Estimation of Poisson parameters: maximum likelihood, Bayes, empirical Bayes or a compromise?

Francisco J. Zamudio S.
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/8443

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Zamudio S., Francisco J.

ESTIMATION OF POISSON PARAMETERS: MAXIMUM LIKELIHOOD, BAYES, EMPIRICAL BAYES OR A COMPROMISE?

Iowa State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
Estimation of Poisson parameters:

Maximum likelihood, Bayes,
empirical Bayes or a compromise?

by

Francisco J. Zamudio S.

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1983
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DEDICATION</strong></td>
</tr>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
</tr>
<tr>
<td>A. Background</td>
</tr>
<tr>
<td>B. Outline</td>
</tr>
<tr>
<td><strong>II. COMPROMISE BETWEEN THE BAYES AND MAXIMUM LIKELIHOOD ESTIMATORS</strong></td>
</tr>
<tr>
<td>A. The Proposed &quot;Limited Translation Rule&quot;</td>
</tr>
<tr>
<td>B. Risk, Bayes Risk, and Relative Saving Loss of the Limited Translation Rule (LTR)</td>
</tr>
<tr>
<td>1. Risk of the LTR</td>
</tr>
<tr>
<td>2. Bayes risk of the LTR</td>
</tr>
<tr>
<td>3. Relative saving loss of the LTR</td>
</tr>
<tr>
<td>C. Generalizations</td>
</tr>
<tr>
<td>1. Sample size equal to n</td>
</tr>
<tr>
<td>2. Estimation of p Poisson means</td>
</tr>
<tr>
<td><strong>D. Conclusions</strong></td>
</tr>
<tr>
<td><strong>E. Figures and Tables</strong></td>
</tr>
<tr>
<td><strong>III. COMPROMISE BETWEEN EMPIRICAL BAYES AND MAXIMUM LIKELIHOOD ESTIMATORS</strong></td>
</tr>
<tr>
<td>A. The Limited Translation Compound Bayes Estimator</td>
</tr>
<tr>
<td>B. Risk, Bayes Risk, and Relative Saving Loss of the Limited Translation Compound Bayes Estimator</td>
</tr>
<tr>
<td>C. Conclusions</td>
</tr>
<tr>
<td>D. Figures and Tables</td>
</tr>
</tbody>
</table>
IV. COMBINED VERSUS SEPARATE EMPIRICAL BAYES ESTIMATORS 78
   A. Separate and Combined Empirical Bayes Estimators 78
   B. Bayes Rules 79
   C. Relative Saving Loss of a Group 81
   D. Total Bayes Risk When $p_1 = p_2$ 85
   E. Conclusions 86
   F. Figures and Tables 87

V. REFERENCES 99

VI. ACKNOWLEDGMENTS 101
DEDICATION

To The Memory of My
Mother
And My Uncle
José B. Zamudio A.
And
With Love and Gratitude
To My
Father.
I. INTRODUCTION

A. Background

For estimating a p-variate normal mean, the usual (maximum likelihood, minimum variance unbiased or best invariant) estimate is the sample mean. For p = 1, Blyth (1951) showed that the usual estimator was admissible under losses of the form $L(\theta, \delta) = W(||\delta - \delta_0||)$, where $|| \cdot ||$ denotes the usual Euclidean norm, W is nonnegative real-valued function satisfying $W(||u||) = W(-||u||)$ for all $u \in \mathbb{R}^p$, and $W(||u||)$ is nondecreasing in each $u_j$ ($j=1, \ldots, p$). The squared error loss is a special case of the above loss. For p = 2, the admissibility of the maximum likelihood estimator (MLE) was proved by Stein (1955) under the sum of squared error losses when the variance-covariance matrix was an identity matrix. These results were later strengthened by Stein (1959a, 1959b), Brown (1966) and Brown and Fox (1974). For $p \geq 3$, the usual estimator is no longer admissible for a fairly general kind of loss. The above fact was first discovered by Stein (1955) for the sum of squared error losses, i.e., when $L(\theta, \delta) = ||\delta - \delta_0||^2$. James and Stein (1961) showed that if $X \sim N(\theta_0, \Sigma)$, then the estimator of $\Theta_0$ given by $\hat{\Theta}_0 = (1-(p-2)S^{-1})X$, where $S = \sum_{i=1}^{p} X_i^2$ uniformly improves on the usual estimator $\hat{X}_0$ of $\Theta_0$.

The James-Stein estimator can be viewed as an empirical Bayes estimator, a notion introduced first by H. Robbins (1955). In this case, the Bayes estimators are obtained first under the assumption of a prior. Then, the unknown prior parameters are estimated from the
data, by using their marginal distribution, and are substituted in the Bayes estimators. In the normal example, assuming the prior distribution of \( \theta \) to be \( N_p(Q, A I_p) \) where \( A > 0 \) is a positive constant, the Bayes estimator of \( \theta \) under the sum of squared error losses is

\[
\hat{\theta}_{\nu B} = E[\theta | x] = (1 - (1 + A)^{-1}) x \tag{1.2}
\]

When \( A \) is unknown, it can be estimated from the marginal distribution of \( X \), which is

\[
X \sim N_p(Q, (1 + A) I_p)
\]

Thus, \( ||X||^2 / (1 + A) \) follows a chi-square distribution with \( p \) degrees of freedom, and hence for \( p \geq 3 \), \( E[(p - 2) / ||X||^2] = 1 / (1 + A) \). Then, using the estimator \( (p - 2) / ||x||^2 \) for \( 1 / (1 + A) \) in (1.2) one gets the James–Stein estimator.

Brown (1966) showed that the inadmissibility of the usual estimator for estimating the mean was still valid for a wider class of loss functions and for a general location family of distributions under certain regularity conditions.

While the MLE of \( \theta \) is inadmissible in three or higher dimensions, the Bayes estimator of \( \theta \) under squared error loss as given in (1.2) is admissible, irrespective of the dimensionality. Observe that \( \hat{\theta}_{\nu B} \) has the risk

\[
R(\theta, \hat{\theta}_{\nu B}) = p A^2 (1 + A)^{-2} + (1 + A)^{-2} ||\theta||^2.
\]

Thus, while the risk performance of the Bayes estimator \( \hat{\theta}_{\nu B} \) of \( \theta \) is much better than that of the MLE (which is minimax with constant risk \( p \)) around the prior mean, namely \( Q \), \( \hat{\theta}_{\nu B} \) performs much worse than the MLE for large \( ||\theta|| \), i.e., as \( \theta \) moves further and further away from the
prior mean. The same phenomenon continues to hold when the prior mean is different from zero. Thus, the Bayes estimator lacks minimaxity. Also, if the true prior distribution $\xi^*$ is $N(\mu, A\eta^p)$, then the Bayes risk of $\hat{\theta}_B$ is given by

$$r(\xi^*, \hat{\theta}_B) = (\frac{A}{A+1} - \frac{A^*}{A^*+1})^2 (A^*+1)a_p^* A^* + \frac{1}{(A+1)^2} ||\xi^*||^2$$

Observe that for large $||\mu^*||$, $\hat{\theta}_B$ has a large Bayes risk, and then misspecified priors, whereas the MLE is robust against any arbitrarily chosen prior.

With the above considerations in mind, Efron and Morris (1971) proposed a compromise between the MLE and the Bayes estimators. These compromise estimators, referred to as the "Limited Translation Rule" estimators, are given by $\{\delta_{A_1M_1}(X_1), \ldots, \delta_{A_pM_p}(X_p)\}$, where

$$\delta_{A_iM_i}(x_i) = \begin{cases} x_i + M_i & \text{if } x_i < -c_i \\ \frac{A}{A+1}x_i & \text{if } x_i \in [-c_i, c_i] \\ x_i - M_i & \text{if } x_i > c_i \end{cases}$$

with $c_i = (1+A)M_i$ (1.3)

The estimator proposed in (1.3) fixes the maximum allowable deviation say $M_i$ from $x_i$ and uses the Bayes estimator $\frac{A}{A+1}x_i$ subject to the constraint $|\frac{A}{A+1}x_i - x_i| \leq M_i$. The above could be made more general by allowing different prior variances $A_i$'s.

Next, note that although the James–Stein estimator given in (1.1) dominates $\hat{\theta}_B$ under sum of squared error losses, for a specific component, its risk performance could be much worse than that of the corresponding component of $X$. This is demonstrated by an example in Efron and Morris (1972). Accordingly, they proposed a compromise between
the James-Stein estimator and the MLE which has good componentwise risk performance without sacrificing good ensemble properties. The estimator proposed is similar to (1.3).

Finally, consider the situation where the $p$ parameters can be divided into two natural groups (e.g., right handed and left handed baseball players as in the case of Efron and Morris (1973a)) of sizes $p_1$ and $p_2$ with $p_1 + p_2 = p$. In this case, the statistician wonders whether the James-Stein estimator should be applied separately to the two sets or once to the combined problem. Specifically, let $\theta = (\theta_1, \theta_2)$ where $\theta_i = (\theta_{i1}, \theta_{i2}, \ldots, \theta_{ip_i})$, $(i=1,2)$. For each $\theta_{ij}$, there is an observation from the random variable $X_{ij}$ which given $\theta_{ij}$ has a normal distribution with parameters $\theta_{ij}$ and $D_i$, $i=1,2$ and $j=1,\ldots,p_i$. For simplicity, take $D_1 = D_2 = D$. There will be squared error loss functions for $\theta_1$ and $\theta_2$ separately and also for $\theta = (\theta_1, \theta_2)$. If $a = (a_1, a_2)$ is an estimate of $\theta$, then these loss functions are

$$L_i(\theta_{\hat{\nu}_1}, \hat{\nu}_1) = \frac{1}{p_i} \| \hat{\nu}_1 - \nu_1 \|^2, \quad i=1,2$$
$$L(\hat{\theta}_{\hat{\nu}}, \nu) = \frac{1}{p} \| a - \theta \|^2$$

The estimators will be of the form

$$\delta_{\hat{\nu}_1}^{(\text{sep})}(X_{\hat{\nu}}) = [1 - \frac{D(p_i - 2)}{S_1}] X_{\hat{\nu}_1}$$
$$\delta_{\hat{\nu}_1}^{(\text{comb})}(X_{\hat{\nu}}) = [1 - \frac{D(p-2)}{S}] X_{\hat{\nu}_1}$$

where $S_1 = \| X_{\hat{\nu}_1} \|^2$, and $S = S_1 + S_2$. Efron and Morris (1973a) consider a bigger class of estimator, i.e.,

$$\delta_{\hat{\nu}_1}(X) = [1 - \frac{D(p-2)}{S}] \rho_1(V_{\hat{\nu}_1}) X_{\hat{\nu}_1}$$

(1.4)
where $V_i = S_i / S$. The rule $\delta_i^{(\text{comb})}$ is obtained from (1.4) putting $p_i(V_i) = 1$, while the choice

$$p_i(V_i) = (p_i - 2)/(p - 2) V_i$$

defines $\delta_i^{(\text{sep})}$.

Efron and Morris (1973a) studied among other things the circumstances under which the "separate" estimators are better than the "combined" estimators or vice versa. Further, they showed that there is a class of compromise estimators, Bayesian in nature, which will usually be preferred to either alternative.

Next, consider the Poisson means estimation problem. In this case, we consider $p$ independent Poisson variables $X_1, \ldots, X_p$ with means $\theta_1, \ldots, \theta_p$ where $\theta_i \in (0, \infty)$ is unknown for each $i = 1, \ldots, p$. Taking one observation from each population involves no loss of generality because if $(i = 1, \ldots, p; j = 1, \ldots, n_i)$ are independent with $(j = 1, \ldots, n_i)$ identically distributed Poisson $(\theta_i), i = 1, \ldots, p$, then the minimal sufficient statistic for $\theta = (\theta_1, \ldots, \theta_p)$ is $X = (X_1, \ldots, X_p)$ where $X_i = \sum_{j=1}^{n_i} X_{ij}$ $\sim$ Poisson $(n_i \theta_i)$, $i = 1, \ldots, p$. If we estimate $\theta$ by $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_p)$, then consider losses of the form

$$L(\theta, \hat{\theta}) = \sum_{i=1}^{p} m_i (\theta_i - \hat{\theta}_i)^2$$

(1.5)

where the $m_i$'s are known nonnegative integers. The usual (maximum likelihood, minimum variance unbiased or best invariant) estimator of $\theta$ is $X$. However, although $X$ is an admissible estimator of $\theta$ under any loss of the type (1.1), when $p = 1$, it is not so for higher values of $p$. The critical value of $p$ for inadmissibility depends of course
on the $m_i$'s.

Clevenson and Zidek (1975) observed the above fact when $m = \ldots = m = 1$ in (1.5). In this case, $X$ turned out to be an inadmissible estimator of $\theta$ for $p \geq 2$. The case $m = \ldots = m = 0$ was studied by Peng (1975) [see also Hudson (1978)]. Peng proved the admissibility of $X$ for $p = 2$ and its inadmissibility for $p \geq 3$.

Tsui and Press (1982) studied the case $m = \ldots = m = m$ when $m$ is a nonnegative integer. They proved the inadmissibility of $X$ for $p \geq 2$ when $m \geq 1$ and for $p \geq 3$ when $m = 0$. In all these cases, they gave a general class of estimators dominating $X$. Finally, Hwang (1982) considered (1.5) in its most general form and gave a general class of estimators dominating $X$ when the $m_i$'s were not necessarily equal.

B. Outline

In this thesis, we will be interested in three point estimation problems, where the random variables will be assumed to be Poisson. Also, we assume conjugate prior distributions for the parameters. The problems are:

1) Estimators compromising between the maximum likelihood and the Bayes estimators. This problem is worked out in part II.

   ii) Estimators compromising between the maximum likelihood and empirical Bayes estimators. These estimators are treated in part III.

   iii) Study of combined against separate estimators. Part IV of the thesis addresses this problem.

The above problems are described more fully below. In what
follows, we will confine ourselves to the case \( m_1 = \ldots = m_p = 1 \) in the loss function given in (1.5). We also consider a componentwise loss function of the form \( L(\theta, \delta_i) = (\delta_i - \theta_i)^2 / \theta_i \).

