π	0.833	0.831	0.822	0.826	0.822	2.839
	50	δ^{PRGM}	0.432	0.244	0.195	0.175	0.149	0.13
		δ^π	0.432	0.43	0.429	0.429	0.426	1.4
EAB	20	δ^{PRGM}	1.097	0.83	0.738	0.655	0.578	0.448
		δ^π	1.097	1.088	1.093	1.09	1.093	2.096
	30	δ^{PRGM}	0.869	0.613	0.533	0.479	0.429	0.374
		δ^π	0.869	0.868	0.863	0.865	0.862	1.661
	50	δ^{PRGM}	0.605	0.401	0.353	0.332	0.306	0.287
		δ^π	0.605	0.603	0.603	0.601	0.6	1.154
		Γ_ϵ^4		$\pi_0 \sim N(6, 0.5)$		$q \sim N(8, 0.3)$		
	n	ϵ	0	0.2	0.4	0.6	0.8	1
EMSE	20	δ^{PRGM}	0.819	0.707	0.643	0.561	0.479	0.314
		δ^π	0.819	0.799	0.81	0.803	0.812	4.472
	30	δ^{PRGM}	0.504	0.439	0.384	0.337	0.285	0.218
		δ^π	0.504	0.502	0.494	0.497	0.494	2.839
	50	δ^{PRGM}	0.256	0.223	0.194	0.179	0.153	0.13
		δ^π	0.256	0.254	0.253	0.254	0.251	1.4
EAB	20	δ^{PRGM}	0.836	0.738	0.686	0.626	0.566	0.448
		δ^π	0.836	0.824	0.831	0.827	0.83	2.096
	30	δ^{PRGM}	0.634	0.565	0.515	0.474	0.432	0.374
		δ^π	0.634	0.633	0.629	0.631	0.627	1.661
	50	δ^{PRGM}	0.434	0.391	0.358	0.341	0.313	0.287
		δ^π	0.434	0.432	0.431	0.432	0.429	1.154

Table 7. The PRGM predicted values of the finite population variance over $I\Gamma_\beta$ and $I\Gamma_\alpha$ under the SE loss function. Bayes prediction of the variance is obtained under the $I\Gamma(10, 3)$ prior distribution.

$\delta^{\pi_{\alpha_0,\beta_0}}$	$\delta^{PRGM}_{I\Gamma_\beta}$	$\delta^{PRGM}_{I\Gamma_\alpha}$
5.650688	5.67791	6.198716

of the underlying superpopulation models.

The proposed methodology was used to study the impact of the prior distribution as an input to the Bayesian prediction process on the predicted values

Table 8. Simulated MSE and absolute bias for the Bayes and PRGM predictors of the finite-population variance for $\alpha_0 = 4, 6, 8, 10$, $\beta_0 = 1, 3, 5, 7$, and $\beta \in [1, 7]$ over $I\Gamma_\beta$ (losing weight data).

		δ^π				δ^{PRGM}
	n	$\beta_0 = 1$	$\beta_0 = 3$	$\beta_0 = 5$	$\beta_0 = 7$	
$\alpha_0 = 4$	20	4.496	4.074	3.699	3.372	3.881
EMSE	30	2.888	2.675	2.485	2.318	2.577
	50	1.611	1.529	1.455	1.391	1.491
	20	1.804	1.706	1.616	1.532	1.660
EAB	30	1.418	1.358	1.304	1.254	1.330
	50	1.043	1.013	0.986	0.962	0.999
$\alpha_0 = 6$	20	6.284	5.738	5.229	4.754	5.479
EMSE	30	3.933	3.637	3.361	3.103	3.497
	50	2.067	1.946	1.832	1.726	1.888
	20	2.225	2.112	2.002	1.897	2.057
EAB	30	1.715	1.639	1.567	1.497	1.603
	50	1.208	1.167	1.128	1.091	1.147
$\alpha_0 = 8$	20	8.290	7.690	7.118	6.573	7.401
EMSE	30	5.219	4.875	4.546	4.232	4.708
	50	2.685	2.535	2.391	2.255	2.462
	20	2.653	2.542	2.433	2.325	2.487
EAB	30	2.048	1.969	1.891	1.814	1.929
	50	1.419	1.372	1.326	1.281	1.349
$\alpha_0 = 10$	20	10.256	9.637	9.040	8.464	9.335
EMSE	30	6.584	6.210	5.850	5.503	6.028
	50	3.399	3.227	3.061	2.902	3.144
	20	3.026	2.923	2.821	2.720	2.872
EAB	30	2.370	2.292	2.215	2.138	2.253
	50	1.644	1.595	1.547	1.499	1.571

when the prior distribution ranges in certain classes of priors. If the impact is considerable, there is sensitivity, and one should use a robust Bayesian prediction, which is relatively insensitive to the uncertainty in determining the prior distribution. Note that a Bayesian analysis depends on other subjective inputs, such as the loss function and/or the model. A future research direction is to study the sensitivity of the Bayesian prediction jointly with respect to the prior and the loss function. Another important research direction is to study the effects of model-misspecification or imprecise probability models on the Bayesian prediction of the parameters of interest. One can also use other classes of prior distributions to obtain a robust Bayesian analysis. These include classes of priors with given

Table 9. Simulated MSE and absolute bias for the Bayes and PRGM predictors of the finite-population variance for $\beta_0 = 1, 3, 5, 7$, $\alpha_0 = 4, 6, 8, 10$, and $\alpha \in [4, 10]$ over $I\Gamma_\alpha$ (losing weight data).

