Poisson Counts, Square Root Transformation and Small Area Estimation

Malay Ghosh and Tamal Ghosh
University of Florida, Gainesville, USA
Masayo Y. Hirose
Kyushu University, Nishi-ku, Japan

Abstract

The paper intends to serve two objectives. First, it revisits the celebrated Fay-Herriot model, but with homoscedastic known error variance. The motivation comes from an analysis of count data, in the present case, COVID-19 fatality for all counties in Florida. The Poisson model seems appropriate here, as is typical for rare events. An empirical Bayes (EB) approach is taken for estimation. However, unlike the conventional conjugate gamma or the log-normal prior for the Poisson mean, here we make a square root transformation of the original Poisson data, along with square root transformation of the corresponding mean. Proper back transformation is used to infer about the original Poisson means. The square root transformation makes the normal approximation of the transformed data more justifiable with added homoscedasticity. We obtain exact analytical formulas for the bias and mean squared error of the proposed EB estimators. In addition to illustrating our method with the COVID-19 example, we also evaluate performance of our procedure with simulated data as well.

AMS (2000) subject classification. Primary: 62C12; Secondary: 62D05.
Keywords and phrases. COVID19, Empirical Bayes, Fay-Herriot model, Random Effects Model, Stein-type shrinkage estimators.

The third author’s research was partially supported by JSPS KAKENHI grant number 18K12758.
1 Introduction

Small area estimation is now a topic of global importance. Methodologies abound, and many of these are finding real-life applications.

Normal theory small area estimation pervades the literature, and the pioneering (Fay and Herriot, 1979) method is most often used in real life applications. The Fay-Herriot model is a normal theory mixed effects area-level model. The model needs to assume known error variances in order to avoid non-identifiability, whereas in reality, these are sample estimates.

The present article deals with the Fay-Herriot model with known constant error variance. The motivation came from an analysis of count data, in particular, COVID-19 data, to find estimates of fatality for all counties in the state of Florida. The Poisson model is used, but we make a square transformation of the original data, and the corresponding mean parameters to attain a closer approximation to normality with added homoscedasticity. This is in contrast to the log-transformation, where also one typically assumes normality of the transformed data. However, transformation of the original data in the log-scale bears the potential hazard of leaving out zero counts, which on most occasions, can affect the conclusion significantly. Further, our approach allows one to develop Stein-type shrinkage estimators for small area means and study their properties analytically.

Variable transformation in the small area context has been addressed before. The logarithmic transformation with the assumption of log-normal distribution of the original data is most commonly used, for example, in modeling income distributions. See for example, Slud and Maiti (2006) and Ghosh et al. (2015). Recently, Hirose et al. (2021) considered an arc-sin transformation of binomial proportions for small area estimation. Sugasawa and Kubokawa (2017) suggested a non-explicit EB estimator and performed an analysis based on the dual power transformation similar to that of Hirose et al. (2021).

The remaining sections are as follows. In Section 2, we introduce the square root transformation, develop Stein-type shrinkage estimators for the transformed data motivated from an empirical Bayes point of view, and then back transform properly to estimate the original parameters of interest. In Sections 3 and 4, we obtain exact expressions for the bias and the mean squared error of our shrinkage estimator. Finally, Section 5 we also obtain an estimator of the mean squared error correct up to order $O(m^{-1})$ where $m$ is the number of small areas. Section 6 contains an illustration of the proposed method to estimate the number of deaths due to COVID-19 in each county. A simulation study is undertaken in Section 7. Some final remarks are made in Section 8.
2 Empirical Bayes Estimators

Suppose there are \( m \) areas with counts \( y_i \) for \( i \)-th area. We assume \( y_i \) are independently distributed from \( Poisson(\lambda_i) \) for \( i \)-th area. We transform \( z_i = \sqrt{y_i} \) and with the usual variance stabilizing square root transformation so that \( V(z_i) \) is approximately 1/4. We begin with

\[
z_i | \theta_i \overset{ind}{\sim} N(\theta_i, 1/4) \text{ where } \theta_i = \sqrt{\lambda_i} \quad (i = 1, ..., m).
\]

Following the customary approach, we consider independent \( N(x_i^\top \beta, A) \) priors for the \( \theta_i \) with \( p \)-dimensional auxiliary variables \( x_i \) and regression parameter \( \beta \in \mathbb{R}^p \) where \( m > p + 4 \). The posterior \( \theta_i | z_i \overset{ind}{\sim} N((1 - B)z_i + Bx_i^\top \beta, (1 - B)/4) \) where \( B = \frac{1/4}{1/4 + A} \). Thus, Bayes estimator of \( \lambda_i \) is

\[
\hat{\lambda}_i^B = E(\lambda_i | z_i) = E(\theta_i^2 | z_i) = (1 - B)/4 + \{(1 - B)z_i + Bx_i^\top \beta \}^2
\]

We now turn towards empirical Bayes (EB) estimation of the \( \lambda_i \). Writing \( X = (x_1, ..., x_m)^\top, Z = (z_1, ..., z_m)^\top \) and \( \hat{\beta} = (X^\top X)^{-1}X^\top Z \), it follows that marginally \( \|Z - X\hat{\beta}\|^2 \sim \frac{1}{4B} \chi^2_{m-p} \). Here \( X \) is a \( m \times p \) matrix with rank \( p \). Following Efron and Morris (1973), an EB estimator of \( B \) is \( \hat{B} = \frac{m-p-2}{4\|Z - X\hat{\beta}\|^2} \). Thus an EB estimator of \( \lambda_i \) is

\[
\hat{\lambda}_i^{EB} = (1 - \hat{B})/4 + \{(1 - \hat{B})z_i + \hat{B}x_i^\top \hat{\beta} \}^2
\]

For proving our technical results, we find it convenient also to define

\[
\tilde{\hat{\lambda}}_i^{EB} = (1 - \hat{B})/4 + \{(1 - \hat{B})z_i + Bx_i^\top \hat{\beta} \}^2,
\]

which is also an EB estimator of \( \lambda_i \) if the shrinkage factor \( B \) were known. Also, for notational simplicity we write \( m_0 = m - p \) hereafter.

3 Bias of \( \hat{\lambda}^{EB} \)

For both bias and mean squared error (MSE) calculations for \( \hat{\lambda}_i^{EB} \) we need the following lemmas.

**Lemma 1.** Let \( \begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right) \). Then

\[
\begin{align*}
Var(X^2) &= 2\sigma_1^4 + 4\mu_1^2\sigma_1^2 \\
Var(Y^2) &= 2\sigma_2^4 + 4\mu_2^2\sigma_2^2 \\
Cov(X^2, Y^2) &= 2\sigma_{12}^2 + 4\mu_1\mu_2\sigma_{12}
\end{align*}
\]
Lemma 2. Define \( s_i = x_i^\top (X^\top X)^{-1}x_i \), \( u_{i1} = (1 - B)z_i + Bx_i^\top \hat{\beta} \) and \( u_{i2} = (1 - B)z_i + Bx_i^\top \beta \) then,

\[
\begin{bmatrix}
u_{i1} \\
u_{i2}
\end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} x_i^\top \beta \\
x_i^\top \beta \end{bmatrix}, \begin{bmatrix} A(1-B) + (2-B)s_i/4 & (1-B)(A + s_i/4) \\
(1-B)(A + s_i/4) & A(1-B) \end{bmatrix} \right). \tag{3.1}
\]

Proof. The result follows from the independence of \( \hat{\beta} \) and \( z_i - x_i^\top \hat{\beta} \) and noting that \( \hat{\beta} \sim N(\beta, (1/4 + A)(X^\top X)^{-1}) \) while \( z_i - x_i^\top \hat{\beta} \sim N(0, (1/4 + A)(1-s_i)) \).

Lemma 3. \((i)\) \( \|Z - X\hat{\beta}\| \) is distributed independently of \( \begin{bmatrix} (z_1 - x_1^\top \hat{\beta})/\|Z - X\hat{\beta}\|, \ldots, (z_m - x_m^\top \hat{\beta})/\|Z - X\hat{\beta}\| \end{bmatrix} \).