Let us add in the first problem the assumption that \( \theta_i \)'s are independent random variables with Gamma \((a, k_i)\) prior distributions for \( i = 1, \ldots, p \), i.e., the prior probability density function of \( \theta_i \) is given by

\[
\theta_i > 0, \quad a > 0 \quad \text{and} \quad k_i > 0.
\]

Then, the posterior distribution of \( \theta_i \) given \( X_i = x_i \) \((i = 1, \ldots, p)\) is Gamma \((1 + a, x_i + k_i)\) and consequently the Bayes estimate of \( \theta \) under the prior (1.6) is given by

\[
\hat{\theta}_{n, k_i} = \frac{k_i}{n} \exp\left(-a\theta_i \right) \theta_i^{-k_i}
\]

where \( \theta_i > 0, a > 0 \) and \( k_i > 0 \). Then, the posterior distribution of \( \theta_i \) given \( X_i = x_i \) \((i = 1, \ldots, p)\) is Gamma \((1 + a, x_i + k_i)\) and consequently the Bayes estimate of \( \theta \) under the prior (1.6) is given by

\[
\hat{\theta}_{n, k_i} = \frac{(x_i + k_i - 1)/(1 + a)}{\theta_i}, \quad \text{if } k_i \geq 1;
\]

\[
0, \quad \text{if } k_i < 1 \text{ and } x_i = 0
\]

(1.7)

Note that the componentwise risk for \( \hat{\theta}_{n, k_i} \) is

\[
R(\theta, \hat{\theta}_{n, k_i}) = E_{\theta} \left[ (\hat{\theta}_{n, k_i} - \theta_i)^2 / \theta_i \right]
\]

\[
= \left(2b_i - b_i^2 / \theta_i^2 \right) e^{-\theta_i} \mathbb{I}(k_i < 1) + (1 + a)^{-2} \left[ (1 + a)^{-1} \left( \theta_i^2 - (k_i - 1)a^{-1} \right)^2 \right]
\]

(1.8)

where \( b_i = (k_i - 1)/(1 + a) \).

Observe that for \( k_i > 1 \) \((k_i \leq 1)\) the prior mode is \((k_i - 1)a^{-1}\) (zero). Thus, componentwise the Bayes estimator performs quite satisfactorily for \( \theta_i \) close to the prior mode. This is easy to see for \( k_i \geq 1 \) from
(1.8). For $k_1 < 1$, from (1.8) we get,

$$R(\theta, \delta_{\alpha, k_1}) = 2b_1 e^{-\theta_1} + \theta_1^{-1} b_1^2 (1-e^{-\theta_1}) + (1+\alpha)^{-2} (1+\alpha_1^2 - 2(k_1-1)\alpha)$$

and so, taking the limit of the above risk when $\theta_1 \to 0$ one gets

$$\lim_{\theta_1 \to 0} R(\theta, \delta_{\alpha, k_1}) = 2b_1 + \alpha b_1^2 (1-a)^{-2 - 2b_1 a}$$

with $a = \alpha/(1+a)$. However, (1.7) performs very poorly by the componentwise risk criterion for values of $\theta_1$ far from the prior mode. In fact, for $k_1 > 1$, $R(\theta, \delta_{\alpha, k_1}) \to \infty$ as $\theta_1 \to 0$ or $\theta_1 \to \infty$ and for $k_1 \leq 1$, $R(\theta, \delta_{\alpha, k_1}) \to \infty$ when $\theta_1 \to \infty$. On the other hand, the usual componentwise estimator of $\theta_1$, namely $X_1$, is a minimax estimator of $\theta_1$ with constant risk 1.

The above unpleasant characteristic of the estimator $\delta_{\alpha, k_1}$ can also be described in a Bayesian framework. Suppose, for instance, $\theta_1$ has a Gamma $(\alpha_1, k_{11})$ prior. For simplicity, let us assume in this specific context that $k_1$ and $k_{11}$ are bigger or equal to one. Then, the componentwise Bayes risk of the estimator $\delta_{\alpha, k_1}$ with respect to the Gamma $(\alpha_1, k_{11})$ prior (denoted by $g_{\alpha_1, k_{11}}$) is given by

$$r(g_{\alpha_1, k_{11}}, \delta_{\alpha, k_1}) = (1+a)^{-2} [1+a_1^2 + a_1 (k_{11}-1)^{-1} - (k_{11}-1)\alpha_1^{-1} - (k_1-1)\alpha^{-1}]^2$$

which exceeds 1 (the Bayes risk of the usual estimator $X_1$) if and only if $[(k_{11}-1)\alpha_1^{-1} - (k_1-1)\alpha^{-1}]^2 > (k_{11}-1)\alpha_1^{-1} (1+2\alpha_1^{-1} - \alpha_1^{-1})$.

Expression (1.9) shows the lack of "robustness" of the usual Bayes procedure as compared to the usual one with respect to the choice of
priors. As noted earlier, a similar lack of robustness property of Bayes estimators in the normal case was shown by Efron and Morris (1971).

Thus, in the Poisson case, minimaxity and robustness of the maximum likelihood estimator with respect to varied choice of priors together with the good performance of the Bayes estimators around the mode of the chosen prior, makes it relevant to suggest procedures compromising between the maximum likelihood and the Bayes procedures.

The objective behind the use of such compromise estimators is that these estimators are robust against misspecified priors, whereas they outperform the MLE if the prior mode is correctly specified, and the variability of the prior distribution is not very high.

One compromise between the Bayes estimator and the maximum likelihood estimator is the use of "Limited Translation Rules" as proposed by Efron and Morris (1971) in the normal case. For any \(a(>0)\) and \(M(>0)\), let \(c=c(a,M)=M(a+1)/\alpha=Ma^{-1}\). Define the estimator \(\delta_{M,\alpha,k_i}\) of \(\theta_i\) as

\[
\delta_{M,\alpha,k_i}(x_i) = \begin{cases} 
  x_i - M & \text{if } x_i < \max(0, -\frac{k_i-1}{\alpha} - c) \\
  \frac{(k_i-1+x_i)/(1+\alpha)}{\alpha} & \text{if } |x_i-(k_i-1)\alpha^{-1}| \leq c \\
  x_i - M & \text{if } x_i > (k_i-1)\alpha^{-1} + c
\end{cases}
\]  

(1.10)

The estimator in (1.10) fixes the maximum allowable deviation, say \(M\) from \(x_i\) and use the Bayes estimator \(\delta_{\alpha,k_i}\) subject to the constraint \(|(k_i-1+x_i)/(1+\alpha)-x_i| \leq M\).

The proposed compromising estimator \(\delta_{M,\alpha,k_i}\) is a shrinkage estimator, it will shrink more towards the usual estimate \(x_i\) for both large
and small values of $x_i$ and indeed the data determine the shrinking factor. We will compare in this work the performance of $\delta_{M,a,k_i}$ with the maximum likelihood estimate $x_i$ in terms of their risk performances. Also, we will compare $\delta_{M,a,k_i}$ with the Bayes estimate $\delta_{a,k_i}$ in terms of the Bayes risk sacrifice one incurs by the use of $\delta_{M,a,k_i}$ instead of $\delta_{a,k_i}$ if $g_{a,k_i}$ were the true prior. A way to compare Bayes risks is to use the criterion of relative saving loss (RSL) as introduced by Efron and Morris (1971, 1972, 1973b) [see also Berger (1982)]. This is defined by

$$RSL(g_{a,k_i}, \delta_{M,a,k_i}) = \frac{r(g_{a,k_i}, \delta_{M,a,k_i}) - r(g_{a,k_i}, \delta_{a,k_i})}{r(g_{a,k_i}, \delta_{M,a,k_i}) - r(g_{a,k_i}, \delta_{a,k_i})}$$

where $\delta^0_1(x) = \delta^0(x_i) = x_i$. This is the proportion of the possible Bayes risk improvement over $\delta^0_1$ that is sacrificed by the use of $\delta_{M,a,k_i}$ instead of the use of the Bayes rule with respect to the prior $g_{a,k_i}$.

Next, for the second problem, consider for $\theta_i$ a Gamma ($\mu=\mu(0,1)$) prior distribution where $\mu$ is an unknown parameter. Assuming $\mu$ to be known, the Bayes estimate of $\theta_i$ is $\delta_{u,k_i}(x_i) = (1-u)(x_i+k_i-1)$ when $x_i+k_i-1 > 0$ and zero otherwise. If $\mu$ is unknown, it can be estimated from the data through the marginal distribution of $X_i$'s. Ghosh (1983) has shown that the $X_i$'s are marginally independent Negative Binomial random variables, where the marginal distribution of $X_i$ involves the parameters $\mu$ and $k_i$. The minimal sufficient statistic for $\mu$ is $T = \sum_{i=1}^{p} X_i$ which is Negative Binomial with parameters $\mu$ and $k = \sum_{i=1}^{p} k_i$. Therefore, an empirical Bayes estimate of $\theta_i$ is of the form $(1-\hat{\mu}(t))(x_i+k_i-1)I(x_i+k_i-1 > 0), \ldots, (1-\hat{\mu}(t))(x_i+k_i-1)$.
For the specific case $k_1 = k_2 = \ldots = k_p = 1$, the empirical Bayes estimate is given by

\[ ((1-u(t))x_1, \ldots, (1-u(t))x_p) \]

As mentioned earlier, the maximum likelihood estimator of $\theta$, namely $\hat{\theta}$, ceases to be an admissible estimator for $p \geq 2$, it being dominated by a class of estimators of the form $\delta_\phi(\hat{\theta}) = (\delta_\phi^1(\hat{\theta}), \ldots, \delta_\phi^p(\hat{\theta}))$ with $\delta_\phi^i(\hat{\theta}) = (1-\frac{\phi(T)}{T+a})x_i, T = \sum_{j=1}^{p} x_j, \quad (i=1, \ldots, p)$, where $\phi$ satisfies certain regularity assumptions. This was shown first by Clevenson and Zidek (1975), and later the class $\phi$ of estimators was widened by Tsui (1979a), Hwang (1982) and Tsui and Press (1982).

A class of estimators dominating $\hat{\theta}$ for $p \geq 3$ is given by

\[
\delta_{m,n}(\hat{\theta}) = \left(1-\frac{(m+p)}{(t+m+n+p-1)}\right) \hat{\theta}^p, \quad -1 \leq m \leq p-2, \: n > 0, \quad \text{where } t = \sum_{j=1}^{p} x_j
\]

The estimator in (1.11) can be shown to be generalized Bayes with respect to the (possibly improper) prior

\[
\Pi_{m,n}(\theta) = \left(\prod_{j=1}^{p} \theta_j\right)^{-\frac{m+p-1}{2}} \exp\left(-\frac{m+n-2}{2}\right) \exp(-u) \, du
\]

Ghosh and Parsian (1981) showed that the estimator $\delta_{m,n}(X)$ is a unique proper Bayes estimator against the prior given in (1.12) for $m > 0, n > 0$, and hence is admissible. Also, the admissibility of such estimators is proved in Clevenson and Zidek (1975) for $m > 0$ and $n = 1$, in Ghosh (1983) for $m > -1$ and $n > 0$, and in Brown and Hwang (1982) for $m > -1$ and $n = 1$. Also, Ghosh (1983) has shown how some of the estimates in (1.11) can be given empirical Bayes interpretation.

A class of admissible estimators dominating $\hat{\theta}$ for $p \geq 2$ was given by Ghosh (1983). They are of the form.
Brown and Hwang (1982) showed that admissibility holds when \( n \leq 1 \).

The empirical Bayes estimate for the case \( k_1 = k_2 = \ldots = k_p = 1 \) when \( 1-u \) is estimated by \( \left( 1 - \frac{p-1}{t+p-1} \right) \) or by \( \left( 1 - \frac{p}{t+p} \right) \) (the minimum variance unbiased estimate and the maximum likelihood estimate of \( 1-u \) respectively) are estimates of the form given in (1.13) with \( \beta = 1 \) and \( \eta = 1 \) in the first case and \( \beta = 1 \) and \( \eta = 0 \) in the second case.

Although the estimators given in (1.11) and (1.13) guarantee a reduction in total risk as compared to the usual estimator \( \hat{X}_n \) of \( \theta \), sometimes they do not perform very well in componentwise estimations. Thus, when estimating a particular component, say \( \theta_j \), by the \( j \)th component of \( \hat{\theta}_m, \) or \( \hat{\theta}^* \), say \( \hat{\theta}_{m,n} \) or \( \hat{\theta}^*_{m,n} \), the maximum risk involved may be quite high in comparison with the usual estimator \( X_i \) which has constant risk 1. In the spirit of the limited translation estimates proposed in (1.10), we will compromise between the maximum likelihood estimates and empirical Bayes estimates.

For estimating \( \theta_i \) (\( i = 1, \ldots, p \)), we will use

\[
\delta^*_m(x) = \left( 1 - \frac{p-\eta}{p-\eta+\beta+t-1} \right) x_i, \beta > 0 \text{ and } \eta < 1. \tag{1.13}
\]

with \( t = \sum_{i=1}^{p} x_i \). Brown and Hwang (1982) showed that admissibility holds when \( \eta \leq 1 \).

\[
\delta^*_m(x) = \left( 1 - \frac{p-\eta}{p-\eta+\beta+t-1} \right) x_i, \beta > 0 \text{ and } \eta < 1. \tag{1.13}
\]

The empirical Bayes estimate for the case \( k_1 = k_2 = \ldots = k_p = 1 \) when \( 1-u \) is estimated by \( \left( 1 - \frac{p-1}{t+p-1} \right) \) or by \( \left( 1 - \frac{p}{t+p} \right) \) (the minimum variance unbiased estimate and the maximum likelihood estimate of \( 1-u \) respectively) are estimates of the form given in (1.13) with \( \beta = 1 \) and \( \eta = 1 \) in the first case and \( \beta = 1 \) and \( \eta = 0 \) in the second case.

Although the estimators given in (1.11) and (1.13) guarantee a reduction in total risk as compared to the usual estimator \( \hat{X}_n \) of \( \theta \), sometimes they do not perform very well in componentwise estimations. Thus, when estimating a particular component, say \( \theta_j \), by the \( j \)th component of \( \hat{\theta}_m, \) or \( \hat{\theta}^* \), say \( \hat{\theta}_{m,n} \) or \( \hat{\theta}^*_{m,n} \), the maximum risk involved may be quite high in comparison with the usual estimator \( X_i \) which has constant risk 1. In the spirit of the limited translation estimates proposed in (1.10), we will compromise between the maximum likelihood estimates and empirical Bayes estimates.

For estimating \( \theta_i \) (\( i = 1, \ldots, p \)), we will use

\[
\delta^*_m(x) = \left( 1 - \frac{p-\eta}{p-\eta+\beta+t-1} \right) x_i, \beta > 0 \text{ and } \eta < 1. \tag{1.13}
\]

where \( M (>0) \) is some constant, \( t = \sum_{j=1}^{p} x_j \) as before, \( c_1(t) \) is the smallest integer greater than or equal to \( \left( 1 - \hat{u}(t) \right) (k_1 - 1) \) and \( \hat{u}(t) \) is the largest integer less than or equal to \( \left( M + (1 - \hat{u}(t)) (k_1 - 1) \right) / \hat{u}(t) \).

Again, the motivation behind the use of the estimates in (1.14) is
that we fix the maximum allowable deviation, say M from $x_i$ and use the empirical Bayes estimate $(1-\hat{u}(t))(x_i+k_i-1)$ subject to the constraint $|(1-\hat{u}(t))(x_i+k_i-1)-x_i| \leq M$. A similar compromise estimate in the normal case was proposed by Efron and Morris (1972).