		δ^{π}				δ^{PRGM}
	n	$\alpha_0 = 4$	$\alpha_0 = 6$	$\alpha_0 = 8$	$\alpha_0 = 10$	
$\beta_0 = 1$ EMSE	20	4.496	6.284	8.290	10.256	6.699
	30	2.888	3.933	5.219	6.584	4.278
	50	1.611	2.067	2.685	3.399	2.271
EAB	20	1.804	2.225	2.653	3.026	2.318
	30	1.418	1.715	2.048	2.370	1.807
	50	1.043	1.208	1.419	1.644	1.279
$\beta_0 = 3$ EMSE	20	4.074	5.738	7.690	9.637	6.138
	30	2.675	3.637	4.875	6.210	3.966
	50	1.529	1.946	2.535	3.227	2.138
EAB	20	1.706	2.112	2.542	2.923	2.204
	30	1.358	1.639	1.969	2.292	1.730
	50	1.013	1.167	1.372	1.595	1.236
$\beta_0 = 5$ EMSE	20	3.699	5.229	7.118	9.040	5.611
	30	2.485	3.361	4.546	5.850	3.673
	50	1.455	1.832	2.391	3.061	2.013
EAB	20	1.616	2.002	2.433	2.821	2.093
	30	1.304	1.567	1.891	2.215	1.655
	50	0.986	1.128	1.326	1.547	1.193
$\beta_0 = 7$ EMSE	20	3.372	4.754	6.573	8.464	5.117
	30	2.318	3.103	4.232	5.503	3.396
	50	1.391	1.726	2.255	2.902	1.895
EAB	20	1.532	1.897	2.325	2.720	1.985
	30	1.254	1.497	1.814	2.138	1.582
	50	0.962	1.091	1.281	1.499	1.153

marginals when dealing with multi-parameter cases, classes of ϵ-contaminated priors with shape constraints, or generalized moment classes of priors in which one considers classes of prior distributions that satisfy some moment conditions with a priori specified moments (e.g., Berger (1990)).

Supplementary Material

The online Supplementary Material provides derivations of the results presented in this paper. It also contains a comprehensive account of the real-data applications associated with our second and third studies, discussed in Section 5. Lastly, it provides complementary numerical results that show how the PRGM predictors of the finite population mean perform compared with their corresponding Bayes predictors.

Acknowledgments

The authors would like to thank Professor William Strawderman from Rutgers University for the useful discussion and his help with the proof of Lemma 1. We would also like to thank the associate editor and three anonymous referees for their constructive comments and suggestions. Mohammad Jafari Jozani gratefully acknowledges the research support of NSERC Canada. The research of Nader Nematollahi was supported by the Research Council of the Allameh Tabataba'i University, and part of this work was done during his visit to the University of Manitoba, Department of Statistics.

References

Berger, J. O. (1984). *The Robust Bayesian Viewpoint (with Discussion)*. Robustness of Bayesian Analysis (Edited by J. Kadane), 63–124. Amsterdam, North-Holland.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Berger, J. O. (1990). Robust Bayesian analysis: Sensitivity to the prior. *Journal of Statistical Planning and Inference* **26**, 303–328.

Betro, B. and Ruggeri, F. (1992). Conditional γ-minimax actions under convex losses. *Communications in Statistics: Theory and Methods* **21**, 1051–1066.

Bolfarine, H. (1990). Bayesian linear prediction in finite populations. *Annals of The Institute of Statistical Mathematics* **42**, 435–444.

Boratynska, A. (2006). Robust Bayesian prediction with asymmetric loss function in Poisson model of insurance risk. *Acta Universitatis Lodziensis, Folia Oeconomica* **196**, 123–138.

Boratynska, A. and Drozdowicz, M. (1999). Robust Bayesian estimation in a normal model with asymmetric loss function. *Applicationes Mathematicae* **26**, 85–92.

Chen, Q., Elliott, M. R. and Little, R. J. A. (2012). Bayesian inference for finite population quantiles from unequal probability samples. *Survey Methodology* **38**, 203–214.

Engelhardt, M. and Bain, L. J. (1977). Uniformly most powerful unbiased tests on the scale parameter of a gamma distribution with a nuisance shape parameter. *Technometrics* **19**, 77–81.

Ericson, W. A. (1969). Subjective Bayesian models in sampling finite populations (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)* **31**, 195–233.

Ghosh, M. (2008). Robust estimation in finite population sampling. *Institute of Mathematical Statistics* **1**, 116–122.