\((ii)\) \( E \left( \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^k} \right) = \frac{E((z_i - x_i^\top \hat{\beta})^k)}{E(\|Z - X\hat{\beta}\|^k)} \) for all positive integers \( k \).

\((iii)\) \( E \left( \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^l} \right) = 0 \) if \( k \) is an odd positive integer and \( 0 < l < k \).

Proof. Marginally, \( Z \sim N(X\beta, (1/4 + A)I_m) \). Hence, \( (\hat{\beta}, \|Z - X\hat{\beta}\|) \) is complete sufficient for \( (\beta, A) \), while \( \begin{bmatrix} (z_1 - x_1^\top \hat{\beta})/\|Z - X\hat{\beta}\|, \ldots, (z_m - x_m^\top \hat{\beta})/\|Z - X\hat{\beta}\| \end{bmatrix} \) is ancillary. This proves (i) by an application of Basu’s Theorem (Basu, 1955).

Now for any positive integer \( k \), and by part (i) in Lemma 3, we have,

\[
E(z_i - x_i^\top \hat{\beta})^k = E \left[ \|Z - X\hat{\beta}\|^k \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^k} \right] = E \left[ \|Z - X\hat{\beta}\|^k \right] E \left( \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^k} \right).
\]

This leads to \( E \left( \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^k} \right) = \frac{E(z_i - x_i^\top \hat{\beta})^k}{E(\|Z - X\hat{\beta}\|^k)} \).

Next noting that \((z_1 - x_1^\top \hat{\beta}), \ldots, (z_m - x_m^\top \hat{\beta}) \equiv -((z_1 - x_1^\top \hat{\beta}), \ldots, (z_m - x_m^\top \hat{\beta})) \) it follows that

\[
\begin{bmatrix}
(\hat{z}_1 - x_1^\top \hat{\beta})/\|Z - X\hat{\beta}\|, \\
(\hat{z}_m - x_m^\top \hat{\beta})/\|Z - X\hat{\beta}\|
\end{bmatrix} \doteq \begin{bmatrix}
-(\hat{z}_1 - x_1^\top \hat{\beta})/\|Z - X\hat{\beta}\|, \\
-(\hat{z}_m - x_m^\top \hat{\beta})/\|Z - X\hat{\beta}\|
\end{bmatrix}
\]

Therefore, \( \frac{(z_i - x_i^\top \hat{\beta})}{\|Z - X\hat{\beta}\|} \) is symmetric random variable around 0 and its all odd moments are 0. Hence,

\[
E \left( \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^l} \right) = E \left[ \|Z - X\hat{\beta}\|^{k-l} \frac{(z_i - x_i^\top \hat{\beta})^k}{\|Z - X\hat{\beta}\|^k} \right]
\]
\[ E(\sqrt{|Z - X\hat{\beta}|}) = E\left(|Z - X\hat{\beta}|^{k-l}\right) E\left[\frac{(z_i - x_i^T \hat{\beta})^k}{|Z - X\hat{\beta}|^k}\right] \]

We get the second equality in the above equations using part (i) of this Lemma. Proof of (iii) is complete after observing \( k \) is an odd integer.

**Lemma 4.**

(i) \( E(\hat{B}) = B \) if \( m_0 > 2 \).

(ii) \( E(\hat{B}^2) = \frac{m_0 - 2}{m_0 - 4} B^2 \) if \( m_0 > 4 \).

(iii) \( E(1/\hat{B}) = \frac{m_0}{(m_0 - 2)B} \) if \( m_0 > 2 \).

(iv) \( E(1/\hat{B}^2) = \frac{m_0(m_0 + 2)}{(m_0 - 2)^2 B^2} \) if \( m_0 > 2 \).

(v) \[ E\left[\frac{(z_i - x_i^T \hat{\beta})^2}{|Z - X\hat{\beta}|^2}\right] = \frac{1 - s_i}{m_0}. \]

(vi) \[ E\left[\frac{(z_i - x_i^T \hat{\beta})^4}{|Z - X\hat{\beta}|^4}\right] = \frac{3(1 - s_i)^2}{m_0(m_0 + 1)} . \]

**Proof.** The proof (i) to (iv) follows by noting that \( \hat{B} = \frac{m_0 - 2}{4|Z - X\hat{\beta}|^2} \) and \( |Z - X\hat{\beta}|^2 \sim \frac{1}{4B}\lambda_m^2 \). To prove (v) and (vi) we also need part (ii) of Lemma 3.

Now we start with calculation of the bias \( E(\hat{\lambda}_i^{EB} - \lambda_i) \). The following theorem is proved.

**Theorem 1.** Suppose \( z_i|\theta_i \sim N(\theta_i, 1/4) \) with priors \( \theta_i \sim N(x_i^T \beta, A) \) and \( m_0 > 2 \). Then bias of the EB estimator \( \hat{\lambda}_i^{EB} \) in Eq. 2.2 for \( \lambda_i = \theta_i^2 \) is given by

\[ \text{Bias}(\hat{\lambda}_i^{EB}) = E(\hat{\lambda}_i^{EB} - \lambda_i) = \frac{2 - B}{4} \left( s_i + \frac{2(1 - s_i)}{m_0} \right). \] (3.2)

**Proof.** We begin with the partition

\[ E(\hat{\lambda}_i^{EB} - \lambda_i) = E(\hat{\lambda}_i^{B} - \lambda_i) + E(\bar{\lambda}_i^{EB} - \hat{\lambda}_i^{B}) + E(\hat{\lambda}_i^{EB} - \bar{\lambda}_i^{EB}) \] (3.3)

By Lemmas 1 and 2,

\[ E(\bar{\lambda}_i^{EB} - \bar{\lambda}_i^{B}) = E(u_{i1}^2 - u_{i2}^2) = Var(u_{i1}) + (x_i\beta)^2 - Var(u_{i2}) - (x_i\beta)^2 = (2 - B)s_i/4 \] (3.4)

Noticing \( E(\lambda_i) = E(E(\lambda_i|z_i)) = E(\hat{\lambda}_i^{B}) \), we have \( E(\bar{\lambda}_i^{B} - \lambda_i) = 0 \). It is easy to see that

\[ \hat{\lambda}_i^{EB} - \bar{\lambda}_i^{EB} = (B - \hat{B})/4 + 2x_i^T \hat{\beta}(B - \hat{B})(z_i - x_i^T \hat{\beta}) \]
\[ + \{(\hat{B}^2 - B^2) - 2(\hat{B} - B)\}(z_i - x_i^\top \hat{\beta})^2 \]  
\[ = A_1 + A_2 + A_3, \]

where \( A_1 = (B - \hat{B})/4, A_2 = 2x_i^\top \beta (B - \hat{B}) \) and \( A_3 = \{(\hat{B}^2 - B^2) - 2(\hat{B} - B)\}(z_i - x_i^\top \hat{\beta})^2 \). The expectation of \( A_1 \) is 0 since \( \hat{B} = \frac{m_0 - 2}{4\|Z - X\beta\|^2} \) is unbiased estimator of \( B \). Now,

\[ E(A_2) = E(x_i^\top \hat{\beta})E[(B - \hat{B})(z_i - x_i^\top \hat{\beta})] \]
\[ = E(x_i^\top \hat{\beta})E(B(z_i - x_i^\top \hat{\beta})) - E(x_i^\top \hat{\beta})E\left[\frac{m_0 - 2}{4\|Z - X\hat{\beta}\|^2}(z_i - x_i^\top \hat{\beta})\right] \]
\[ = 0 \]  
\[ (3.6) \]

The first equality in Eq. 3.6 is by independence of \( \hat{\beta} \) and \((z_1 - x_1^\top \hat{\beta}), \ldots, (z_n - x_n^\top \hat{\beta})\)). The third equality holds by part (iii) of Lemma 3 since \((z_i - x_i^\top \hat{\beta}) \sim N(0, (1/4 + A)(1 - s_i))\).