We believe that the estimators $\delta_{\lambda_M}(\lambda) = (\delta_{\lambda_M,1}, \ldots, \delta_{\lambda_M,p})$ with $\delta_{\lambda_M,i}$ $(i=1, \ldots, p)$ given in (1.14) have both good ensemble as well as good component properties. We will investigate the component property using the risk criterion and the ensemble property in terms of the Bayes risk sacrifice one incurs by using the compromise estimate $\delta_{\lambda_M}$ instead of the empirical Bayes estimate which will be denoted by $\delta^1_{\lambda}$. Recall that the prior distribution to the $\theta_i$'s is Gamma $(\frac{u}{1-u}, k_i)$. Writing $\delta_{\lambda}^0(X) = X$ as before, the Bayes risk comparison can be made by using the relative savings loss (RSL) criterion

$$RSL((u,k), \delta_{\lambda_M}) = \frac{r((u,k), \delta_{\lambda_M}) - r((u,k), \delta^1_{\lambda})}{r((u,k), \delta^0_{\lambda}) - r((u,k), \delta^1_{\lambda})}$$

(1.15)

where $k=(k_1, \ldots, k_p)$ and $r((u,k), \delta)$ denotes the Bayes risk of $\delta$ with respect to the joint prior

$$g_{u,k}(\theta) = \prod_{i=1}^{p} \{\exp(-u\theta_i) \theta_i^{k_i-1} u^{k_i}/\Gamma(k_i)\}$$

(1.16)

Once again, (1.15) is the proportion of the possible ensemble Bayes risk improvement over $\delta_{\lambda}^0$ that is sacrificed by the use of $\delta_{\lambda_M}$ instead of $\delta^1_{\lambda}$ if (1.16) were the true prior.

In particular, we will give numerical results for the case when $k_1=k_2=\ldots=k_p$ and two possible estimators of $u$, i.e., $\hat{u}(t) = (p-1)/(t+p-1)$ and $\hat{u}(t) = p/(t+p)$. 
Finally, in the third problem, we consider the situation when the \( p \) parameters can be divided into two natural groups of sizes \( p_1 \) and \( p_2 \) with \( p_1 + p_2 = p \). We want to investigate whether it is better to apply empirical Bayes estimators separately to the two groups or apply it all at once to the combined problem. Specifically, we will deal with empirical Bayes estimators where \((1-u)\) is estimated through its minimum variance unbiased estimator. In other words, let \( \hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2) \) where \( \hat{\theta}_i = (\hat{\theta}_{i1}, \hat{\theta}_{i2}, \ldots, \hat{\theta}_{ip_i}) \), \( i = 1, 2 \). Corresponding to each \( \hat{\theta}_{ij} \), there is a single \( X_{ij} \) such that \( X_{ij} \sim \text{Poisson} (\theta_{ij}) \), \( j = 1, \ldots, p_i \), \( i = 1, 2 \). We assume independence of all the \( X_{ij} \)'s. If we estimate \( \theta_{11} \) and \( \theta_{22} \) respectively by \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \), then the losses considered are of the form

\[
L_i(\hat{\theta}_1, \hat{\theta}_2) = p_i \sum_{j=1}^{p_i} (\hat{\theta}_{ij} - a_{ij})^2 / \theta_{ij}, \quad i = 1, 2, \quad (1.17)
\]

while for estimating \( \theta \) by \( \hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2) \), consider the loss function

\[
L(\hat{\theta}, \theta) = p \sum_{i=1}^{2} \frac{p_i}{2} \sum_{j=1}^{p_i} (\hat{\theta}_{ij} - a_{ij})^2 / \theta_{ij} \quad (1.18)
\]

Let \( x = (x_1, x_2) = (x_{11}, \ldots, x_{1p_1}, x_{21}, \ldots, x_{2p_2}) \), one can either estimate \( \theta_{11} \) by

\[
\hat{\delta}_{11}^{(sep)}(x, \gamma) = (1 - \hat{u}(t_1)) (x_{11} - l_1), \quad t_1 = \sum_{j=1}^{p_1} \frac{x_{1j}}{\gamma_{1j}}
\]

\( i = 1, 2 \), or estimate \( \theta_{11} \) by

\[
\hat{\delta}_{11}^{(comb)}(x) = (1 - \hat{u}(t_1, t_2)) (x_{11} + k_{11} - l_1)
\]

where \( k_{11} = (k_{11}, \ldots, k_{1p_1}) \) and \( l_1 \) is a \( p_1 \)-dimensional vector of ones.

Of course, one can estimate \( \theta \) either by \( \hat{\delta}_1 = (\hat{\delta}_{11}^{(sep)}, \hat{\delta}_{12}^{(sep)}) \) or by \( \hat{\delta}_2 = (\hat{\delta}_{11}^{(comb)}, \hat{\delta}_{12}^{(comb)}) \). Bayes risk comparisons can be made
by using separate priors for $\theta_{1j}$ and $\theta_{2j}$, say $\theta_{i,j}'s$ are independent 
with $\theta_{1j} \sim \text{Gamma}(u_1(1-u_1)^{-1},k_{1j}),(j=1,\ldots,p_1)$ and $\theta_{2j} \sim \text{Gamma}$ 
$(u_2(1-u_2)^{-1},k_{2j}),(j=1,\ldots,p_2)$. The Bayes risks comparisons will 
be made through the "partial" and "total" relative saving loss 
criterion as introduced by Efron and Morris (1973a). The "partial"
relative saving loss of $\delta_i$ for group $i$ is defined to be

$$RSL_i((u_i,k_{i1}),\delta_{i4}) = \frac{r_i((u_i,k_{i1}),\delta_{i4})-r_i((u_{i1},\delta_{i4}^o),\delta_{i4}^*)}{r_i((u_i,k_{i1}),\delta_{i4}^o)-r_i((u_{i1},\delta_{i4}^o),\delta_{i4}^*)}$$

$i=1,2$. Where $k_{i1}=(k_{i11},k_{i12},\ldots,k_{i1p_1}),\delta_{i4}^*$ is the Bayes estimator of $\theta_{ij}$
with respect to the Gamma prior introduced earlier for $\theta_{i1}\delta_{i4}^o$ is the 
maximum likelihood estimator for $\theta_{i1}$, i.e., $\chi_{i1}$ and finally $r_i((u_i,k_{i1}),\delta_{i4})$
is the Bayes risk of $\delta_{i4}$ under the loss (1.17) and the Gamma prior for
$\theta_{i1}$ given earlier.

The "total" relative saving loss of $\delta=(\delta_1,\delta_2)$ is defined to be

$$RSL((u,k),\delta) = \frac{r((u,k),\delta)-r((u,k),\delta^o)}{r((u,k),\delta^o)-r((u,k),\delta^*)}$$

where $u=(u_1,u_2), k=(k_{11},k_{22}),\delta^*=(\delta_{11}^*,\delta_{22}^*),\delta^o=(\delta_{11}^o,\delta_{22}^o)$ and $r((u,k),\delta)$
is the Bayes risk of $\delta$ under the loss in (1.18) with respect to the
joint Gamma prior for $(\theta_{11},\theta_{22})$.

For the case $k_{11}=1$, and using the minimum variance estimate for
$1-u$, the separate and combined estimates turn to be

$$\delta_{i1}^{(sep)}(x_1)=[1-(p_1-1)/(t_1+t_1p_1)]x_1 \quad (1.19)$$

$$\delta_{i1}^{(comb)}(x)=[1-(p-1)/(t+t+p-1)]x_{\delta_1} \quad t=t_1+t_2 \quad (1.20)$$

since $T_1$ is the minimal sufficient estimator of $u_1$, one would expect
that if $u_1=u_2$ then (1.20) would do better than (1.19) while if $u_1 \neq u_2$.
the opposite would occur. We will study the circumstances under which the "separate" estimators are better than the "combined" estimators or vice versa.

The study of the combined versus separate James-Stein estimators in the normal case was undertaken in Efron and Morris (1973a), and later on in Berger and Dey (1980), and in Dey (1981).

It is clear that generalizations should be possible when the number of groups in which the parameter vector $\theta$ is partitioned, is an arbitrary number, not necessarily equal to 2.
II. COMPROMISE BETWEEN THE BAYES AND MAXIMUM LIKELIHOOD ESTIMATORS

A. The Proposed "Limited Translation Rule"

We have motivated in the introduction why it is useful to consider estimators compromising between the Maximum Likelihood and Bayes estimators. One such compromise is to fix the maximum allowable deviation from the MLE, say $M$, and use the Bayes estimator only if its deviation from the MLE does not exceed $M$. Otherwise, use an estimator closer to the MLE than the Bayes estimator. This will be made more specific in the later paragraphs.

To start with, let $X$ be a Poisson random variable with mean $\theta$. Consider a Gamma $(\alpha, k)$, with $\alpha > 0$ and $k > 0$, prior for $\theta$. Also, if we estimate $\theta$ by $a$, assume a loss function of the type

$$L(\theta, a) = (\theta - a)^2/\theta$$

We will denote the MLE by $\delta^0$ and the Bayes estimator by $\delta_{\alpha, k}$, i.e.,

$$\delta^0(x) = x \text{ and } \delta_{\alpha, k}(x) = \begin{cases} \frac{x + k - 1}{1 + \alpha} & \text{if } x + k - 1 > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

As mentioned earlier, for a fixed $M > 0$, we want to use the Bayes estimator when the restriction $|\delta^0(x) - \delta_{\alpha, k}(x)| < M$ holds, i.e., when

$$\frac{((k-1)-M(1+\alpha))}{\alpha} \leq x \leq \frac{M(1+\alpha)+(k-1)}{\alpha} \text{ or}$$

$$(b-M)/a \leq x \leq (M+b)/a, \text{ if we define } a = \alpha/(1+\alpha) \text{ and }$$

$$b = (k-1)/(1+\alpha).$$

For estimating $\theta$, we propose the "Limited Translation Rule"
\[ x + M \quad \text{if } x < \frac{(b-M)}{a} \]

\[ \delta_{M,\alpha,k}(x) = \begin{cases} 
  x(1-a) + b & \text{if } \frac{(b-M)}{a} \leq x < \frac{(M+b)}{a} \\
  x-M & \text{if } x > \frac{(M+b)}{a}
\end{cases} \quad (2.2) \]

For the case \( k=1 \), since \( b=0 \), the rule turns to be

\[ \delta_{M,\alpha,1}(x) = \begin{cases} 
  x(1-a) & \text{if } x \leq \frac{M}{a} \\
  x-M & \text{if } x > \frac{M}{a}
\end{cases} \quad (2.3) \]

while for \( k < 1 \), the rule is

\[ \delta_{M,\alpha,k}(x) = \begin{cases} 
  x(1-a) + b & \text{if } 1 < x < \frac{(M+b)}{a} \\
  x-M & \text{if } x > \frac{(M+b)}{a}
\end{cases} \quad (2.4) \]

Finally, for the case \( k > 1 \), we proceed as follows. Note that since the Bayes estimate is not zero at \( x=0 \), the corresponding estimator has infinite risk as \( \theta \to 0 \). Hence, first use the following modified Bayes estimate

\[ \tilde{\delta}_{\alpha,k}(x) = \begin{cases} 
  0 & \text{if } x=0 \\
  \frac{x(1-a)}{1+(1+\alpha)} & \text{if } x \geq 1
\end{cases} \quad (2.5) \]

Next, for simplicity, choose \( M > b \). Then, the proposed limited translation rule, corresponding to the modified Bayes rule as proposed in (2.5) is

\[ \delta_{M,\alpha,k}(x) = \begin{cases} 
  x(1-a) + b & \text{if } 1 < x < \frac{(M+b)}{a} \\
  x-M & \text{if } x > \frac{(M+b)}{a}
\end{cases} \quad (2.6) \]

The above estimate is the same as the ones proposed in (2.3) and (2.4), i.e., when \( k=1 \) and \( k < 1 \), respectively. Henceforth, we consider the estimate proposed in (2.6) as the limited translation estimate.
B. Risk, Bayes Risk, and Relative Saving Loss of the Limited Translation Rule (LTR)

1. Risk of the LTR

First, we derive the risk function of the LTR. Recall \( a = \alpha/(1+\alpha) \) and \( b = (k-1)/(1+\alpha) \).

Theorem 2.1 Let \( c \) be the greatest integer less than or equal to \( (M + b)/a \), i.e., \( c = \lfloor (M + b)/a \rfloor \), then the risk of the LTR is

\[
R(\theta, \delta, M, \alpha, k) = e^{-\theta} + \frac{(\theta - 2b + b^2(\theta - 1))}{(1+\theta)}(F_\theta(c) - F_\theta(0)) + (1-a)^2 \theta F_\theta(c-2) - \left[ 2(\theta - b)(1-a) - (1-a)^2 \right] F_\theta(c-1) + (\theta + 2M + \theta^{-1})(1-F_\theta(c)) + \theta (1-F_\theta(c-2)) - \\
(2M + 2\theta)(1-F_\theta(c-1))
\]

(2.7)

where \( F_\theta(x) = \sum_{j=0}^{\infty} e^{-\theta j}/j! \).