Ghosh, M. and Kim, D. H. (1993). Robust Bayes estimation of the finite population mean. *The Indian Journal of Statistics* **55**, 322–342.

Glaser, R. E. (1973). Inferences for a gamma distributed random variable with both parameters unknown with applications to reliability. Technical Report 154. Standford University.

Godambe, V. P. (1955). A unified theory of sampling from finite populations. *Journal of the Royal Statistical Society. Series B (Methodological)* **17**, 268–278.

Godambe, V. P. and Joshi, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations. I. *Annals of Mathematical Statistics* **36**, 1707–1722.

Godambe, V. P. and Thompson, M. E. (1971). The specification of prior knowldege by classes of prior distribution in survey sampling estimation. In *Foundations of Statistical Inference* (Edited by V. P. Godambe and D. A. Sprott), 243–258. Holt, Rinehart and Winston, Toronto.

Goldstein, M. (1980). The linear Bayes regression estimator under weak prior assumptions. *Biometrika* **67**, 621–628.

Gross, A. J. and Clark, V. A. (1975). *Survival Distributions: Reliability Applications in the Biomedical Services*. Wiley, New York.

Hamner, M. S., Seaman, W. and Young, M. (2001). Bayesian methods in finite population sampling. In *Proceedings of the Annual Meeting of the American Statistical Association.*

Hill, B. (1968). Posterior distribution of percentiles: Bayes theorem for sampling from a population. *Journal of the American Statistical Association* **63**, 677–691.

Jafaraghaie, R. and Nematollahi, N. (2018). Robust Bayesian prediction under a general linear-exponential posterior risk function and its application in finite population. *Communications in Statistics: Theory and Methods* **47**, 3269–3292.

Jafari Jozani, M. and Jafari Tabrizi, N. (2013). Intrinsic posterior regret gamma-minimax estimation for the exponential family of distributions. *Electronic Journal of Statistics* **7**, 1856–1874.

Jafari Jozani, M., Marchand, E. and Parsian, A. (2012). Bayesian and robust Bayesian analysis under a general class of balanced loss functions. *Statistical Papers* **53**, 51–60.

Jafari Jozani, M. and Parsian, A. (2008). Posterior regret γ-minimax estimation and prediction based on k-record data under entropy loss function. *Communications in Statistics: Theory and Methods* **37**, 2202–2212.

Kiapour, A. and Nematollahi, N. (2011). Robust Bayesian prediction and estimation under a squared log error loss function. *Statistics and Probability Letters* **81**, 1717–1724.

Kim, H. M. and Saleh, A. M. E. (2008). Prediction of finite population total with measurement error models. In *Recent Advances in Linear Models and Related Areas*, 79–93.

Lee, E. T. and Wang, J. (2003). *Statistical Methods for Survival Data Analysis*. Wiley, New York.

Lemonte, A. J. and Cordeiro, G. M. (2013). An extended lomax distribution. *Statistics* **47**, 800–816.

Little, R. J. A. (2004). To model or not to model? Competing modes of inference for finite population sampling. *Journal of the American Statistical Association* **99**, 546–556.

Liu, X. Q. and Rong, J. Y. (2007). Quadratic prediction problems in finite populations. *Statistics and Probability Letters* **77**, 483–489.

Meczarski, M. and Zielinski, R. (1991). Stability of the Bayesian estimator of the Poisson mean under the inexactly specified gamma prior. *Statistics and Probability Letters* **12**, 329–333.

Ozturk, O., Bilgin, O. C. and Wolfe, D. A. (2005). Estimation of population mean and variance in flock management: A ranked set sampling approach in a finite population setting. *Journal of Statistical Computation and Simulation* **75**, 905–919.

Pefeffermann, D. and Rao, C. R. (2009). *Handbook of Statistics*. Elsevier Science, Amsterdam.

Insua, D. R., Ruggeri, F. and Vidakovic, B. (1995). Some results on posterior regret Γ-minimax estimation. *Statistics and Decisions* **13**, 315–331.

Si, Y., Pillai, N. and Gelman, A. (2015). Bayesian nonparametric weighted sampling inference. *Bayesian Analysis* **1**, 1–21.

Zangeneh, S. Z. and Little, R. J. A. (2015). Bayesian inference for the finite population total from a heteroscedastic probability proportional to size sample. *Journal of Survey Statistics and Methodology* **3**, 162–192.

Zen, M. M. and DasGupta, A. (1993). Estimating a binomial parameter: Is robust Bayes real Bayes? *Statistics and Decisions* **11**, 37–60.

Nader Nematollahi

Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.

E-mail: nematollahi@atu.ac.ir

Mohammad Jafari Jozani

Department of Statistics, University of Manitoba, Winnipeg, MB R3T 2N2, Canada.

E-mail: m_jafari_jozani@umanitoba.ca

Razieh Jafaraghaie

Department of Mathematics and Statistics, Shoushtar Branch, Islamic Azad University, Shoushtar, Iran.

E-mail: r.aghaee2588@gmail.com

(Received February 2018; accepted December 2019)