Finally, we simplify \( A_3 \). By (i) and (iii) of Lemma 3 and Lemma 4,

\[ E(\hat{\lambda}_i^{EB} - \bar{\lambda}_i^{EB}) = E[(\hat{B}^2 - B^2) - 2(\hat{B} - B)](z_i - x_i^\top \hat{\beta})^2] \]
\[ = E\left[\frac{(z_i - x_i^\top \hat{\beta})^2}{\|Z - X\hat{\beta}\|^2}\right] E\left[\{(\hat{B}^2 - B^2) - 2(\hat{B} - B)\}\frac{m_0 - 2}{4\hat{B}}\right] \]
\[ = \frac{(1 - s_i)}{m_0} \frac{(m_0 - 2)}{4} E(\hat{B}^2 - B^2/\hat{B} - 2 + 2B/\hat{B}) \]
\[ = \frac{(1 - s_i)}{m_0} \frac{(m_0 - 2)}{4} \left[ B - B - \frac{m_0}{m_0 - 2} - 2 + 2\frac{m_0}{m_0 - 2} \right] \]
\[ = \frac{1 - s_i}{2m_0}(2 - B). \]  
\[ (3.7) \]

The proof of Eq. 3.2 follows now by combining (3.3) with Eqs. 3.4–3.7.

Remark 1. With the usual assumption, \( s_i = x_i^\top (X^\top X)^{-1}x_i = O(m^{-1}) \) for large \( m \), the bias of the EB estimator \( \hat{\lambda}_i^{EB} \), \( E(\hat{\lambda}_i^{EB} - \lambda_i) = O(m^{-1}) \).

Remark 2. We can estimate the bias in Eq. 3.2 by replacing the \( B \) by \( \hat{B} \). An unbiased estimator of the bias is

\[ \hat{\text{bias}} = \frac{2 - \hat{B}}{4} \left( s_i + \frac{2(1 - s_i)}{m_0} \right). \]  
\[ (3.8) \]

Thus, from Eq. 3.8, the EB estimator has positive bias, and the bias-corrected estimator of \( \lambda_i \) is

\[ \hat{\lambda}_i^{CEB} = \hat{\lambda}_i^{EB} - \hat{\text{bias}}. \]
4 MSE of $\hat{\lambda}^{EB}$

The following theorem provides an exact expression for the MSE of $\hat{\lambda}^{EB}$.

**Theorem 2.** Suppose $z_i|\theta_i \sim N(\theta_i, 1/4)$ with priors $\theta_i \sim N(x_i^\top \beta, A)$ and $m_0 > 4$. Then MSE of the EB estimator $\hat{\lambda}_i^{EB}$ in Eq. 2.2 for $\lambda_i = \theta_i^2$ is given by

$$MSE(\hat{\lambda}_i^{EB}) = E(\hat{\lambda}_i^{EB} - \lambda_i)^2$$

$$= (1 - B) \left[ (x_i^\top \beta)^2 + \frac{(1 - B)(2 - B)}{8B} \right] + B s_i (x_i^\top \beta)^2 + (1 - B)^2 s_i/4 + s_i^2 (1/2 - B/4 - B^2/16)$$

$$+ \frac{B^2}{8(m_0 - 4)} + \frac{1 - s_i}{2m_0} \left[ (x_i^\top \beta)^2 B + s_i \right] + 3 (1 - s_i)^2 \left[ \frac{2m_0^2 - 9m_0 + 6}{4m_0(m_0 + 2)(m_0 - 4)} B^2 - \frac{B}{m_0 + 2} + \frac{1}{2m_0} \right]$$

$$+ \frac{(1 - s_i)B}{2m_0} \left[ 1 - \frac{B(m_0 - 3)}{m_0 - 4} \right] + \frac{s_i(1 - s_i)}{m_0} (2 - 2B + B^2/4).$$

The proof of Theorem 2 is given in Appendix A.

**Remark 3.** Theorem 2 shows that the MSE of EB estimator $\hat{\lambda}_i^{EB}$, $E(\hat{\lambda}_i^{EB} - E(\lambda_i))^2 = O(1)$ due to the first term in Eq. 4.1 for large $m$.

5 Estimation of the MSE of $\hat{\lambda}^{EB}$

In this section we estimate the MSE of $\hat{\lambda}_i^{EB}$ provided in Theorem 2 up to the order $O(m^{-1})$ for large $m$. We now assume that $s_i = O(m^{-1})$. Ignoring the $O(m^{-2})$ terms, we rewrite

$$MSE(\hat{\lambda}_i^{EB}) = (1 - B)(x_i^\top \beta)^2 + \frac{(1 - B)^2(2 - B)}{8B} + s_i \left[ B(x_i^\top \beta)^2 + \frac{(1 - B)^2}{4} \right]$$

$$+ \frac{B^2}{8m} + \frac{2B(x_i^\top \beta)^2}{m} + \frac{3(1 - B)^2}{2m} + \frac{B(1 - B)}{2m} + O(m^{-2}).$$

It is easy to see that only first two terms in Eq. 5.1 do not depend on $m$ and remaining terms are of $O(m^{-1})$. Using Lemma 4 we get

$$E \left( \frac{1}{B} - \frac{2}{m B} \right) = \frac{1}{B} + O(m^{-2})$$

$$E \left( \hat{B}^2 - \frac{2}{m} \hat{B}^2 \right) = B^2 + O(m^{-2}).$$
Using Eq. 5.1, we find
\[
E \left[ \frac{(1 - \hat{B})^2(2 - \hat{B}) - 4}{m \hat{B}} + 2 \hat{B}^2 \right] = \left[ \frac{(1 - B)^2(2 - B)}{B} \right] + O(m^{-2}). \tag{5.2}
\]
Now by \(x_i^\top \hat{\beta} \sim N(x_i^\top \beta, \frac{s_i}{4\hat{B}})\), Lemma 4 and the independence of \(\hat{\beta}\) and \(\hat{B}\), we also have
\[
E[(1 - \hat{B})(x_i^\top \hat{\beta})^2 - s_i/4\hat{B}] = (1 - B)(x_i^\top \beta)^2 + O(m^{-2}). \tag{5.3}
\]

Since we are ignoring \(O(m^{-2})\) terms in MSE estimation, we can estimate the \(O(m^{-1})\) terms in Eq. 5.1 simply by replacing \(B^2\) by \(\hat{B}^2\) and \((x_i^\top \beta)^2\) by \((x_i^\top \hat{\beta})^2\). By Eqs. 5.2 and 5.3, we derive estimator of the MSE of \(\hat{\lambda}_i^{EB}\) in Theorem 3.

**Theorem 3.** Assume conditions of Theorem 2. Then \(MSE(\hat{\lambda}_i^{EB})\) as given in (13) is estimated by
\[
\widehat{MSE}(\hat{\lambda}_i^{EB}) = (1 - \hat{B})(x_i^\top \hat{\beta})^2 + \left[ \frac{(1 - \hat{B})^2(2 - \hat{B})}{8\hat{B}} \right] + s_i \left[ \hat{B}(x_i^\top \hat{\beta})^2 + \frac{(1 - \hat{B})^2}{4} - \frac{(1 - \hat{B})}{4\hat{B}} \right] + \frac{3\hat{B}^2}{8m} + \frac{2\hat{B}}{m}(x_i^\top \hat{\beta})^2 + \frac{3}{2m}(1 - \hat{B})^2 + \frac{\hat{B}(1 - \hat{B})}{2m} - \frac{1}{2m\hat{B}} + O_p(m^{-2}).
\]

The next theorem shows that MSE of the bias-corrected estimator \(\hat{\lambda}_i^{CEB}\) equals \(MSE(\hat{\lambda}_i^{EB}) + O(m^{-2})\). In other words, bias correction does not lead to any significant improvement over \(MSE(\hat{\lambda}_i^{EB})\), at least, when calculated up to \(O(m^{-1})\).