Proof. Write \( p_\theta(x) = e^{-\theta x}/x! \). Then,

\[
R(\theta, \delta, M, \alpha, k) = E_\theta(\theta - \delta, M, \alpha, k(x))^2/\theta \\
= \frac{1}{\theta} \left[ \theta e^{-\theta} + \sum_{x=1}^{c} (\theta - b - x(1-a))^2 p_\theta(x) + \\
\sum_{x=c+1}^{\infty} (\theta - (x-1))^2 p_\theta(x) \right]
\]

(2.8)

Write \( T_1 = \sum_{x=1}^{c} (\theta - b - x(1-a))^2 p_\theta(x) \), \( T_2 = \sum_{x=c+1}^{\infty} (\theta - (x-M))^2 p_\theta(x) \), and assume \( c \geq 2 \).

\[
T_1 = \sum_{x=1}^{c} (\theta - b - x(1-a))^2 p_\theta(x) = (1-a)^2 \sum_{x=1}^{c} (\theta - b)^2/(1-a)^2 - 2(\theta - b)x/(1-a) + x^2 p_\theta(x)
\]

\[
= (\theta - b)^2 (F_\theta(c) - F_\theta(0)) + (1-a)^2 \sum_{x=1}^{c} [x(x-1) - x(2(\theta - b)/(1-a) - 1)] p_\theta(x)
\]

\[
= (\theta - b)^2 (F_\theta(c) - F_\theta(0)) + (1-a)^2 \left[ \theta^2 \sum_{x=0}^{c-1} p_\theta(x) - (2(\theta - b)/(1-a) - 1) \theta x \sum_{x=0}^{c-1} p_\theta(x) \right]
\]

\[
= (\theta - b)^2 (F_\theta(c) - F_\theta(0)) + (1-a)^2 \theta^2 F_\theta(c-2) - (1-a)(2(\theta - b) - (1-a))(1-F_\theta(c-1))
\]

\[
T_2 = \sum_{x=c+1}^{\infty} (\theta + M)^2 p_\theta(x) = \sum_{x=c+1}^{\infty} [(\theta + M)^2 - 2x(\theta + M) + x^2] p_\theta(x)
\]

(2.9)

\[
= (\theta + M)^2 (1-F_\theta(c)) + \sum_{x=c+1}^{\infty} [x(x-1) - x(2(\theta + M) - 1)] p_\theta(x)
\]
Hence, for $c \geq 2$

$$R(\theta, \delta_{M, \alpha, k}) = \theta e^{-\theta} + \frac{1}{\theta}(T1+T2)$$

$$= \theta e^{-\theta} + (\theta-2b+b^2\theta^{-1})(F_\theta(c)-F_\theta(0)) + (1-a)^2\theta F_\theta(c-2) - [2(\theta-b)(1-a)-(1-a)^2]F_\theta(c-1) + (\theta+2M+M^2\theta^{-1})(1-F_\theta(c)) +$$

$$\theta(1-F_\theta(c-2)-(2(\theta+M)-1)(1-F_\theta(c-1))$$

(2.11)

When $c=0$, the risk is

$$= \frac{1}{\theta}[\theta^2 e^{-\theta} + \sum_{x=1}^{\infty} (\theta-(x-M))^2 p_\theta(x)]$$

$$= \frac{1}{\theta}[(\theta-b-(1-a))^2 \theta e^{-\theta} + \sum_{x=1}^{\infty} (\theta-(x-M))^2 p_\theta(x)]$$

$$= \theta e^{-\theta} - (\theta+M)^2 \theta e^{-\theta} + \frac{\theta}{\theta}(\text{Var}_\theta(X-M)+(E_\theta(X-M)-\theta)^2)$$

$$= \theta e^{-\theta} - (\theta+M)^2 \theta e^{-\theta} + \frac{\theta}{\theta}[\theta+M]^2$$

$$= 1+M^2 \theta^{-1}(1-e^{-\theta})-2Me^{-\theta}$$

(2.12)

which agrees with the rhs of (2.7) when $c=0$.

When $c=1$, the risk is

$$= \frac{1}{\theta}[\theta^2 e^{-\theta} + (\theta-b-(1-a))^2 \theta e^{-\theta} + \sum_{x=1}^{\infty} (\theta-(x-M))^2 p_\theta(x)]$$

$$= \frac{1}{\theta}[(\theta-b-(1-a))^2 \theta e^{-\theta} - (\theta-(1-M))^2 \theta e^{-\theta} + \sum_{x=1}^{\infty} (\theta-(x-M))^2 p_\theta(x)]$$

$$= \theta e^{-\theta} [b+(1-a)]^2 - (1-M)^2 + 2\theta e^{-\theta} [(1-M)-(b+(1-a))] +$$

$$1+M^2 \theta^{-1}(1-e^{-\theta})-2Me^{-\theta}$$

(2.13)

which agrees with the expression in the rhs of (2.7) when $c=1$. The proof of the theorem is complete.

Theorem 2.2 The risk of the limited translation rule is bounded.

Further,

$$\lim_{\theta \to \infty} R(\theta, \delta_{M, \alpha, k}) = 1$$

(2.14)
and,
\[
\lim_{\theta \to 0} R(\theta, \delta_M, \alpha, k) = \begin{cases} 
(1-M)^2 & \text{if } c = 0 \\
(b+(1-a))^2 & \text{if } c \geq 1.
\end{cases}
\tag{2.15}
\]

Proof. First, rewrite (2.7) as
\[
R(\theta, \delta_M, \alpha, k) = (b^2-M^2)\theta^{-1}\left(F_\theta(c)-F_\theta(0)\right) + 2\left(ac-(b+M)\right)F_\theta(c) +
2bF_\theta(0) + a^2\theta F_\theta(c-2) + (a^2-2ba-2a)F_\theta(c-1) + 1 + M^2\theta^{-1}\left(1-F_\theta(0)\right)
\tag{2.16}
\]
Hence, for \(c=0\),
\[
R(\theta, \delta_M, \alpha, k) = 1 + M^2\theta^{-1}(1-\exp(-\theta)) - 2M\exp(-\theta) \leq 1 + M^2
\tag{2.17}
\]
noting that \(\theta^{-1}(1-\exp(-\theta)) \to \infty \) for \(\theta > 0\).

Now, \(c=1\) implies \(a \leq b+M\) and consequently \((b^2-M^2) \leq 0\). From (2.16), one gets
\[
R(\theta, \delta_M, \alpha, k) \leq (a^2-2a+2b(1-a))F_\theta(0) + 1 + M^2
\tag{2.18}
\]
For \(c \geq 2\), note \(ac \leq M+b\), \((b^2-M^2) \leq 0\), \((b-M)(c-1)^{-1}+a \leq (2b-a)(c-1)^{-1}\)
and \((b^2-M^2)\theta^{-1}\left(F_\theta(c)-F_\theta(0)\right) \leq (b-M)(c-1)^{-1}a\theta F_\theta(c-2)\). Using (2.16), we have
\[
R(\theta, \delta_M, \alpha, k) \leq 1 + M^2\theta^{-1}(1-\exp(-\theta)) + 2bF_\theta(0) + (2b-a)(c-1)^{-1}a\theta F_\theta(c-2) +
(a^2-2a-2ba)F_\theta(c-1)
\tag{2.19}
\]
Note that when \(b \leq 0\) then \((a^2-2a-2ba) = a(1-a)(2k+a(1-a)^{-1}) < 0\)
and so the above risk is less than or equal to \(1+M^2\). Otherwise, since \(\theta(c-1)^{-1}F_\theta(c-2) \leq F_\theta(c-1)-F_\theta(0)\) one gets from the above expression
\[
R(\theta, \delta_M, \alpha, k) \leq 1 + M^2\theta^{-1}(1-\exp(-\theta)) + 2b(1-a)F_\theta(0) + (a^2-2a)F_\theta(c-1)
\leq 1 + M^2 + 2b(1-a)
\tag{2.19}
\]
Also, taking limits as \(\theta \to \infty\) in (2.16), (2.14) follows. Taking limits as \(\theta \to 0\) in (2.16), one gets for \(c=0\),
\[
\lim_{\theta \to 0} R(\theta, \delta_M, \alpha, k) = 1 - 2M + M^2 = (1-M)^2
\tag{2.20}
while for \( c \geq 1 \),

\[
\lim_{\theta \to 0} R(\theta, \delta_{M,\alpha,k}) = 1 + b^2 + 2b(1-a) + (1-a)^2 - 1 = (b+1-a)^2. 
\] (2.21)

**Remark.** The bounds obtained in Theorem 2.2 are in general very conservative in the sense that usually the supremum of the risk function is much smaller than the bounds given in (2.17) and (2.19). This will be revealed later in the numerical computations undertaken.

2. **Bayes risk of the LTR**

In order to find the relative savings loss of the LTR, we compute the Bayes risk of such estimators with respect to the Gamma \((\alpha,k)\) prior.

**Theorem 2.3** The Bayes risk of the limited translation rule with respect to the Gamma \((\alpha,k)\) prior denoted by \( r(g_{\alpha,k}, \delta_{M,\alpha,k}) \) is given by

\[
r(g_{\alpha,k}, \delta_{M,\alpha,k}) = 1 - 2b[H_k(c) - H_k(0)] + 2\alpha^{-1}h_{k+1}(c-1) \\
+ a^2\alpha^{-1}h_{k+1}(c-2) + \{2b(1-a) + (1-a)^2 - 1\} H_k(c-1) \\
- 2Mh_k(c) + b^2 \overline{I}(k) + M^2 I(k), 
\] (2.22)

where

\[
\overline{I}(k) = \frac{c + k}{x!} \Gamma(k) \Gamma(x+k-1)(1+\alpha)^{-x+k-1} \\
I(k) = \begin{cases} 
\frac{c + k}{x!} \Gamma(k) \Gamma(x+k-1)(1+\alpha)^{-x+k} & \text{if } k \neq 1 \\
-\log \alpha - \frac{c}{x!} \Gamma(k) & \text{if } k = 1 
\end{cases}, \\
H_k(c) = \frac{\Gamma(x+k)}{x!} \Gamma(k) a^{k}(1-a)^x, \text{ and} \\
h_k(c) = \frac{\Gamma(x+k+1)}{c! \Gamma(k)} a^{k}(1-a)^c.
\]

**Proof.** Note that

\[
r(g_{\alpha,k}, \delta_{M,\alpha,k}) = \int_0^{\infty} R(\theta, \delta_{M,\alpha,k}) \exp(-\alpha\theta) a^k \theta^{k-1} \Gamma(k)^{-1} d\theta 
\] (2.23)

Take \( c \geq 2 \). If \( 0 \leq c < 2 \) the terms becoming zero in (2.16) are the ones in (2.22). First,
\[
\begin{align*}
&\int_{0}^{\infty} [F_\theta(c) - F_\theta(0)] e^{-\alpha \theta} \alpha^k k^{-1} (\Gamma(k))^{-1} d\theta \\
&= -\int_{0}^{\infty} \sum_{x=1}^{\infty} e^{-\theta x} \frac{\alpha^k}{x!} \frac{\Gamma(x+k-1)}{\Gamma(k)} x^{-1} d\theta \\
&= \frac{c}{x=1} \frac{\alpha^k}{x! \Gamma(k)} \int_{0}^{\infty} e^{-(a+1)\theta} x^{k-1} d\theta \\
&= \frac{c}{x=1} \frac{\alpha^k}{x! \Gamma(k)} (a+1)^{-k} \frac{\Gamma(x+k-1)}{\Gamma(x+k)} = I(k) \hspace{2cm} (2.24)
\end{align*}
\]

Next,
\[
\begin{align*}
&\int_{0}^{\infty} \theta p_\theta(c-1) e^{-(a-\theta) \alpha^k k^{-1} (\Gamma(k))^{-1} d\theta} \\
&= \int_{0}^{\infty} \exp(-(a+1)\theta) \frac{\alpha^k k^c k^{-1}}{(c-1)! \Gamma(k)} d\theta \\
&= a^k (a+1)^{-c} \frac{\Gamma(c+k)}{(c-1)! \Gamma(k)} = a^k (1-a)^{-c} \frac{\Gamma(k+c)}{(c-1)! \Gamma(k+1)} \frac{k}{\alpha} \\
&= k \frac{h_{c+1}}{\alpha} \hspace{2cm} \frac{2.25}{
\end{align*}
\]

Similarly,
\[
\begin{align*}
&\int_{0}^{\infty} p_\theta(c) e^{-(a-\theta) \alpha^k k^{-1} (\Gamma(k))^{-1} d\theta} = a^k (1-a)^{-c} \frac{\Gamma(k+c)}{c! \Gamma(k)} = h_{c} \hspace{2cm} (2.26)
\end{align*}
\]

Finally,
\[
\begin{align*}
&\int_{0}^{\infty} \theta F_\theta(c-1) e^{-(a-\theta) \alpha^k k^{-1} (\Gamma(k))^{-1} d\theta} = H_{c-1} \hspace{2cm} (2.28)
\end{align*}
\]

Finally,
\[
\begin{align*}
&\int_{0}^{\infty} \theta^{-1} [1-F_\theta(c)] e^{-(a-\theta) \alpha^k k^{-1} (\Gamma(k))^{-1} d\theta} \\
&= \int_{x=c+1}^{\infty} e^{-c+1} \frac{\alpha^k x^{k-2}}{x! \Gamma(k)} x^{-1} d\theta \\
&= \int_{x=c+1}^{\infty} \alpha^k \frac{\Gamma(x+k-1)(a+1)}{x! \Gamma(k)} = H_{c-1} \hspace{2cm} (2.29)
\end{align*}
\]
\[
\sum_{x=1}^{\infty} \frac{\Gamma(x+k-1)}{x! \Gamma(k)} a^k (1-a)^{x-1} - \frac{c}{\sum_{x=1}^{\infty} \frac{\Gamma(k-1+x)}{x! \Gamma(k)}} a^k (1-a)^{x-1}
\]  

(2.30)

Next, observe that
\[
\sum_{x=0}^{\infty} \frac{\Gamma(x+k)}{x! (x+1)! \Gamma(k)} a^k (1-a)^x
\]
\[
= \int_0^1 \left( \sum_{x=0}^{\infty} \frac{\Gamma(x+k)}{x! \Gamma(k)} a^k (1-a) u^x \right) du
\]
\[
= a^k \int_0^1 \left( \sum_{x=0}^{\infty} \frac{\Gamma(x+k)}{x! \Gamma(k)} (1-a) u^x \right) du = a^k \int_0^1 (1-(1-a) u)^{-k} du
\]
\[
= I \text{ (say)}
\]  

(2.31)

For \( k=1 \),
\[
I = a \int_0^1 \frac{du}{1-u(1-a)} = - \frac{a}{1-a} \log a = -a \log a,
\]  

(2.32)

while for \( k \neq 1 \),
\[
I = - \frac{a^k}{(1-a)(-k+1)} [a^{-k+1} - 1]
\]
\[
= \frac{a^k}{(1-a)(k-1)} [a^{-k+1} - 1]
\]
\[
= \frac{a^k}{(1-a)(k-1)} - \frac{a^k}{(1-a)(k-1)} = \frac{a}{b} (1-a)^{k-1}.
\]  

(2.33)

Combining (2.16) and (2.23) - (2.33), one gets (2.22).

Two lemmas are needed to simplify the Bayes risk, so we prove them first.

Lemma 2.1 If \( k > 1 \), and \( h_k(.) \) is the cumulative distribution function of a Negative Binomial random variable with parameter \( (a,k) \), then,
\[ H_k(x) = aH_{k-1}(x) + (1-a)H_k(x-1) \]

**Proof.**

\[
H_k(x) = \sum_{j=0}^{\infty} \frac{\Gamma(k+j+1)}{j! \Gamma(k)} a^k (1-a)^j x^j
= \sum_{j=0}^{\infty} \frac{\Gamma(k+j-1)}{j! \Gamma(k-1)} a^k (1-a)^j x^j + \sum_{j=1}^{\infty} \frac{\Gamma(k+j-1)}{(j-1)! \Gamma(k)} a^k (1-a)^j x^j
= aH_{k-1}(x) + x \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{j! \Gamma(k)} a^k (1-a)^j x^j + (1-a)H_k(x-1).
\]

**Lemma 2.2** If \( k > 0 \), and \( H_k(.) \) is as in Lemma 2.1, then

\[
H_k(m) = \int_a^1 \frac{u^{k-1} (1-u)^{m-1}}{B(k, m+1)} du = \int_0^a \frac{u^{k-1} (1-u)^{m-1}}{B(k, m+1)} du
\]

where \( B(k, m+1) \) is the Beta function.