**Theorem 4.** Assume conditions of Theorem 2. Then
\[
E[(\hat{\lambda}_i^{CEB} - \lambda_i)^2] = E[(\hat{\lambda}_i^{EB} - \lambda_i)^2] + O(m^{-2}).
\]

**Proof.** Write \(\widehat{bias} = \frac{(2 - \hat{B})}{4}d_i\), where \(d_i = s_i + 2\frac{1 - s_i}{m_0} = O(m^{-1})\). We begin with
\[
E[(\hat{\lambda}_i^{CEB} - \lambda_i)^2] = V[\hat{\lambda}_i^{CEB} - \lambda_i] = V[\hat{\lambda}_i^{EB} - \lambda_i] + V[\widehat{bias}] - 2Cov(\hat{\lambda}_i^{EB} - \lambda_i, \widehat{bias}). \tag{5.4}
\]
It is immediate that \(V[\hat{\lambda}_i^{EB} - \lambda_i] = E[(\hat{\lambda}_i^{EB} - \lambda_i)^2] - (bias)^2\), where \((bias)^2 = \frac{1}{16}(2 - B)^2d_i^2 = O(m^{-2})\). Next
\[
V[\widehat{bias}] = E[\hat{B}^2] - B^2 = E \left[ \frac{(m_0 - 2)^2}{16||Z - X\hat{\beta}||^4} \right] - B^2.
\]
Hence Eq. 5.7 reduces to
\[
\frac{(m_0 - 2)^2}{16} \frac{16B^2}{(m_0 - 2)(m_0 - 4)} - B^2 = \frac{2B^2}{m_0 - 4}.
\]

Hence,
\[
V[\text{bias}] = \frac{d_i^2}{8} \frac{B^2}{m_0 - 4} = \mathcal{O}(m^{-3}). \tag{5.5}
\]

Finally, \(\text{Cov}(\hat{\lambda}_i^{EB} - \lambda_i, \text{bias}) = \text{Cov}(\hat{\lambda}_i^{EB}, \text{bias}) - \text{Cov}(\lambda_i, \text{bias})\), But
\[
\text{Cov}(\lambda_i, \text{bias}) = E[\lambda_i \text{bias}] - E[\lambda_i]E[\text{bias}],
\]
\[
= \text{Cov}(\hat{\lambda}_i^2, \text{bias}),
\]
\[
= \text{Cov}((1 - B)z_i + Bx_i^\top \hat{\beta})^2, \text{bias}),
\]
\[
= \frac{(1 - s_i)}{8m_0} (1 - B)^2 d_i = \mathcal{O}(m^{-2}). \tag{5.6}
\]

Thus, \(\text{Cov}(\hat{\lambda}_i^{EB} - \lambda_i, \text{bias}) = \text{Cov}(\hat{\lambda}_i^{EB}, \text{bias}) + \mathcal{O}(m^{-2})\). Now,
\[
\text{Cov}(\hat{\lambda}_i^{EB}, \text{bias}) = \text{Cov}\left(\frac{1 - \hat{B}}{4} + ((1 - \hat{B})z_i + \hat{B}x_i^\top \hat{\beta})^2, \frac{2 - \hat{B}}{4} d_i\right),
\]
\[
= \frac{d_i}{16} V[\hat{B}] - \frac{d_i}{4} \text{Cov}\left(((1 - \hat{B})z_i + \hat{B}x_i^\top \hat{\beta})^2, \hat{B}\right),
\]
\[
= \frac{B^2}{8(m_0 - 4)}d_i - \frac{d_i}{4} \text{Cov}\left(x_i^\top \hat{\beta}\right)^2 + 2(1 - \hat{B})(z_i - x_i^\top \hat{\beta})x_i^\top \hat{\beta}
\]
\[
+ (1 - \hat{B})^2(z_i - x_i^\top \hat{\beta})^2, \hat{B}\right). \tag{5.7}
\]

Due to the independence of \(\hat{\beta}\) and \(Z - X\hat{\beta}\), \(\text{Cov}(x_i^\top \hat{\beta})^2, \hat{B}\) = 0. Next, again invoking the symmetry of \(Z - X\beta\) around 0, and \(E(z_i - x_i^\top \hat{\beta}) = 0\),
\[
\text{Cov}((1 - \hat{B})(z_i - x_i^\top \hat{\beta})x_i^\top \hat{\beta}, \hat{B}) = E[\{(1 - \hat{B})(z_i - x_i^\top \hat{\beta})x_i^\top \hat{\beta}\hat{B}]
\]
\[
- E[\{(1 - \hat{B})(z_i - x_i^\top \hat{\beta})x_i^\top \hat{\beta}\hat{B}]}E[\hat{B}],
\]
\[
= 0.
\]

Hence Eq. 5.7 reduces to
\[
\text{Cov}(\hat{\lambda}_i^{EB}, \text{bias}) = \frac{-d_i}{4} \text{Cov}((1 - 2\hat{B} + \hat{B}^2)(z_i - x_i^\top \hat{\beta})^2, \hat{B}) + \mathcal{O}(m^{-2}), \tag{5.8}
\]
\[
\text{Cov}((z_i - x_i^\top \hat{\beta})^2, \hat{B}) = E\left[\frac{(m_0 - 2)}{4} \frac{(z_i - x_i^\top \hat{\beta})^2}{||Z - X\hat{\beta}||^2}\right] - E[(z_i - x_i^\top \hat{\beta})^2]B.
\]
\[
= \frac{1 - s_i}{2m_0}.
\]

Next
\[
\text{Cov}(\hat{B}(z_i - x_i^T \hat{\beta})^2, \hat{B}) = \text{Cov}\left(\frac{(m_0 - 2) (z_i - x_i^T \hat{\beta})^2}{4||Z - X\hat{\beta}||^2}, \frac{(m_0 - 2)}{4||Z - X\hat{\beta}||^2}\right) = 0
\]

Since \(\frac{(z_i - x_i^T \hat{\beta})^2}{||Z - X\hat{\beta}||^2}\) is ancillary, and \(||Z - X\hat{\beta}||^2\) is a function of the complete sufficient statistic \((\hat{\beta}, Z - X\hat{\beta})\), by Basu’s Theorem,
\[
\text{Cov}(\hat{B}^2(z_i - x_i^T \hat{\beta})^2, \hat{B}) = E[\hat{B}^3(z_i - x_i^T \hat{\beta})^2] - E[\hat{B}^2(z_i - x_i^T \hat{\beta})^2]B
\]
\[
= \frac{(m_0 - 2)^3}{64} E\left[\frac{(z_i - x_i^T \hat{\beta})^2}{||Z - X\hat{\beta}||^2} \cdot \frac{1}{||Z - X\hat{\beta}||^4}\right]
\]
\[
- \frac{(m_0 - 2)^2}{16} E\left[\frac{(z_i - x_i^T \hat{\beta})^2}{||Z - X\hat{\beta}||^2} \cdot \frac{1}{||Z - X\hat{\beta}||^2}\right],
\]
\[
= \frac{(m_0 - 2)(1 - s_i)B^2}{2m_0(m_0 - 4)}
\]

Hence, from Eqs. 5.7–5.10, \(\text{Cov}(\hat{\lambda}_i^{EB}, \hat{\text{bias}}) = O(m^{-2})\). This along with Eqs. 5.4–5.6 proves the theorem.

6 Data Analysis

In this section, we now deploy our approach on the 2020 COVID-19 pandemic dataset, which is available at usafacts.org. This example is used mainly for illustration. We are using the figures provided as the sampled estimates. Our study shows that the coefficient of determination (\(R^2\)) does not increase much if we include other demographic variables such as the population size, number of people over age 60, and income in the linear model for the number of deaths regressing on number of confirmed cases. It suggests that the number of confirmed cases is the most crucial variable in estimation of the number of deaths by Coronavirus than the aforementioned demographic variables. We have also studied a few more county-level data sources\(^1\) and we found out that adjusted gross income (AGI)\(^2\) of the year 2017 is really relevant for estimating the number of deaths. In our model, we have transformed the number of confirmed cases and adjusted

\(^1\)US Census Bureau and Statistics of Income Division (SOI) of the IRS
\(^2\)https://www.irs.gov/pub/irs-soi/17incyallagi.csv
gross income (AGI) by taking the square root. All data are aggregated at the county level. We are interested in estimating the counts of death due to Coronavirus for all counties in Florida. Here $m = 57$ since Florida has 57 counties. From Section 2 we know that $\sqrt{y_i} \sim N(x_i^\top \beta, A + 0.25)$ and we estimate $\beta$ by ordinary least square method. Based on our analysis, we get $\hat{\beta} = (-0.2786, 0.0917, 0.0003)^\top$, the respective estimates for the intercept, number of confirmed cases and AGI. We have summarized our results based on our model in Table 1 and the shrinkage factor $\hat{B} = 0.3777$. It seems that our model-based approach seems to pull the direct estimates towards some grand average, as one anticipates in a typical EB analysis. Figure 1 shows that the estimates are higher in south east Florida than the rest of the state.

7 Simulation

In this section, we will measure the performance of our model via a simulation study. The choice of the auxiliary parameters is guided by the case study of the previous section. For illustration purposes, we have considered only one covariate— the number of confirmed cases to estimate the number of deaths due to COVID 19. This data is available for 3,142 counties of the United States. For simulation purposes, we have taken a random sample from this data without any replacement for each choice of $m$. The number of small areas (counties), $m$, is set to be 25, 50, 100, 200, 500, or 1000. For each choice of $m$, we generated data from the model:

$$z_i | \theta_i \sim N(\theta_i, 1/4), \quad \theta_i = \sqrt{\lambda_i} \sim N(x_i^\top \beta, A)$$

The design matrix $X$ includes a column of ones and one explanatory variable. To set the value of the parameter for $\beta$ and $A$, we first create a linear regression model for the number of deaths on the number of confirmed cases using entire data for 3,142 counties. The estimated value for regression coefficient vector $\beta$ is $(5.281570, 0.000272)^\top$ and mean square residuals is 22.75. For simulation we set $\beta = (5.281570, 0.000272)^\top$ and $A = 22.75 - 0.25 = 22.50$, hence shrinkage factor $B = 0.011$. Due to this variance stabilizing transformation, the shrinkage factor does not change between counties. Now using Eq. 7.1 we generate $\lambda_i$ and $z_i$ for all $i = (1, \ldots, m)$. The explanatory variable is again number of confirmed cases which is simulated randomly without replacement from the entire populations of 3,142 counties in the United States.

Here we will compare the true RMSE and estimated RMSE of $\lambda_i^{EB}$. We examine our findings in Theorems 2, 3 and 4 based on six different settings for $m$. Here we will vary the $m$ and only one dataset is generated for each $m$, the latter taking values 25, 50, 100, 200, 500, 1000. We have estimated the
Table 1: COVID-19 in Florida as of December 28, 2020