**Proof.** Put \( u = az \), then

\[
\int_0^a \frac{u^{k-1} (1-u)^{m-1}}{B(k, m+1)} du = \frac{a^k}{B(k, m+1)} \int_0^1 z^{k-1} (1-az)^{m-1} dz
= \frac{a^k}{B(k, m+1)} \int_0^1 z^{k-1} (1-a)z^{1-z} \left( z + \sum_{x=0}^{m-1} \frac{(m)}{x} x! (1-a)^x z^x \right) dz
= \frac{a^k}{B(k, m+1)} \sum_{x=0}^{m} \frac{(m)}{x} (1-a)^x \int_0^1 z^{x+k-1} (1-z)^{m-x} dz
= \frac{a^k}{\Gamma(k)} \sum_{x=0}^{m} \frac{(m)}{x} (1-a)^x \Gamma(x+k, m-x+1)
= \frac{a^k}{\Gamma(k)} \sum_{x=0}^{m} \frac{1}{x!} (1-a)^x \Gamma(x+k)
= H_k(m).
\]

The integral in the rhs of Lemma 2.2 will be denoted by \( IB(a; k, m+1) \).

Next, note that for \( k > 1 \)

\[
\mathcal{I}(k) = \sum_{x=1}^{\infty} \left( \frac{1}{x! \Gamma(k)} \frac{\Gamma(x+k-1)}{(1+\alpha)^x} \right)
\]
Similarly,

\[ \overline{\Phi}(k) = ab^{-1}(1-a^{k-1}) - ab^{-1}[H_{k-1}(c) - H_{k-1}(0)] \]

Substituting the above expressions for \( \Phi(k) \) and \( \overline{\Phi}(k)(k > 1) \) in the rhs of (2.22) leads to

\[
\begin{align*}
\tau(g_{\alpha,k}, \delta_{\alpha,k}) &= 1 + 2ab^{k-2bH_{k-1}(c)} + 2(b+1-a)[H_{k+1}(c-1) - H_{k+1}(c-2)] \\
&\quad + a(b+1-a)H_{k+1}(c-2) + [2b(1-a) + a^2 - 2a]H_{k-1}(c-1) \\
&\quad - 2M[H_{k-1}(c) - H_{k-1}(c-1)] + ab[H_{k-1}(c) - a^{k-1}] + \\
&\quad M^2 ab^{-1}(1-H_{k-1}(c)) .
\end{align*}
\]

(2.34)

Using Lemma 2.1 and Lemma 2.2 for \( k > 1 \), one gets from (2.34),

\[
\begin{align*}
\tau(g_{\alpha,k}, \delta_{\alpha,k}) &= 1 + ba^k + M^2 ab^{-1} - 2(b+M)[aH_{k-1}(c) + (1-a)H_{k-1}(c-1)] \\
&\quad + 2a(b+1-a)[H_{k-1}(c-1) - H_{k-1}(c-2)] + a(b+1-a)H_{k+1}(c-2) \\
&\quad + [2b(1-a) + a^2 - 2a + 2M]H_{k-1}(c-1) + abH_{k-1}(c) - M^2 ab^{-1}H_{k-1}(c) \\
&\quad = 1 + ba^k + M^2 ab^{-1} - 2a(b+M)IB(a;k-1,c+1) \\
&\quad + [-2(b+M)(1-a) + 2a(b+1-a)]IB(a;k,c) - a(b+1-a)IB(a;k+1,c-1) \\
&\quad + [2b(1-a) + a^2 - 2a + 2M]IB(a;k,c) + (ab - M^2 ab^{-1})IB(a;k-1,c+1) \\
&\quad = 1 + ba^k + M^2 ab^{-1} - a(b+1-a)IB(a;k+1,c-1) \\
&\quad - (b+M)^2 ab^{-1}IB(a;k-1,c+1) + [2(b+M) - a]aIB(a;k,c).
\end{align*}
\]

(2.35)
For \( k=1 \),

\[
\begin{align*}
    r(g_{\alpha,1}, \delta_M, \alpha, 1) &= 1 + \frac{2a}{a} h_2^-(c-1) + \frac{a^2}{a} H_2^-(c-2) + (a^2 - 2a) H_1 (c-1) \\
    &\quad - 2Mh_1 (c) + M^{2-1}(1) \\
    &= 1 + 2(1-a) h_2^-(c-1) + a(1-a) H_2^-(c-2) + (a^2 - 2a) H_1 (c-1) \\
    &\quad - 2Mh_1 (c) + M^{2-1}(1) \\
    &= 1 + 2c a^2 (1-a)^2 + a(1-a) \left[ 1 - \frac{a}{x^{c-1}} (x+1) a^2 (1-a)^x \right] \\
    &\quad + (a^2 - 2a) \left[ 1 - \frac{a}{x^{c-1}} a(1-a)^x \right] - 2Ma(1-a)^c \\
    &\quad + M^2 \left[ a \log a - \frac{c}{x^{c-1}} (1-a)^x \right] \\
    &= 1 + 2ca^2 (1-a)^2 + a(1-a) \frac{x}{x^{c-1}} (1-a)^{x-1} \\
    &= 1 - 2a(1-a)c + a(1-a) \frac{x}{x^{c-1}} (1-a)^{x-1} \\
    &\quad - (a^2 - 2a)(1-a)^c - 2 Ma(1-a)^c - 1M^2 [\log a + \frac{c}{x^{c-1}} (1-a)^{x-1}] \\
    &= 1 + 2ca^2 (1-a)^2 + a(1-a) \frac{x}{x^{c-1}} (1-a)^{x-1} \\
    &\quad - (a^2 - 2a)(1-a)^c - 2 Ma(1-a)^c - 1M^2 [\log a + \frac{c}{x^{c-1}} (1-a)^{x-1}] \\
    &= 1 + 2ca^2 (1-a)^2 + a(1-a) \frac{x}{x^{c-1}} (1-a)^{x-1} \\
    &\quad - a(1-a)^{-1} M^2 [\log a + \frac{c}{x^{c-1}} (1-a)^{x-1}] \tag{2.36}
\end{align*}
\]

where in the penultimate step, we have used \( \sum_{x=c}^{\infty} x (1-a)^{x-1} = a(1-a)^{c-1} + (1-a)c a^{-2} \).

For \( k < 1 \), we simplify (2.22) slightly into

\[
\begin{align*}
    r(g_{\alpha,k}, \delta_M, \alpha, k) &= 1 + 2ba^k + ka(1-a) H_{k+1}^-(c-2) \\
    &\quad + \{a - 2(1+b)\} a H_k (c-1) - 2(b+Maac) H_k^-(c) \\
    &\quad + b^2 I(k) + M^{2-1}(k). \tag{2.37}
\end{align*}
\]
3. Relative saving loss of the LTR

For any rule $\delta$, we defined the relative saving loss (RSL) in the introduction as

$$RSL(g_{\alpha,k}, \delta) = \frac{r(g_{\alpha,k}, \delta) - r(g_{\alpha,k}, \delta^0)}{r(g_{\alpha,k}, \delta^0) - r(g_{\alpha,k}, \delta_{\alpha,k})}$$

which is the proportion of the possible Bayes risk improvement over $\delta^0$ that is sacrificed by the use of $\delta$ instead of the use of the Bayes rule with respect to the prior $g_{\alpha,k}$.

Since, in our case

$$1/(1+\alpha) = 1-\alpha$$

if $k \geq 1$

$$r(g_{\alpha,k}, \delta_{\alpha,k}) = \begin{cases} 
1-(a-ba^k) & \text{if } k < 1 \\
1 & \text{if } k \geq 1
\end{cases}$$

$$r(g_{\alpha,k}, \delta^0) = 1$$

Hence, the relative saving loss, for the limited translation rules $\delta_{M,\alpha,k}$, when $k=1$, is given by

$$RSL(g_{\alpha,1}, \delta_{M,\alpha,1}) = (1-\alpha)^c + ca(1-\alpha)^c - 2M(1-\alpha)^c + M^2(1-\alpha)^{-1}(loga + \sum_{x=1}^{c} (1-\alpha)^x/x)$$

For $k > 1$,

$$RSL(g_{\alpha,k}, \delta_{M,\alpha,k}) = 1 + ba^{k-1} + M^2b^{-1} - (b+(1-\alpha))IB(a;k+1,c-1) + (2(b+M)-a)IB(a;k,c) - (b+M)^2b^{-1}IB(a;k-1,c+1)$$

For $k \in (0,1)$

$$RSL(g_{\alpha,k}, \delta_{M,\alpha,k}) = (a+ba^k+ka(1-\alpha)H_{k+1}(c-2) + (a-2(1+b))aH_{k}(c-1) - 2(b+M-ca)h_k(c) + b^2I(k) + M^2I(k))/(a-ba^k)$$

Theorem 2.4 For $k \geq 1$, and $\delta_{M,\alpha,k}$ the relative saving loss as a
function of $M$ is a strictly decreasing function.

**Proof.** To prove this theorem, first it will be proved that for a fixed $c$, $k \geq 1$, and $\delta_{M,a,k}$ the relative saving loss as a function of $M$ is strictly decreasing function on its range which depends on $c$. Recall that $c = \gamma(M+b)/a$, then $c = m$ implies that $ma-b \leq M < a(m+1)-b$. Next, for two consecutive values of $c$, say $m$ and $m+1$, and at $M = a(m+1)-b$ we will show that the relative saving loss has the same value. Observe that the point $a(m+1)-b$ is the boundary between the sets $[ma-b, a(m+1)-b)$ and $[a(m+1)-b, a(m+2)-b)$ which are the ones where $M$ takes values when $c = m$ and $c = m+1$, respectively. Thus, the two facts before show that the RSL of $\delta_{M,a,k}$ is strictly decreasing on $M$.

\[
\begin{align*}
RSL(g_{\alpha,1}, \delta_{M,a,1}) &= (1-a)^c + ca(1-a)^c - 2M(1-a)^c - M^2(1-a)^{-1}\left(\log a \sum_{x=1}^{\infty} (1-a)^x/x\right) \\
&= (1-a)^c + ca(1-a)^c - 2M(1-a)^c - M^2(1-a)^{-1}\left(\log a \sum_{x=1}^{\infty} (1-a)^x/x\right)
\end{align*}
\]

Note that

\[
\log a = \log(1-(1-a)) = -\sum_{x=1}^{\infty} (1-a)^x/x
\]

Hence,

\[
\begin{align*}
RSL(g_{\alpha,1}, \delta_{M,a,1}) &= (1-a)^c + ca(1-a)^c - 2M(1-a)^c + M^2(1-a)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x/x \\
&= (1-a)^c + ca(1-a)^c - 2M(1-a)^c + M^2(1-a)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x/x
\end{align*}
\]

We want to prove that $\frac{d}{dM} RSL(g_{\alpha,1}, \delta_{M,a,1}) < 0$ which is equivalent to showing that

\[
M(1-a)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x/x < (1-a)^c
\]

For a fixed $c$, we have

\[
c \leq M/a < c+1 \Rightarrow ac \leq M < a(c+1), \text{ then}
\]

\[
M(1-a)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x/x < a(c+1)(1-a)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x/x
\]

\[
< a(c+1)(1-a)^{-1}(c+1)^{-1}\sum_{x=c+1}^{\infty} (1-a)^x
\]

\[
= a(1-a)^{-1}(1-a)^{c+1}a^{-1} = (1-a)^c
\]
Next consider the case $k > 1$. Here,

$$\frac{d}{dM} RSL(g_{α,k,δ M,α,k}) = 2Mb^{-1}(1-IB(a;k-1,c+1)) + 2(IB(a;k,c)-IB(a;k-1,c+1))$$

But,

$$IB(a;k,c) = \int_0^a u^{-k-1}(1-u)^{-c-1}/B(k,c) du$$

Integrating by parts with $y = u^{-k-1}$ and $dv = (1-u)^{-c-1} du$,

$$\int_0^a u^{-k-1}(1-u)^{-c-1}/B(k,c) du = \left[B(k,c)\right]^{-1}\left\{[u^{-k-1}(1-u)^{-c}/c]\right|_0^a + \int_0^a (k-1)/c u^{-k}(1-u)^{-c-1}(1-u) du\right\}$$

$$= -\alpha^{-k-1}(1-a)^{-c}/B(k,c)c + \int_0^a (k-1)(1-u)(c+1)^{-1}/B(k-1,c+1) du$$

$$= -\Gamma(k+c)\alpha^{-k-1}(1-a)^{-c}/\Gamma(k)c! + IB(a;k-1,c+1)$$

then,

$$\frac{d}{dM} RSL(g_{α,k,δ M,α,k}) = 2Mb^{-1}(1-IB(a;k-1,c+1)) - \frac{2\Gamma(k+c)\alpha^{-k-1}(1-a)^{-c}}{\Gamma(k)c!}$$

We want to prove that

$$Mb^{-1}(1-IB(a;k-1,c+1)) < \Gamma(k+c)\alpha^{-k-1}(1-a)^{-c}/\Gamma(k)c!$$

Recall $c < (Mb)/a < c+1$ which implies $M < a(c+1) - b$. Now,

$$1-IB(a;k-1,c+1) = \int_a^1 u^{-k-2}(1-u)^{-c}/B(k-1,c+1) du$$

Integrating by parts with $y = u^{-k-2}$ and $dv = (1-u)^{-c} du$,

$$\int_a^1 u^{-k-2}(1-u)^{-c} du = \frac{k-2}{c+1}(1-a)^{c+1} + \int_a^1 u^{-k-3}(1-u)^{c+1} du$$

Hence,

$$Mb^{-1}(1-IB(a;k-1,c+1)) < \frac{(a(c+1)-b)^{-1}\int_a^1 u^{-k-2}(1-u)^{-c} du}{B(k-1,c+1)}$$

$$= \frac{a^{-k-1}(1-a)^{c+1}b^{-1}}{B(k-1,c+1)} + T$$

$$= \frac{a^{-k-1}(1-a)^{c}\Gamma(k+c) + T}{\Gamma(k)c!}$$

where
Hence, it suffices to show that $T < 0$. For $1 < k < 2$ the result is clear since the left term in $T$ is nonpositive while the right term is positive. For $k > 2$,

$$a(k-2)b^{-1} \int_{a}^{u} k^{-3} (1-u)^{c+1} du = a(k-2)b^{-1} \int_{a}^{u} k^{-2} (1-u)^{c} (u-1) du$$

$$\leq a(k-2)b^{-1} \int_{a}^{u} k^{-2} (1-u)^{c} du$$

$$= (k-2)b^{-1} (1-a) \int_{a}^{u} k^{-2} (1-u)^{c} du$$

$$\leq (k-1)(1-a)b^{-1} \int_{a}^{u} k^{-2} (1-u)^{c} du$$

$$= \int_{a}^{u} k^{-2} (1-u)^{c} du$$

Now, taking expression (2.22) into the relative saving loss, this can be written as

$$RSL(g_{u,k}, \delta_{M,u,k}) = 1-2ba^{-1}(H_{k}(c)-H_{k}(0)) + 2ka^{-1}h_{k+1}(c-1)$$

$$+ a^{-1}H_{k+1}(c-2) + a^{-1}\{2b(1-a)+(1-a)^{2}-1\} H_{k}(c-1)$$

$$- 2Ma^{-1}h_{k}(c) + b^{-1}I(k) + a^{-1}n^{-1}(k)$$

We wish to prove that the relative saving loss evaluated at $M = a(m+l)-b$ has the same value when $c = m$ and when $c = m+1$. Evaluating the difference between the RSL when $c = m+1$ and $c = m$ at $M = a(m+l)-b$ we have

$$-2ba^{-1}h_{k}(m+1)+2ka^{-1}(h_{k+1}(m)-h_{k+1}(m-1))$$

$$+ a^{-1}h_{k+1}(m-1)+a^{-1}\{2b(1-a)+(1-a)^{2}-1\} h_{k}(m)$$

$$-2(a(m+1)-b)a^{-1}(h_{k}(m+1)-h_{k}(m)) + b^{-1}(m+1)^{-1}h_{k}(m)$$

$$-a^{-1}(a(m+1)-b)^{2}(m+1)^{-1}h_{k}(m)$$

since $aa^{-k-1}(m+1)!/(m+1)! = (m+1)^{-1}h_{k}(m)$. Simplifying, the above difference is

$$2ka^{-1}(h_{k+1}(m)-h_{k+1}(m-1))-2h_{k}(m)$$

$$-2(m+1)(h_{k}(m+1)-h_{k}(m)) + ak^{-1}h_{k+1}(m-1)-amh_{k}(m)$$
If \( m = 0 \), then the last expression is zero. Also, when \( m \geq 1 \) the difference is zero since
\[
\begin{align*}
   h_{k+1}(m-1) &= \frac{m}{k} h_k(m), \\
   h_{k+1}(m) &= (m+k) \frac{1}{k} h_k(m), \text{ and} \\
   h_k(m+1) &= (m+k)(1-a)(m+1) \frac{1}{k} h_k(m)
\end{align*}
\]
and so, the proof of the theorem is complete.

C. Generalizations

1. **Sample size equal to \( n \)**

First, for completeness we will compute the risk and Bayes risk of the rules mentioned before when we have a sample of size \( n \). Then, we will generalize our results for the case of estimating \( p \) Poisson means.

Let \( X_1, \ldots, X_n \) be a sample from a Poisson distribution with mean \( \theta \), and as before let us assign to \( \theta \) a prior Gamma \((a,k)\) distribution with probability density function denoted by \( g_{a,k} \).

Note, that the minimal sufficient statistic is \( T = \frac{1}{n} \sum_{i=1}^{n} X_i \sim \text{P}(n\theta) \) and under the above prior \( n\theta \) has a Gamma \((a/n,k)\) distribution. In this case the maximum likelihood estimator of \( n\theta \) is \( T \) and the Bayes estimator is \( (T+k-1)/(1+a/n) \). Equivalently, the maximum likelihood estimator of \( \theta \) is \( T/n \) and the Bayes estimate is \( (T+k-1)/(n+a) \).

As before, assume the loss \( L(\theta, a) = (\theta-a)^2/\theta \).

Denote the maximum likelihood estimator of \( \theta \) by \( \delta^0(X) = T/n = \frac{1}{n} \sum_{i=1}^{n} X_i/n \) and the Bayes estimator by \( \delta_{a,k}(X) = (T+k-1)/(n+a) \).

Then,
\[
\begin{align*}
   i) \quad &R(\theta, \delta^0) = E_\theta (T/n-\theta)^2/\theta = n^{-1} E_\theta (T-n\theta)^2/n\theta = n^{-1}; \\
   ii) \quad &r(g_{a,k}, \delta^0) = ER(\theta, \delta^0) = n^{-1}; \\
   iii) \quad &R(\theta, \delta_{a,k}) = E_\theta [(T+k-1)/(n+a)-\theta]^2/\theta
\end{align*}
\]
\[ n^{-1}\mathbb{E}_\theta \left[ \frac{(T+k-1)/(1+\alpha/n) - n\theta}{n\theta} \right]^2/n\theta \]

Now, we are in a situation as before when we deal with one observation, only require to note that \(n\theta, \alpha/n\) and \(T\) are playing the role of \(\theta, \alpha\) and \(X\), respectively. So, to get the above risk we can use the result corresponding to the Bayes estimator based on one observation with the appropriate changes. Therefore,

\[ R(\theta, \delta_{\alpha,k}) = n^{-1}(1+\alpha/n)^{-2}(1+\alpha/n)^{-2}(n\theta)^{-1}(n\theta - (k-1)(\alpha/n)^{-1})^2 \]

\[ = (n+\alpha)^{-2}(n+\alpha)^{-2}(\theta - (k-1)\alpha^{-1})^2. \]

iv) \[ r(g_{\alpha,k}', \delta_{\alpha,k}) = \mathbb{E}_\theta \left[ \frac{(T+k-1)/(1+\alpha/n) - n\theta}{n\theta} \right]^2/n\theta \]

Since the Bayes risk of the Bayes estimator is \((1+\alpha)^{-1}\), then

\[ r(g_{\alpha,k}', \delta_{\alpha,k}) = n^{-1}(1+\alpha/n)^{-1} = (n+\alpha)^{-1} \]

v) The limited translation rule for this case will be

\[ \delta_{M,\alpha,k}(x) = \begin{cases} 0 & \text{if } t = 0 \\ \frac{t}{n} - M & \text{if } t > (M^1+b^1)/a^1 \\ \frac{t}{n} + M & \text{if } t < (M^1+b^1)/a^1 \end{cases} \]

Put \(M^1 = nM\), \(a^1 = (\alpha/n)/(1+\alpha/n)\) and \(b^1 = (k-1)/(1+\alpha/n)\), then

\[ \delta_{M,\alpha,k}(x) = \begin{cases} 0 & \text{if } t = 0 \\ \frac{t}{n} - M & \text{if } t \leq (M^1+b^1)/a^1 \\ \frac{t}{n} + M & \text{if } t > (M^1+b^1)/a^1 \end{cases} \]

If we call \(c^1 = [(M^1+b^1)/a^1]\), then

\[ R(\theta, \delta_{M,\alpha,k}) = \mathbb{E}_\theta \left[ \delta_{M,\alpha,k}(x) - \theta \right]^2/\theta \]

\[ = \mathbb{E}_n \left[ (\delta_{M,\alpha,k}(x) - \theta)^2/\theta \right] \quad (\delta_{M,\alpha,k} \text{ is a function of } T) \]
\[
R(\theta, \delta_{M, \alpha}, k) = n^{-1} n^1 e^{-n^0} + \sum_{t=1}^{n^1} \left( \left( \frac{(t+k-1)/(n^0) - \theta}{\theta} \right)^2 \right)
\]

where \(\delta_{M, \alpha/n, k}(t)\) is a limited translation rule of the kind introduced in part A., i.e., sample size equal to one, but now \(n^0, \alpha/n, t, M^1\) and \(c^1\) are playing the role of \(\theta, \alpha, x, M\) and \(c\), respectively. So, to calculate the above risk, we can use the expression obtained earlier for the limited translation rule with the appropriate changes. Of course, \(a^1\) and \(b^1\) will play the role of \(a\) and \(b\) for the case at hand.

Thus,

\[
R(\theta, \delta_{M, \alpha}, k) = n^{-1} (n^0 e^{-n^0} + (n^0 - 2n^0 + (b^1)^2 (n^0)^{-1}) (F_T(c^1) - F_T(0)) + (1-a^1)^2 (n^0)^{-1} (2(n^0 - b^1)(1-a^1)/-(1-a^1)^2) F_T(c^1 - 1) + (n^0 + 2M^1 + (M^1)^2 (n^0)^{-1}) (1 - F_T(c^1)) + n^0 (1 - F_T(c^1 - 2)) - (2M^1 - 1 + 2n^0) (1 - F_T(c^1 - 1)))
\]

where \(\delta_{M, \alpha/n, k}(t)\) is a limited translation rule of the kind introduced in part A., i.e., sample size equal to one, but now \(n^0, \alpha/n, t, M^1\) and \(c^1\) are playing the role of \(\theta, \alpha, x, M\) and \(c\), respectively. So, to calculate the above risk, we can use the expression obtained earlier for the limited translation rule with the appropriate changes. Of course, \(a^1\) and \(b^1\) will play the role of \(a\) and \(b\) for the case at hand.

Thus,

\[
R(\theta, \delta_{M, \alpha}, k) = n^{-1} (n^0 e^{-n^0} + (n^0 - 2n^0 + (b^1)^2 (n^0)^{-1}) (F_T(c^1) - F_T(0)) + (1-a^1)^2 (n^0)^{-1} (2(n^0 - b^1)(1-a^1)/-(1-a^1)^2) F_T(c^1 - 1) + (n^0 + 2M^1 + (M^1)^2 (n^0)^{-1}) (1 - F_T(c^1)) + n^0 (1 - F_T(c^1 - 2)) - (2M^1 - 1 + 2n^0) (1 - F_T(c^1 - 1)))
\]

where \(\delta_{M, \alpha/n, k}(t)\) is a limited translation rule of the kind introduced in part A., i.e., sample size equal to one, but now \(n^0, \alpha/n, t, M^1\) and \(c^1\) are playing the role of \(\theta, \alpha, x, M\) and \(c\), respectively. So, to calculate the above risk, we can use the expression obtained earlier for the limited translation rule with the appropriate changes. Of course, \(a^1\) and \(b^1\) will play the role of \(a\) and \(b\) for the case at hand.
\[ r(g_{\alpha,1}, \delta_M, \alpha, 1) = n^{-1}((1-a) + a(1-a) c^1 + c^1 (a^2) (1-a) c^1 - 2M a^1 (1-a) c^1 - \frac{(a/n)(M^1)^2}{t/(1-a)} \] 
\[ = n^{-1}((1-a) + n^{-1} a(1-a) c^1 + n^{-1} c^1 (a^2) (1-a) c^1 - 2M a^1 (1-a) c^1 - \frac{(a/n)(M^1)^2}{t/(1-a)} \]

For \( k > 1, \)
\[ r(g_{\alpha,k}, \delta_M, \alpha, k) = n^{-1}(1 + b^1 a^1 k + (M^1)^2 a^1 (b^1)^{-1} - (b^1 + (1-a^1) a^1 IB(a^1; k+1, c^1-1) + 2(b^1 + M^1 a^1 IB(a^1; k, c^1) - (b^1 + M^1 a^1 b^1)^{-1} IB(a^1; k-1, c^1+1)) \] 
\[ = n^{-1} b^1 a^1 k + (M^1)^2 a^1 (b^1)^{-1} IB(a^1; k+1, c^1-1) + 2(b^1 + M^1 - n^{-1} a^1 a^1 IB(a^1; k, c^1) - (b^1 + M^1)^2 a^1 (b^1)^{-1} IB(a^1; k-1, c^1+1) \]

Finally, for \( k < 1, \)
\[ r(g_{\alpha,k}, \delta_M, \alpha, k) = n^{-1}(1 + 2b^1 a^1 k + ka^1 (1-a^1) IB(a^1; k+1, c^1-1) + (a^1 - 2(b^1 + a^1) a^1 IB(a^1; k, c^1) - (b^1 + M^1 a^1 (b^1)^{-1} IB(a^1; k-1, c^1) + 2p^*_k(c^1) + (b^1)^2 \frac{n^1}{k+1} (M^1)^2 \frac{1}{k} (k) \] 
\[ = n^{-1} + 2b^1 a^1 k + n^{-1} k a^1 (1-a^1) IB(a^1; k+1, c^1+1) + (n^{-1} a^1 - 2(n^{-1} b^1) a^1 IB(a^1; k, c^1) - (b^1 + M^1 + n^{-1} (1-a^1) c^1)^2 p^*_k(c^1) + n^2 b^1 \frac{n^1}{k+1} n^2 \frac{1}{k} (k) \]

where
\[ p^*_k(c^1) = \frac{(k+c^1)^{k-1}}{c^1^k} (1-a) c^1 \]
\[ \frac{1}{k} = \frac{c^1}{t=k} \left( \frac{(a/n)^{k}}{(t+k-1)/(1+a/n)^{t+k-1}} \right) \]
\[ \frac{1}{k} = a^1 (b^1)^{-1} (1-(a^1)^{k-1})^{-1} \]

The corresponding relative saving losses can be written down from the above expression.

2. **Estimation of p Poisson means**

Let \( X_1, \ldots, X_p \) be \( p \) independent Poisson variables with means \( \theta_1, \theta_2, \ldots, \theta_p \) where \( \theta_i \in (0, \infty) \) is unknown for each \( i = 1, \ldots, p \).
Consider in this estimation problem a loss function of the form
\[ L(\theta_i, \delta_i) = \sum_{i=1}^{p} (\theta_i - \delta_i)^2 / \theta_i \]
where \( \theta = (\theta_1, \theta_2, \ldots, \theta_p) \) and \( a = (a_1, a_2, \ldots, a_p) \).

Let us assume \( \theta_1, \theta_2, \ldots, \theta_p \) are independent random variables with \( \theta_1 \)
having a Gamma \((\alpha, k_1)\) prior distribution.

Note that the maximum likelihood estimate \( \delta_i^0 \) and the Bayes estimate \( \delta_{a,k_1}^0 \)
of each component are
\[ \delta_i^0 (x) = x_i \quad \text{and} \quad \delta_{a,k_1}^0 (x) = \begin{cases} \frac{x_i + k_1 - 1}{1 + \alpha} & \text{if } x_i + k_1 - 1 > 0 \\ 0 & \text{otherwise} \end{cases} \]

As in part A., for a fixed \( M > 0 \), a limited translation rule for
each component can be given by
\[ \delta_{M,a,k_1} (x) = \begin{cases} 0 & \text{if } x_i = 0 \\ \frac{x_i + k_1 - 1}{1 + \alpha} & \text{if } 1 \leq x_i \leq \frac{M + b_1}{a} \\ x_i - M & \text{if } x_i > \frac{M + b_1}{a} \end{cases} \]
where \( a = \alpha / (1 + \alpha) \) and \( b_1 = (k_1 - 1) / (1 + \alpha) \) and \( x = (x_1, x_2, \ldots, x_p) \).

So, taking a limited translation rule in each component can be thought as a compromise rule for estimating \( p \) means.

It is clear that the risk and Bayes risk of the compromise rule
(limited translation rule) in this case will be
\[ R(\theta_i, \delta_i) = \sum_{i=1}^{p} R(\theta_i, \delta_{M,a,k_1} (x)) \]
and
\[ r(g, \delta_i) = \sum_{i=1}^{p} r(g, \delta_{M,a,k_1} (x)) \]
where \( \delta_{M,a,k_1} (x) = (\delta_{M,a,k_1} (x_1), \delta_{M,a,k_1} (x_2), \ldots, \delta_{M,a,k_1} (x)) \) and \( g \) is the prior
distribution on \( (\theta_1, \theta_2, \ldots, \theta_p) \).