| County name | Confirm | AGI | Deaths | EB | BIAS | Corrected BIAS | RMSE | Death rate | Estimated rate |
|-------------|---------|-----|--------|----|------|----------------|------|------------|----------------|
| Alachua     | 15473   | 7.294066 | 128 | 128.55 | 0.023 | 128.527 | 12.82 | 4.76 | 4.78 |
| Baker       | 2414    | 0.554345 | 36  | 36.68  | 0.025 | 36.655  | 9.37  | 12.32 | 12.56 |
| Bay         | 11403   | 4.571565 | 203 | 201.86 | 0.024 | 201.836 | 11.48 | 11.62 | 11.55 |
| Bradford    | 2118    | 0.478624 | 23  | 23.75  | 0.025 | 23.725  | 9.32  | 8.16  | 8.42 |
| Brevard     | 20355   | 17.959964 | 513 | 509.65 | 0.027 | 509.623 | 17.42 | 8.52  | 8.47 |
| Broward     | 133480  | 68.207212 | 1828 | 1832.21 | 0.092 | 1832.118 | 45.89 | 9.36  | 9.38 |
| Calhoun     | 1169    | 0.207429 | 28  | 28.70  | 0.025 | 28.675  | 9.16  | 19.85 | 20.35 |
| Charlotte   | 7128    | 5.233347 | 231 | 229.35 | 0.023 | 229.327 | 11.46 | 12.23 | 12.14 |
| Citrus      | 6467    | 3.447436 | 260 | 257.45 | 0.024 | 257.426 | 10.73 | 17.37 | 17.20 |
| Clay        | 10829   | 6.127044 | 183 | 182.44 | 0.023 | 182.417 | 12.05 | 8.35  | 8.32 |
| Collier     | 22004   | 27.616972 | 328 | 329.86 | 0.048 | 329.812 | 21.46 | 8.52  | 8.57 |
| Columbia    | 5837    | 1.304030 | 117 | 116.71 | 0.025 | 116.685 | 9.86  | 16.32 | 16.28 |
| DeSoto      | 2843    | 0.527165 | 52  | 52.52  | 0.025 | 52.495  | 9.38  | 13.68 | 13.82 |
| Dixie       | 1088    | 0.540883 | 11  | 11.73  | 0.025 | 11.705  | 9.28  | 6.54  | 6.97 |
| Duval       | 58491   | 28.523854 | 722 | 720.17 | 0.030 | 720.140 | 24.32 | 7.54  | 7.52 |
| Escambia    | 21507   | 8.068327 | 360 | 357.02 | 0.024 | 356.996 | 13.52 | 11.31 | 11.22 |
| Flagler     | 3565    | 3.411798 | 48  | 48.81  | 0.024 | 48.786  | 10.53 | 4.17  | 4.24 |
| Franklin    | 878     | 0.283821 | 4   | 4.61   | 0.025 | 4.585   | 9.17  | 3.30  | 3.80 |
| Gadsden     | 3836    | 0.824111 | 59  | 59.48  | 0.025 | 59.455  | 9.56  | 12.92 | 13.03 |
| Gilchrist   | 984     | 0.328444 | 23  | 23.73  | 0.025 | 23.705  | 9.19  | 12.38 | 12.77 |
| County name | Confirm | AGI$^1$ | Deaths$^2$ | EB$^3$ | BIAS | Corrected BIAS | RMSE$^5$ | Death$^6$ | Estimated rate $^7$ |
|-------------|---------|---------|-----------|--------|------|----------------|----------|-----------|-------------------|
| Glades      | 762     | 0.232870| 11        | 11.72  | 0.025| 11.695         | 9.14     | 7.96      | 8.49              |
| Gulf        | 1273    | 0.321757| 27        | 27.71  | 0.025| 27.685         | 9.21     | 19.80     | 20.32             |
| Hamilton    | 1224    | 0.190548| 12        | 12.73  | 0.025| 12.705         | 9.15     | 8.32      | 8.82              |
| Hardee      | 2057    | 0.417308| 20        | 20.75  | 0.025| 20.725         | 9.29     | 7.42      | 7.70              |
| Hendry      | 3235    | 0.709752| 49        | 49.57  | 0.025| 49.545         | 9.48     | 11.66     | 11.80             |
| Hernando    | 7074    | 3.958244| 264       | 261.49 | 0.024| 261.466        | 10.96    | 13.61     | 13.48             |
| Highlands   | 4822    | 1.882105| 199       | 197.30 | 0.025| 197.275        | 10.02    | 18.73     | 18.57             |
| Hillsborough| 74788   | 46.284944| 1064     | 1064.35| 0.053| 1064.297       | 32.76    | 7.23      | 7.23              |
| Holmes      | 1603    | 0.280332| 24        | 24.73  | 0.025| 24.705         | 9.21     | 12.23     | 12.61             |
| Indian      | 6553    | 7.607307| 156       | 155.99 | 0.024| 155.966        | 12.36    | 9.75      | 9.75              |
| Jackson     | 4502    | 0.819253| 115       | 114.65 | 0.025| 114.625        | 9.60     | 24.78     | 24.70             |
| Jefferson   | 957     | 0.309690| 16        | 16.74  | 0.025| 16.715         | 9.18     | 11.23     | 11.75             |
| Lafayette   | 1414    | 0.106457| 21        | 21.73  | 0.025| 21.705         | 9.13     | 24.93     | 25.80             |
| Lake        | 14841   | 9.534133| 293       | 291.32 | 0.023| 291.297        | 13.67    | 7.98      | 7.94              |
| Lee         | 39332   | 27.089733| 652     | 649.84 | 0.031| 649.809        | 22.42    | 8.46      | 8.43              |
| Leon        | 18742   | 7.918907| 173       | 173.09 | 0.023| 173.067        | 13.28    | 5.89      | 5.90              |
| Levy        | 1798    | 0.721669| 19        | 19.77  | 0.025| 19.745         | 9.39     | 4.58      | 4.76              |
| Liberty     | 716     | 0.113382| 14        | 14.73  | 0.025| 14.705         | 9.09     | 16.76     | 17.63             |
| Madison     | 1516    | 0.324513| 33        | 33.68  | 0.025| 33.655         | 9.22     | 17.84     | 18.21             |
| Manatee     | 21539   | 13.804729| 412     | 409.42 | 0.023| 409.397        | 15.82    | 10.22     | 10.15             |
| Marion      | 17125   | 7.971880| 456       | 450.97 | 0.023| 450.947        | 13.20    | 12.47     | 12.34             |
| County name | Confirm | AGI$^1$ | Deaths$^2$ | EB$^3$ | BIAS | Corrected BIAS | RMSE$^5$ | Death rate$^6$ | Estimated rate$^7$ |
|-------------|---------|---------|-----------|-------|------|----------------|---------|---------------|------------------|
| Martin      | 7656    | 9.243197| 207       | 206.49| 0.024| 206.466        | 13.08   | 12.86         | 12.83            |
| Miami-Dade  | 290363  | 92.544722| 4155      | 4160.13| 0.427| 4159.703        | 67.08   | 15.29         | 15.31            |
| Monroe      | 4168    | 5.227991| 35        | 36.04 | 0.024| 36.016         | 11.27   | 4.72          | 4.86             |
| Nassau      | 4521    | 3.389901| 66        | 66.65 | 0.024| 66.626         | 10.58   | 7.45          | 7.52             |
| Okaloosa    | 12407   | 6.812245| 224       | 222.92| 0.023| 222.897        | 12.43   | 10.63         | 10.58            |
| Okeechobee  | 2347    | 0.720267| 51        | 51.54 | 0.025| 51.515         | 9.42    | 12.09         | 12.22            |
| Orange      | 73691   | 41.729653| 732      | 735.09| 0.044| 735.046        | 30.80   | 5.25          | 5.28             |
| Osceola     | 24512   | 7.340821| 286       | 284.32| 0.026| 284.294        | 13.43   | 7.61          | 7.57             |
| Palm        | 80865   | 84.441622| 1866     | 1873.59| 0.279| 1873.311       | 49.08   | 12.47         | 12.52            |
| Pasco       | 21222   | 13.333365| 360      | 358.20| 0.023| 358.177        | 15.61   | 6.50          | 6.47             |
| Pinellas    | 43480   | 35.959433| 1035     | 1029.81| 0.049| 1029.761       | 26.35   | 10.62         | 10.56            |
| Polk        | 35942   | 15.491829| 767      | 758.92| 0.024| 758.896        | 17.47   | 10.58         | 10.47            |
| Putnam      | 3819    | 1.217151| 71        | 71.36 | 0.025| 71.335         | 9.70    | 9.53          | 9.58             |
| County name | Confirm | AGI$^1$ | Deaths$^2$ | EB$^3$ | BIAS | Corrected BIAS | RMSE$^5$ | Death rate | Estimated rate |
|-------------|---------|---------|-----------|-------|------|----------------|---------|------------|----------------|
| St.         | 12481   | 12.966328 | 112       | 113.41 | 0.025 | 113.385        | 14.88   | 4.23       | 4.28           |
| St.         | 13864   | 7.683491  | 393       | 389.01 | 0.023 | 388.987        | 12.87   | 11.97      | 11.85          |
| Santa       | 10575   | 5.192317  | 122       | 122.24 | 0.023 | 122.217        | 11.67   | 6.62       | 6.63           |
| Sarasota    | 17916   | 20.149336 | 502       | 499.32 | 0.033 | 499.287        | 18.14   | 11.57      | 11.51          |
| Seminole    | 17445   | 16.199289 | 311       | 310.49 | 0.026 | 310.464        | 16.51   | 6.59       | 6.58           |
| Sumter      | 4819    | 4.709867  | 118       | 118.11 | 0.024 | 118.086        | 11.11   | 8.91       | 8.92           |
| Suwannee    | 3913    | 0.701966  | 94        | 93.97  | 0.025 | 93.945         | 9.51    | 21.16      | 21.16          |
| Taylor      | 1914    | 0.362501  | 25        | 25.73  | 0.025 | 25.705         | 9.26    | 11.59      | 11.93          |
| Union       | 1416    | 0.217377  | 65        | 65.30  | 0.025 | 65.275         | 9.18    | 42.66      | 42.86          |
| Volusia     | 21365   | 14.096377 | 430       | 427.17 | 0.023 | 427.147        | 15.93   | 7.77       | 7.72           |
| Wakulla     | 2032    | 0.710564  | 21        | 21.77  | 0.025 | 21.745         | 9.40    | 6.22       | 6.45           |
| Walton      | 4707    | 3.317614  | 43        | 43.86  | 0.024 | 43.836         | 10.57   | 5.81       | 5.92           |
| Washington  | 1907    | 0.397700  | 30        | 30.71  | 0.025 | 30.685         | 9.27    | 11.78      | 12.06          |

1 Calculated in millions
2 Number of deaths is our direct estimates
3 Empirical Bayes estimator based on our methodology
4 Corrected BIAS EB = EB - BIAS
5 RMSE = \(\sqrt{MSE}\)
6 Death rate = Deaths/(Population Size)
7 Estimated rate = EB/(Population Size)

This rate is calculated for every 10,000 population
Figure 1: The top two maps compare the EB and BIAS corrected EB for the number of deaths due to COVID-19 in each county of Florida, and the bottom two maps show the BIAS and RMSE of the EB.

true RMSE of $\hat{\lambda}^{CEB}_i$ based on 1,000 simulated samples since we do not have exact expression for RMSE of $\hat{\lambda}^{CEB}_i$.

Figure 2 substantiates that the approximations given in Theorems 3 and 4 are fairly close to the true RMSE. In addition, they also point out one particular small area where the MSE is significantly higher than rest of the small areas.

8 Conclusion

The paper introduces square root transformation of Poisson count data, and attains approximately both normality of the transformed data as well as variance stabilization. In this way, we obtain explicit estimates of bias and MSE for Poisson means. Based on the simulation, it seems that our estimates closely resemble the truth. Data analysis part tells us that estimates are higher on south-east Florida when the model appropriate.
Figure 2: figure compares True RMSE: root of Eq. 4.1, estimated RMSE: root of Eq. 5.4 and simulated RMSE of bias corrected estimators. The RMSE of bias corrected EB estimators ($\hat{\lambda}_{CEB}^i$) is based on 1,000 simulations provided in Fig. 2 and it also verifies the result in Theorem 4.

There are many potential extensions. One that immediately comes to mind is consideration of unit level models with corresponding square root transformation. Gonçalves and Ghosh (2021) have addressed this problem using a pure hierarchical Bayesian framework, but an empirical Bayes approach with all its theoretical properties should also be a topic of future investigation. Even under the present framework, one may add a spatial component and using something like a CAR model (see for example, Ghosh et al., 1999). A final interesting problem is to consider an overdispersed Poisson model, i.e. a negative binomial model for count data with variable transformation as in Yu (2009) which also leads to homoscedasticity.