We have already computed \( R(\theta_i, \delta_{M,a,k_1} ) \) and \( r(g, \delta_{M,a,k_1} ) \) in
part B., so we only add up all of them to get risk and Bayes risk when
we estimate $p$ Poisson means.

Let $\hat{\theta}$ and $\hat{\theta}^0$ be the Bayes and maximum likelihood estimates for the above estimation problem of $p$ Poisson means, i.e.,

$\hat{\theta}(x) = (\delta_{\alpha, k_1}(x_1), \delta_{\alpha, k_2}(x_2), \ldots, \delta_{\alpha, k_p}(x_p))$ and

$\hat{\theta}^0(x) = x$

then, the relative saving loss is defined by

$$\text{RSL}(\hat{\theta}, \hat{\theta}^0) = \frac{r(\hat{\theta}, \hat{\theta}^0) - r(\hat{\theta}, \hat{\theta})}{r(\hat{\theta}, \hat{\theta}^0) - r(\hat{\theta}, \hat{\theta})}$$

Since we know all the terms in the above expression, we can calculate the relative saving loss of the rule $\hat{\theta}^0_M$ as in part B. Section 3.

Finally, if we take a sample of size $n_1$ from each Poisson population with mean $\theta_i$, i.e., if $X_{ij}(j=1, \ldots, n_i; i=1, \ldots, p)$ are independent with $X_{ij}(j=1, \ldots, n_i)$ i.i.d. Poisson $(\theta_i)$, $i=1, \ldots, p$ then we can redefine the problem as $T_1, \ldots, T_p$ are $p$ independent Poisson variables with means $\lambda_1, \lambda_2, \ldots, \lambda_p$, where $T_i = \sum_{j=1}^{n_i} X_{ij}$ and $\lambda_i = n_i \theta_i$ and use the formulae we obtained in part C. Section 1.

D. Conclusions

The limited translation rules are designed to perform much better than the MLEs around the prior mode and should not be much worse off than the MLEs at the tails. In part B., we showed that

$$\sup_{\theta > 0} R(\theta, \hat{\theta}^0_M, \alpha, k) \leq (1+M)^2 \quad \text{if } k \leq 1$$

$$\leq (1+M)^2 \quad \text{if } k > 1$$

These bounds are quite conservative in nature, and in actuality, the LTRs have as evidenced in Figures 2.1-2.4 upper bounds of the risks usually much smaller than the ones given in (2.38). Note also that $\text{RSL}(\hat{\theta}, \hat{\theta}^0_M, \alpha, k)$ is strictly decreasing in $M$. See Figures 2.5-2.10.
For the case $k < 1$, we have only numerical evidence. This is anticipated because, larger the value of $M$, closer the LTR gets to the usual Bayes estimator with respect to the conjugate prior and accordingly the sacrifice one makes in terms of Bayes risk savings by the use of the LTR rather than the Bayes estimator becomes smaller with increasing $M$. The triumph of the LTR as compared to the Bayes estimator lies in its robustness against misspecified priors, since its risk remains very close to the risk of the minimax estimator, namely the MLE.

E. Figures and Tables

Figures 2.1-2.4 show the risk as a function of $\theta$ of the limited translation rule for some values of $k$, $\alpha$ and $M$. Each figure has three graphs corresponding to three different values for $M$. Both parameters $k$ and $\alpha$ take two values, $k$ the values 3 and 7, $\alpha$ the values 1/3 and 3. For $k$ we take only values bigger than one since our interest is to observe the behavior of the risk around the prior mode $(k-1)\alpha^{-1}$. For $k \leq 1$, we have already seen the limited translation rules behave well. To distinguish the values of $M$ in each figure, we use three different symbols to join some points of the graph. Table 2.1 shows values of the risk displayed in Figures 2.1-2.4 for some values of $\theta$.

Figures 2.5-2.10 are plots of the relative saving loss against values of $M$ of the limited translation rule for some values of $\alpha$ and $k$. Figures 2.5-2.8 contain three graphs corresponding to three values of $\alpha$, i.e., $\alpha=1/3$, $\alpha=1$, and $\alpha=3$. As before, different values of $\alpha$ are identified by different symbols. Each figure uses a different value for $k$; these are $k=.25$, $k=.5$, $k=.75$, and $k=1$. Table 2.2 contains the
numerical values in Figures 2.5-2.8. Finally, Figures 2.9-2.10 each have three graphs corresponding to some combinations of $\alpha$ and $k$ values. These combinations are shown in each figure. Table 2.3 has the numerical values of Figures 2.9-2.10.
Triangle $M = 2$
Square $M = 4$
Diamond $M = 6$

**FIGURE 2.1** RISK OF THE LIMITED TRANSLATION RULE $\delta_{M,1/3,3}$
Triangle $M = 1$
Square $M = 5$
Diamond $M = 9$

Figure 2.2 Risk of the Limited Translation Rule $\delta_{M,3,3}$
Triangle  \( M = 4.5 \)
Square  \( M = 7 \)
Diamond  \( M = 9 \)

**FIGURE 2.3** RISK OF THE LIMITED TRANSLATION RULE \( \delta_{M,1/3,7} \)
Triangle $M = 2$
Square $M = 6$
Diamond $M = 10$

Figure 2.4 Risk of the Limited Translation Rule $\delta_{M,3,7}$
Triangle $\alpha = 1/3$
Square $\alpha = 1$
Diamond $\alpha = 3$

FIGURE 2.5 RELATIVE SAVING LOSS OF $\delta_{M,\alpha,.25}$
Triangle $\alpha = \frac{1}{3}$
Square $\alpha = 1$
Diamond $\alpha = 3$

Figure 2.6 Relative Saving Loss of $\delta_{M,\alpha,.5}$
Figure 2.7 Relative saving loss of $\delta_{N,\alpha,0.75}$

Triangle $\alpha = 1/3$
Square $\alpha = 1$
Diamond $\alpha = 3$
FIGURE 2.8 RELATIVE SAVING LOSS OF $\delta_{M,\alpha,1}$

Triangle $\alpha = 1/3$
Square $\alpha = 1$
Diamond $\alpha = 3$
VALUES OF M

Figure 2.9 Relative Saving Loss of $\delta_{M, a, k}$

Triangle $a=3, k=3$
Square $a=11, k=7$
Diamond $a=7, k=5$
VALUES OF $M$

FIGURE 2.10 RELATIVE SAVING LOSS OF $\delta_{M,\alpha,k}$

Triangle $\alpha=1/3, k=3$
Square $\alpha=3, k=7$
Diamond $\alpha=25/15, k=5$
### Table 2.1 Values for the Risk of the Limited Translation Rule

\((k > 1)\)

<table>
<thead>
<tr>
<th>(M)</th>
<th>(\theta)</th>
<th>(2)</th>
<th>4</th>
<th>6</th>
<th>(\alpha=1/3, k=3)</th>
<th>(\alpha=3, k=3)</th>
<th>(\alpha=1/3, k=7)</th>
<th>(\alpha=3, k=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.001</td>
<td>5.058</td>
<td>5.058</td>
<td>5.058</td>
<td>0.562</td>
<td>0.562</td>
<td>0.562</td>
<td>27.544</td>
</tr>
<tr>
<td>2.001</td>
<td>1.001</td>
<td>2.399</td>
<td>2.399</td>
<td>2.399</td>
<td>0.552</td>
<td>0.401</td>
<td>0.401</td>
<td>14.478</td>
</tr>
<tr>
<td>5.001</td>
<td>2.001</td>
<td>1.316</td>
<td>1.316</td>
<td>1.316</td>
<td>0.832</td>
<td>0.682</td>
<td>0.681</td>
<td>8.406</td>
</tr>
<tr>
<td>6.001</td>
<td>5.001</td>
<td>0.592</td>
<td>0.592</td>
<td>0.592</td>
<td>1.138</td>
<td>2.111</td>
<td>2.182</td>
<td>2.708</td>
</tr>
<tr>
<td>10.001</td>
<td>6.001</td>
<td>0.579</td>
<td>0.579</td>
<td>0.574</td>
<td>1.132</td>
<td>2.806</td>
<td>3.274</td>
<td>1.648</td>
</tr>
<tr>
<td>17.001</td>
<td>10.001</td>
<td>1.111</td>
<td>1.042</td>
<td>1.008</td>
<td>1.059</td>
<td>2.464</td>
<td>5.400</td>
<td>0.566</td>
</tr>
<tr>
<td>18.001</td>
<td>17.001</td>
<td>1.136</td>
<td>1.109</td>
<td>1.064</td>
<td>1.056</td>
<td>2.385</td>
<td>5.267</td>
<td>0.563</td>
</tr>
<tr>
<td>19.001</td>
<td>18.001</td>
<td>1.152</td>
<td>1.175</td>
<td>1.122</td>
<td>1.053</td>
<td>2.314</td>
<td>5.117</td>
<td>0.566</td>
</tr>
<tr>
<td>30.001</td>
<td>19.001</td>
<td>1.133</td>
<td>1.498</td>
<td>1.774</td>
<td>1.033</td>
<td>1.833</td>
<td>3.700</td>
<td>0.906</td>
</tr>
<tr>
<td>55.001</td>
<td>30.001</td>
<td>1.073</td>
<td>1.291</td>
<td>1.655</td>
<td>1.018</td>
<td>1.455</td>
<td>2.473</td>
<td>1.367</td>
</tr>
<tr>
<td>75.001</td>
<td>55.001</td>
<td>1.053</td>
<td>1.213</td>
<td>1.480</td>
<td>1.013</td>
<td>1.333</td>
<td>2.080</td>
<td>1.270</td>
</tr>
<tr>
<td>M</td>
<td>α</td>
<td>1/3</td>
<td>1</td>
<td>3</td>
<td>1/3</td>
<td>1</td>
<td>3</td>
<td>1/3</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>0.5326</td>
<td>0.5540</td>
<td>0.5618</td>
<td>0.5238</td>
<td>0.5541</td>
<td>0.5631</td>
<td>0.5502</td>
</tr>
<tr>
<td>0.50</td>
<td>0.50</td>
<td>0.2220</td>
<td>0.2474</td>
<td>0.2539</td>
<td>0.2381</td>
<td>0.2612</td>
<td>0.2613</td>
<td>0.3144</td>
</tr>
<tr>
<td>0.75</td>
<td>0.75</td>
<td>0.0682</td>
<td>0.0801</td>
<td>0.0762</td>
<td>0.1250</td>
<td>0.1213</td>
<td>0.0948</td>
<td>0.1951</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.0329</td>
<td>0.0361</td>
<td>0.0247</td>
<td>0.0740</td>
<td>0.0692</td>
<td>0.0476</td>
<td>0.1256</td>
</tr>
<tr>
<td>1.25</td>
<td>1.25</td>
<td>0.0186</td>
<td>0.0187</td>
<td>0.0131</td>
<td>0.0460</td>
<td>0.0394</td>
<td>0.0253</td>
<td>0.0827</td>
</tr>
<tr>
<td>1.50</td>
<td>1.50</td>
<td>0.0111</td>
<td>0.0107</td>
<td>0.0059</td>
<td>0.0294</td>
<td>0.0239</td>
<td>0.0121</td>
<td>0.0553</td>
</tr>
<tr>
<td>1.75</td>
<td>1.75</td>
<td>0.0069</td>
<td>0.0061</td>
<td>0.0030</td>
<td>0.0192</td>
<td>0.0144</td>
<td>0.0069</td>
<td>0.0374</td>
</tr>
<tr>
<td>2.00</td>
<td>2.00</td>
<td>0.0044</td>
<td>0.0037</td>
<td>0.0017</td>
<td>0.0127</td>
<td>0.0090</td>
<td>0.0038</td>
<td>0.0256</td>
</tr>
<tr>
<td>2.25</td>
<td>2.25</td>
<td>0.0029</td>
<td>0.0022</td>
<td>0.0009</td>
<td>0.0085</td>
<td>0.0056</td>
<td>0.0020</td>
<td>0.0176</td>
</tr>
<tr>
<td>2.50</td>
<td>2.50</td>
<td>0.0019</td>
<td>0.0014</td>
<td>0.0005</td>
<td>0.0058</td>
<td>0.0036</td>
<td>0.0012</td>
<td>0.0122</td>
</tr>
<tr>
<td>2.75</td>
<td>2.75</td>
<td>0.0012</td>
<td>0.0009</td>
<td>0.0003</td>
<td>0.0039</td>
<td>0.0023</td>
<td>0.0007</td>
<td>0.0085</td>
</tr>
<tr>
<td>3.00</td>
<td>3.00</td>
<td>0.0008</td>
<td>0.0005</td>
<td>0.0001</td>
<td>0.0027</td>
<td>0.0015</td>
<td>0.0004</td>
<td>0.0060</td>
</tr>
</tbody>
</table>
TABLE 2.3 VALUES FOR THE RELATIVE SAVING LOSS OF $\delta_{M, \alpha, k}$ ($k > 1$)

<table>
<thead>
<tr>
<th>$b = .50$</th>
<th>$k=3$</th>
<th>$k=5$</th>
<th>$k=7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha=3$</td>
<td>.3152</td>
<td>.2981</td>
<td>.2915</td>
</tr>
<tr>
<td>$M=1.00$</td>
<td>.2120</td>
<td>.1884</td>
<td>.1824</td>
</tr>
<tr>
<td>$M=1.25$</td>
<td>.1447</td>
<td>.1214</td>
<td>.1124</td>
</tr>
<tr>
<td>$M=1.50$</td>
<td>.0951</td>
<td>.0809</td>
<td>.0747</td>
</tr>
<tr>
<td>$M=1.75$</td>
<td>.0632</td>
<td>.0507</td>
<td>.0471</td>
</tr>
<tr>
<td>$M=2.00$</td>
<td>.0428</td>
<td>.0308</td>
<td>.0279</td>
</tr>
<tr>
<td>$M=2.25$</td>
<td>.0279</td>
<td>.0202</td>
<td>.0170</td>
</tr>
<tr>
<td>$M=2.50$</td>
<td>.0183</td>
<td>.0129</td>
<td>.0110</td>
</tr>
<tr>
<td>$M=2.75$</td>
<td>.0124</td>
<td>.0077</td>
<td>.0066</td>
</tr>
<tr>
<td>$M=3.00$</td>
<td>.0080</td>
<td>.0047</td>
<td>.0037</td>
</tr>
<tr>
<td>$M=3.25$</td>
<td>.0052</td>
<td>.0031</td>
<td>.0023</td>
</tr>
<tr>
<td>$M=3.50$</td>
<td>.0035</td>
<td>.0019</td>
<td>.0014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$b = 1.5$</th>
<th>$k=3$</th>
<th>$k=5$</th>
<th>$k=7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha=1/3$</td>
<td>.1401</td>
<td>.1303</td>
<td>.1250</td>
</tr>
<tr>
<td>$M=1.75$</td>
<td>.1094</td>
<td>.0978</td>
<td>.0920</td>
</tr>
<tr>
<td>$M=2.00$</td>
<td>.0853</td>
<td>.0734</td>
<td>.0680</td>
</tr>
<tr>
<td>$M=2.25$</td>
<td>.0664</td>
<td>.0550</td>
<td>.0503</td>
</tr>
<tr>
<td>$M=2.50$</td>
<td>.0516</td>
<td>.0408</td>
<td>.0365</td>
</tr>
<tr>
<td>$M=2.75$</td>
<td>.0401</td>
<td>.0305</td>
<td>.0266</td>
</tr>
<tr>
<td>$M=3.00$</td>
<td>.0311</td>
<td>.0225</td>
<td>.0194</td>
</tr>
<tr>
<td>$M=3.25$</td>
<td>.0241</td>
<td>.0167</td>
<td>.0139</td>
</tr>
<tr>
<td>$M=3.50$</td>
<td>.0186</td>
<td>.0123</td>
<td>.0100</td>
</tr>
<tr>
<td>$M=3.75$</td>
<td>.0144</td>
<td>.0090</td>
<td>.0073</td>
</tr>
<tr>
<td>$M=4.00$</td>
<td>.0111</td>
<td>.0067</td>
<td>.0052</td>
</tr>
<tr>
<td>$M=4.25$</td>
<td>.0085</td>
<td>.0049</td>
<td>.0037</td>
</tr>
</tbody>
</table>
III. COMPROMISE BETWEEN EMPIRICAL BAYES AND MAXIMUM LIKELIHOOD ESTIMATORS