Acknowledgements. The authors are grateful to the editor and anonymous reviewer(s) for their constructive comments and suggestions which greatly improved an earlier version of this article.
Next we compute the second term in the right side of Eq. 5.11. By Lemmas 1, 2, and 3, each term in the right side of Eq. 5.11 will be computed separately. Since \( E((\hat{\lambda}_i - \lambda_i)(\hat{\lambda}_i - \lambda_i)) = 0 \) since \( E((\hat{\lambda}_i - \lambda_i)|z_i) = 0 \). Thus we have

\[
E(\hat{\lambda}_i^{EB} - \lambda_i)^2 = E(\hat{\lambda}_i^B - \lambda_i)^2 + E(\hat{\lambda}_i^{EB} - \hat{\lambda}_i^B)^2 + E(\hat{\lambda}_i^{EB} - \hat{\lambda}_i^{EB})^2 + 2E((\hat{\lambda}_i^{EB} - \hat{\lambda}_i^{EB})(\hat{\lambda}_i^{EB} - \hat{\lambda}_i^B)) \tag{5.11}
\]

Each term in the right side of Eq. 5.11 will be computed separately. Since \( \theta_i|z_i \sim N((1 - B)z_i + Bx_i^\top \beta, (1 - B)/4) \), by Lemma 1,

\[
E(\hat{\lambda}_i^B - \lambda_i)^2 = E(E(\lambda_i|z_i) - \lambda_i)^2 = E[E(E(\lambda_i) - \lambda_i)^2|z_i] = E[\text{Var}(\lambda_i|z_i)] = E[(1 - B)^2/8 + u_i^2/(1 - B)] = (1 - B)[(1 - B)/8 + (A(1 - B) + (x_i^\top \beta)^2)] = (1 - B) \left( (x_i^\top \beta)^2 + \frac{(1 - B)(2 - B)}{8B} \right).
\]

Next we compute the second term in the right side of Eq. 5.11. By Lemmas 2 and Eq. 3.4,

\[
E(\hat{\lambda}_i^{EB} - \hat{\lambda}_i^{EB})^2 = \text{Var}(\hat{\lambda}_i^{EB}) + \text{Var}(\hat{\lambda}_i^{EB}) - 2\text{Cov}(\hat{\lambda}_i^{EB}, \hat{\lambda}_i^{EB}) + (E(\hat{\lambda}_i^{EB} - \hat{\lambda}_i^{EB}))^2 = \text{Var}(u_i^{21}) + \text{Var}(u_i^{22}) - 2\text{Cov}(u_i^{21}, u_i^{22}) + (E(u_i^{21} - u_i^{21}))^2. \tag{5.12}
\]

Also,

\[
\text{Var}(u_i^{21}) = 2(A(1 - B) + (2 - B)s_i/4)^2 + (4A(1 - B) + (2 - B)s_i)(x_i^\top \beta)^2 \tag{5.13}
\]
\[
\text{Var}(u_i^{22}) = 2(A(1 - B))^2 + 4A(1 - B)(x_i^\top \beta)^2 \tag{5.14}
\]
\[
\text{Cov}(u_i^{21}, u_i^{22}) = 2(1 - B)^2(A + s_i/4)^2 + 4(x_i^\top \beta)^2(1 - B)(A + s_i/4) \tag{5.15}
\]
\[
[E(u_i^{21} - u_i^{22})]^2 = (2 - B)^2 s_i^2/16 \tag{5.16}
\]
By Eqs. 5.13–5.16, 5.12 simplifies to

\[ E(\tilde{\lambda}_i^{EB} - \hat{\lambda}_i^B)^2 = B s_i(x_i^T \hat{\beta})^2 + (1 - B)^2 s_i/4 + s_i^2(1/2 - B/4 - B^2/16) \]  

(5.17)

Now we evaluate the last expression in right side of Eq. 5.11.

\[
\tilde{\lambda}_i^{EB} - \hat{\lambda}_i^B = ((1 - B)z_i + Bx_i^T \hat{\beta})^2 - ((1 - B)z_i + Bx_i^T \beta)^2 \\
= B^2[(x_i^T \hat{\beta} - x_i^T \beta)^2 + 2x_i^T \beta(x_i^T \hat{\beta} - x_i^T \beta)] \\
+ 2B(1 - B)(x_i^T \hat{\beta} - x_i^T \beta)x_i^T \hat{\beta} \\
+ 2B(1 - B)(x_i^T \hat{\beta} - x_i^T \beta)(z_i - x_i^T \hat{\beta}) \\
= B_1 + B_2 + B_3
\]

(5.18)

We define \( B_1 = B^2[(x_i^T \hat{\beta} - x_i^T \beta)^2 + 2x_i^T \beta(x_i^T \hat{\beta} - x_i^T \beta)] \), \( B_2 = 2B(1 - B)(x_i^T \hat{\beta} - x_i^T \beta)x_i^T \hat{\beta} \), \( B_3 = 2B(1 - B)(x_i^T \hat{\beta} - x_i^T \beta)(z_i - x_i^T \hat{\beta}) \). The expressions \( A_1, A_2 \) and \( A_3 \) in Eq. 3.5 are functions of residuals \( ((z_i - x_i^T \hat{\beta}), \ldots, (z_i - x_i^T \hat{\beta})) \) and the expressions \( B_1 \) and \( B_2 \) in Eq. 5.18 are functions of \( \hat{\beta} \). Therefore, \( (A_1, A_2, A_3) \) is independent of \( (B_1, B_2) \) and their covariances are 0. The expectation of \( B_3 \) is 0 since \( (x_i^T \hat{\beta} - x_i^T \beta) \) is independent of \( (z_i - x_i^T \hat{\beta}) \). Also,

\[
Cov(A_1, B_3) = \frac{2B(1 - B)}{4} E((B - \hat{B})(z_i - x_i^T \hat{\beta}))E(x_i^T \hat{\beta} - x_i^T \beta) = 0.
\]

\[
Cov(A_3, B_3) = 2B(1 - B)E \left[ (\hat{B}^2 - B^2 - 2(\hat{B} - B))(z_i - x_i^T \hat{\beta})^3 \right] \\
\times E(x_i^T \hat{\beta} - x_i^T \beta) \\
= 0.
\]

Hence, again using part (v) of Lemma 4, Eqs. 3.5 and 5.18, we have

\[
Cov(\tilde{\lambda}_i^{EB} - \hat{\lambda}_i^{EB}, \tilde{\lambda}_i^{EB} - \hat{\lambda}_i^B) = Cov(A_2, B_3) \\
= E(2x_i^T \hat{\beta}(B - \hat{B})(z_i - x_i^T \hat{\beta})2B(1 - B) \\
\times (x_i^T \hat{\beta} - x_i^T \beta)(z_i - x_i^T \hat{\beta})) \\
= 4B(1 - B)E((B - \hat{B})(z_i - x_i^T \hat{\beta})^2) \\
\times E(x_i^T \hat{\beta}(x_i^T \hat{\beta} - x_i^T \beta)) \\
= 4B(1 - B)\frac{(1 - s_i)}{2m_0} Var(x_i^T \hat{\beta}) \\
= 4B(1 - B)\frac{(1 - s_i)}{2m_0}(1/4 + A)s_i \\
= (1 - B)\frac{s_i(1 - s_i)}{2m_0}.
\]
Hence, from Eqs. 3.4 and 3.7,

\[
E((\hat{\lambda}_i^E - \tilde{\lambda}_i^E)(\tilde{\lambda}_i^E - \hat{\lambda}_i^E)) = (1 - B) \frac{s_i(1 - s_i)}{2m_0} \\
+ \frac{2 - B}{2} (1 - s_i) \frac{2 - B}{4} s_i \\
= \frac{s_i(1 - s_i)}{2m_0} (2 - 2B + B^2/4).
\]

Next we compute the remaining third term in the right side of Eq. 5.11.

\[
E(\hat{\lambda}_i^E - \tilde{\lambda}_i^E)^2 = E(A_1 + A_2 + A_3)^2 \\
= E(B - \hat{B})^2/16 + 4E(x_i^\top \hat{\beta}(B - \hat{B})(z_i - x_i^\top \hat{\beta}))^2 \\
+ E[(\hat{\beta}^2 - B^2)^2 - 2(\hat{B} - B)^2(z_i - x_i^\top \hat{\beta})]^2 \\
+ E(B - \hat{B})((\hat{B}^2 - B^2) - 2(\hat{B} - B))(z_i - x_i^\top \hat{\beta})^2]/2. \quad (5.19)
\]

In the above calculation, the cross terms \(E(A_1 A_2)\) and \(E(A_2 A_3)\) in Eq. 5.19 vanish by part (iii) of Lemma 3. Again, by part (ii) of Lemma 4, we have,

\[
E(B - \hat{B})^2 = B^2 - 2BE(\hat{B}) + E(\hat{B}^2) = -B^2 + \frac{m_0 - 2}{m_0 - 4} B^2 = \frac{2B^2}{(m_0 - 4)}. \quad (5.20)
\]

Again, applying Lemmas 3 and 4,

\[
E[x_i^\top \hat{\beta}(B - \hat{B})(z_i - x_i^\top \hat{\beta})]^2 = E[(x_i^\top \hat{\beta})^2]E[(B - \hat{B})^2(z_i - x_i^\top \hat{\beta})^2] \\
= \left[(x_i^\top \hat{\beta})^2 + \frac{s_i}{4\hat{B}} \right] \times E \left[ \frac{(z_i - x_i^\top \hat{\beta})^2}{\|Z - X\hat{\beta}\|^2} (B - \hat{B})^2 \frac{m_0 - 2}{4B} \right] \\
= \left[(x_i^\top \hat{\beta})^2 + \frac{s_i}{4\hat{B}} \right] \frac{(1 - s_i)(m_0 - 2)}{4m_0} \\
\times E[B^2/\hat{B} - 2B + \hat{B}^2] \\
= \left[(x_i^\top \hat{\beta})^2 + \frac{s_i}{4\hat{B}} \right] \frac{(1 - s_i)(m_0 - 2)}{4m_0} \frac{2B}{m_0 - 2} \\
= \frac{1 - s_i}{2m_0} \left[(x_i^\top \hat{\beta})^2 B + \frac{s_i}{4} \right].
\]
Now we calculate the third term, $1/2E[((\hat{B}^2 - B^2) - 2(\hat{B} - B))^2(z_i - x_i^T\hat{\beta})^4]$ in Eq. 5.19. By Lemma 3 and 4 and recalling $||Z - X\hat{\beta}||^2 = \frac{m_0-2}{4B}$, we obtain,

\[
E[((\hat{B}^2 - B^2) - 2(\hat{B} - B))^2(z_i - x_i^T\hat{\beta})^4] = E \left[ \frac{(z_i - x_i^T\hat{\beta})^4}{||Z - X\hat{\beta}||^4} \right] E \left[ ((\hat{B}^2 - B^2) - 2(\hat{B} - B))^2 \right] \left( \frac{m_0 - 2}{16B^2} \right) = \frac{3(1 - s_i)^2(m_0 - 2)^2}{16m_0(m_0 + 2)} \left[ B^2 \left( \frac{m_0}{m_0 - 2} - \frac{m_0(m_0 + 2)}{(m_0 - 2)^2} \right) + 4B \left( \frac{m_0}{m_0 - 2} - \frac{m_0(m_0 + 2)}{(m_0 - 2)^2} \right) + 4 \left( \frac{m_0}{m_0 - 2} - \frac{m_0(m_0 + 2)}{(m_0 - 2)^2} \right) \right] = \frac{3(1 - s_i)^2(m_0 - 2)^2}{16m_0(m_0 + 2)} \left[ \frac{2m_0^2 - 9m_0 + 6}{4m_0(m_0 + 2)(m_0 - 4)} B^2 - \frac{B}{m_0 + 2} + \frac{1}{2m_0} \right]. \tag{5.21}
\]

Finally, once again by Lemmas 3 and 4, and recalling $||Z - X\hat{\beta}||^2 = \frac{m_0-2}{4B}$, we get (5.19),

\[
E[(B - \hat{B})((\hat{B}^2 - B^2) - 2(\hat{B} - B)) (z_i - x_i^T\hat{\beta})^2] = E \left[ \frac{(z_i - x_i^T\hat{\beta})^2}{||Z - X\hat{\beta}||^2} \right] E \left[ (B - \hat{B})((\hat{B}^2 - B^2) - 2(\hat{B} - B))^2 \right] \left( \frac{m_0 - 2}{4B} \right) = \frac{(1 - s_i)(m_0 - 2)}{4m_0} \left[ B\hat{B} - \hat{B}^2 - B^3/\hat{B} + B^2 - 2(\hat{B} - B) - 2B + \frac{2B^2}{\hat{B}} \right] = \frac{(1 - s_i)(m_0 - 2)}{4m_0} \left[ B^2 - B^2 \frac{m_0 - 2}{m_0 - 4} - \frac{m_0 B^2}{m_0 - 2} + B^2 - 2B + \frac{m_0 2B}{m_0 - 2} \right] = \frac{(1 - s_i)(m_0 - 2)}{4m_0} \left[ \frac{4B}{m_0 - 2} - \frac{4B^2(m_0 - 3)}{(m_0 - 2)(m_0 - 4)} \right] = \frac{(1 - s_i)B}{m_0} \left[ 1 - \frac{B(m_0 - 3)}{m_0 - 4} \right]. \tag{5.22}
\]
From Eqs. 5.19–5.22, we obtain,

$$E(\hat{\lambda}_i^{EB} - \tilde{\lambda}_i^{EB})^2 = \frac{B^2}{8(m_0 - 4)} + \frac{1 - s_i}{2m_0} \left[ 4(x_i^\top \beta)^2 B + s_i \right]$$
$$+ 3(1 - s_i)^2 \left[ \frac{2m_0^2 - 9m_0 + 6}{4m_0(m_0 + 2)(m_0 - 4)} B^2 - \frac{B}{m_0 + 2} + \frac{1}{2m_0} \right]$$
$$+ \frac{(1 - s_i)B}{2m_0} \left[ 1 - B(m_0 - 3) \right].$$

Theorem 2 follows from Eqs. 5.11, 5.12, 5.17, 5.19 and 5.23.

References

Basu, D (1955). On statistics independent of a complete sufficient statistic. Sankhyā: The Indian Journal of Statistics (1933-1960) 15, 377–380.

Efron, B and Morris, C (1973). Stein’s estimation rule and its competitors—an empirical bayes approach. Journal of the American Statistical Association 68, 117–130.

Fay, R E and Herriot, R A (1979). Estimates of income for small places: an application of james-stein procedures to census data. Journal of the American Statistical Association 74, 269–277. https://doi.org/10.5705/ss.202020.0446

Ghosh, M, Natarajan, K, Waller, L A and Kim, D (1999). Hierarchical bayes glms for the analysis of spatial data: an application to disease mapping. Journal of Statistical Planning and Inference 75, 305–318.

Ghosh, M, Kubokawa, T and Kawakubo, Y (2015). Benchmarked empirical bayes methods in multiplicative area-level models with risk evaluation. Biometrika 102, 647–659.

Gonçalves, K. C. and Ghosh, M. (2021). Unit level model for small area estimation with count data under square root transformation. Brazilian Journal of Statistics and Probability In Press.

Hirose, M, Ghosh, M and Ghosh, T (2021). Arc-sin transformation for binomial sample proportions in small area estimation. Statistica Sinica Preprint No: SS-2020-0446. https://doi.org/10.5705/ss.202020.0446

Slud, E V and Maiti, T (2006). Mean-squared error estimation in transformed fay–herriot models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68, 2, 239–257.

Sugasawa, S and Kubokawa, T (2017). Transforming response values in small area prediction. Computational Statistics & Data Analysis 114, 47–60.

Yu, G (2009). Variance stabilizing transformations of poisson, binomial and negative binomial distributions. Statistics & Probability Letters 79, 14, 1621–1629.

Publisher’s Note. Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Paper received: 7 September 2021; accepted 14 September 2021.