A. The Limited Translation Compound Bayes Estimator

Suppose $X_1, \ldots, X_p$ are independent Poisson variables with respective parameters $\theta_1, \ldots, \theta_p$. Assume the total and componentwise loss functions as follows

$$L(\theta, a) = \sum_{i=1}^{p} \frac{(\theta_i - a_i)^2}{\theta_i}$$

$$L(\theta_i, a_i) = \frac{(\theta_i - a_i)^2}{\theta_i}$$

where $a = (a_1, \ldots, a_p)$, and $\theta = (\theta_1, \ldots, \theta_p)$. The prior distributions of the $\theta_i$'s are assumed to be independent Gamma $(u/(1-u), k_i)$, where $k_i > 0$.

The Bayes estimate of $\theta$ is given by

$$\hat{\theta}(x) = (\hat{\theta}_1(x), \ldots, \hat{\theta}_p(x))$$

where

$$\hat{\theta}_i(x) = \begin{cases} \frac{(1-u)(x_i+k_i-1)}{x_i} & \text{if } x_i > 1 \\ 0 & \text{if } x_i = 0 \text{ and } k_i < 1 \end{cases} \quad 1 \leq i \leq p \quad (3.2)$$

If $u$ is unknown, we estimate $u$ from the data $(X_1, \ldots, X_p)$. Under the assumed prior, $X_1, \ldots, X_p$ are marginally independent, and negative binomial with parameters $u$ and $k_i$, with $i=1,2,\ldots,p$. The minimal sufficient statistic for $u$ is $T = \sum_{i=1}^{p} X_i$, which marginally is negative binomial with parameters $u$, and $k = \sum_{i=1}^{p} k_i$.

Here, as in II we will use a modified Bayes estimate instead of (3.2). This is

$$\tilde{\delta}_i(x) = \begin{cases} 0 & \text{if } x_i = 0 \\ \frac{(1-u)(x_i+k_i-1)}{x_i} & \text{if } x_i > 1 \end{cases} \quad 1 \leq i \leq p \quad (3.3)$$

In view of (3.3), and the minimal sufficiency of $T$ for $u$, a modified empirical Bayes estimate of $\theta$ is of the form
Now, we want to compromise between the above empirical Bayes estimate and the maximum likelihood estimate \( \hat{\theta}(x) = x = (x_1, \ldots, x_p) \).

Our compromise will be componentwise and in the same fashion as in the earlier chapter where we compromised between Bayes and maximum likelihood estimates. So, for every component and for every \( M > 0 \), we use the empirical Bayes estimate whenever 
\[
|x_i - (1 - \hat{\alpha}(t))(x_i + k_i - 1)| \leq M,
\]
otherwise use \( x_i - M \) or \( x_i + M \) depending on whether \( x_i - (1 - \hat{\alpha}(t))(x_i + k_i - 1) > M \) or \( < -M \).

Note that the condition 
\[
|x_i - (1 - \hat{\alpha}(t))(x_i + k_i - 1)| \leq M
\]
is equivalent to the condition 
\[
((1 - \hat{\alpha}(t))(k_i - 1) - M)/\hat{\alpha}(t) < x_i < ((M + (1 - \hat{\alpha}(t))(k_i - 1)))/\hat{\alpha}(t),
\]
so if we take \( M > q - 1 \), where \( q = \max k_i \), then 
\[
(1 - \hat{\alpha}(t))(k_i - 1) - M < 0
\]
and hence the compromise estimate for each component which we call componentwise "limited translation compound Bayes estimate", is given by
\[
\delta_{M, i}(x) = \begin{cases} 
0 & \text{if } x_i = 0 \\
\hat{\alpha}(t)(x_i + k_i - 1) & \text{if } 1 \leq x_i \leq c_i(t) \\
x_i - M & \text{if } x_i > c_i(t) 
\end{cases}
\]
where \( c_i(t) = [(M + (1 - \hat{\alpha}(t))(k_i - 1))/\hat{\alpha}(t)] \) with \([a]\) defined as before.

Defining \( \rho_{M, i}(v) = \min\{1, (M + (k_i - 1))/v\} \), the above estimate, from now on denoted by \( \delta_{M, i} \), can be written as
\[
\delta_{M, i}(x) = \begin{cases} 
0 & \text{if } x_i = 0 \\
(1 - \hat{\alpha}(t))\rho_{M, i}(x_i + k_i - 1) & \text{if } x_i \geq 1
\end{cases}
\]
\( (3.4) \)

Note that the subscript \( i \) in the function \( \rho_{M, i}(\cdot) \) determines the component we consider.
Hence, the limited translation compound Bayes estimate for $\hat{\theta}_o$ will be $\hat{\delta}^M_o(x) = (\hat{\delta}^M_{1,1}(x), \ldots, \hat{\delta}^M_{1,p}(x))$ where $\hat{\delta}^M_{1,i}(x)$ is as in (3.4) for $1 \leq i \leq p$.

B. Risk, Bayes Risk, and Relative Saving Loss of the Limited Translation Compound Bayes Estimator

First, we compute the risk of the limited translation compound Bayes estimator $\hat{\delta}^M_o$. Since the total loss function defined in (3.1) is the sum of componentwise loss functions, the risk of $\hat{\delta}^M_o$ is the summation of the component risks, so we compute the risk for a component of $\hat{\delta}^M_o$ and then add them up.

Theorem 3.1 The risk of $\hat{\delta}^M_{1,i}$ the $i$th component of the limited translation compound Bayes rule, denoted by $R(\hat{\delta}_o, \hat{\delta}^M_{1,i})$, is given by

$$R(\hat{\delta}_o, \hat{\delta}^M_{1,i}) = 1 + M^2 \theta_1^{-1} \left[ (1 - \hat{\theta}(T+2))^2 - 1 \right] G_{T, \lambda_1} \left[ c(T+2) - 2 \right]$$

where $G_{T, \lambda}(.)$ is a Binomial cumulative distribution with parameter $t$ and $\lambda$ and $E_{\lambda}(.)$ is expectation over $T$ which is distributed as Poisson ($\lambda$).

Proof.

$$R(\hat{\delta}_o, \hat{\delta}^M_{1,i}) = E_{\hat{\delta}_o} \left[ \hat{\delta}^M_{1,i}(x) - \theta_1^{-1} \right]^2$$

$$= \theta_1^{-1} E_{\hat{\delta}_o} \left[ \left. \left( \hat{\delta}^M_{1,i}(x) - \theta_1^{-1} \right)^2 \right| X_i = 0 \right] + \left( \hat{\theta} - \theta_1^{-1} \right)^2 I_{1 \leq X_i \leq c_i(T)}$$

$$+ \left( X_i - \hat{\theta} - \theta_1^{-1} \right)^2 I_{X_i > c_i(T)}$$

(3.5)
Note that the distribution of $X_i$ given $T = t$ is Binomial with parameters $t$ and $\lambda_i = (\theta_i / \sum_{j=1}^{p} \theta_j) = \theta_i / \lambda$ where $\lambda = \sum_{j=1}^{p} \theta_j$. When $c_i(t) \geq t$ put $c_i(t) = t$, then

$$E(X_i | X_i = c_i(T)) = \sum_{x=0}^{x_i} \binom{x_i}{x} \binom{x_i (t-x_i)}{x} \lambda_i^{x_i} (1-\lambda_i)^{t-x_i}$$

Similarly,

$$E(X_i | X_i = c_i(T)) = \lambda_i t G_{t-1}, \lambda_i (c_i(t)-1)$$

Hence, taking in (3.6) the conditional expectation with respect to $T$, using (3.7) and (3.8), and then the unconditional expectation over $T$ one gets

$$R(\theta, \delta M, i) = \theta_i^{-1} \left[ \sum_{x=0}^{x_i} \binom{x_i}{x} \binom{x_i (t-x_i)}{x} \lambda_i^{x_i} (1-\lambda_i)^{t-x_i} \right]$$

$$= \lambda_i \sum_{x=0}^{x_i} \binom{x_i}{x} \binom{x_i (t-x_i)}{x} \lambda_i^{x_i} (1-\lambda_i)^{t-x_i}$$

$$= \lambda_i \sum_{x=0}^{x_i} \binom{x_i}{x} \binom{x_i (t-x_i)}{x} \lambda_i^{x_i} (1-\lambda_i)^{t-x_i}$$
\[
\left( (1-\hat{\theta}(T))^2 (k_1-1)^2 - M^2 \right) - 2 \hat{\theta}_1 (M + (1-\hat{\theta}(T))(k_1-1)) \left[ G_{T, \lambda_1} \left( c_1(T) \right) \right] \]
\[
= 1 + M^2 \hat{\theta}_1^{-1} (1 - \exp(-\hat{\theta}_1)) - 2 M \exp(-\hat{\theta}_1) + \\
\hat{\theta}_1^{-1} E \left[ 2 X_1 \left( (1-\hat{\theta}(T))^2 (k_1-1) + M \right) I(c_1(T) \leq X_1) \right] \\
\leq 1 + M^2 \hat{\theta}_1^{-1} (1 - \exp(-\hat{\theta}_1)) + \hat{\theta}_1^{-1} E \left[ 2 X_1 (k_1-1) + M \right] \\
= 1 + M^2 \hat{\theta}_1^{-1} (1 - \exp(-\hat{\theta}_1)) + 2 (k_1-1) + 2 M \\
\leq (1+M)^2 + 2(k_1-1) 
\]

(3.9)

So, the risk of any component of the limited translation compound Bayes estimator is bounded and does not depend on \( p \), the number of parameters we are estimating.
The total risk for \( \delta_M^{\rho} \), denoted by \( R(\theta, \delta_M^{\rho}) \), is \( R(\theta, \delta_M^{\rho}) = \sum_{i=1}^{p} R(\theta, \delta_M^{\rho}, i) \leq p(1+M)^2 + 2 \sum_{i=1}^{p} (k_i - 1) \). Also, observe that if \( k_i < 1 \) for all \( i \) then there exist \( M > 0 \) but small enough such that \( p(1+M)^2 + 2 \sum_{i=1}^{p} (k_i - 1) < p \) which tells us limited translation compound Bayes rules dominate the maximum likelihood estimator for the above specific case.

Next, we compute the Bayes risk of \( \delta_M^{\rho} \) with respect to the priors of the \( \theta_i \)'s given in section A. part III. We will denote the Bayes risk of any rule \( \delta \) with respect to the above mentioned priors by \( r(u, k, \delta) \), where \( k = (k_1, \ldots, k_p) \). For computing the Bayes risk of \( \delta_M^{\rho} \), we need the following lemma.

**Lemma 3.1** Under the assumptions in Section A. the following equalities hold, when \( k_i > 0 \).

\[
E(\theta_1 | X_1 = x_1) = (1-u)(x_1 + k_1) \quad \text{and} \quad E(\theta_1^{-1} | X_1 = x_1) = (1-u)^{-1}(x_1 + k_1 - 1)^{-1}
\]

for \( x_1 + k_1 - 1 > 0 \) \hspace{1cm} (3.10)

**Proof.** It is clear that conditional on \( X_1 = x_1 \), \( \theta_1 \) has a Gamma distribution with parameters \( (1-u)^{-1} \) and \( x_1 + k_1 \). Hence,

\[
E(\theta_1 | X_1 = x_1) = \int_0^\infty \theta_1 e^{-(1-u)^{-1} \theta x_1 + k_1 - 1} / \Gamma(x_1 + k_1)(1-u)^{x_1 + k_1} \, d\theta
\]

\[
= \frac{\Gamma(x_1 + k_1 + 1)(1-u)^{x_1 + k_1 + 1}}{\Gamma(x_1 + k_1)(1-u)^{x_1 + k_1}} \quad \forall \, k_i > 0 \, \text{and} \, x_i > 0.
\]

\[
= (x_1 + k_1)(1-u) \quad \forall \, k_i > 0 \, \text{and} \, x_i > 0. \hspace{1cm} (3.11)
\]

Again,
\[ E(\theta_1^{-1}|X_1=x_1) = \frac{1}{\Gamma(x_1+k_1)(1-u)x_1+k_1} \int_0^\infty (1-u)\theta_1 (x_1+k_1-1)\, d\theta_1 \quad (3.12) \]

Now,
\[ \int_0^\infty (1-u)\theta_1 (x_1+k_1-1)\, d\theta_1 = \Gamma(x_1+k_1-1)(1-u)x_1+k_1 -1 \text{ if } x_1+k_1-1 > 0 \]

Hence, from (3.12) and (3.13) we have
\[ E(\theta_1^{-1}|X_1=x_1) = (x_1+k_1-1)^{-1}(1-u)^{-1} \text{ if } x_1+k_1-1 > 0 \]

which completes the proof.

Theorem 3.1 The Bayes risk of the limited translation compound Bayes estimator \( \theta_M \) is given by

\[ r(u,k),\theta_M = (1-u)^{-1} \sum_{i=1}^p \left( (u-\hat{u}(T))\rho_{M,i} \right)^2 x_i(\hat{k}_i-1) + p(1-u) \quad (3.14) \]

where \( \hat{u} \) is expectation over \( T(1) = \sum X_j^i \), \( \hat{u} \) is expectation over \( X_i \), and

\[ g_i(T) = \hat{u}(T)\rho_{M,i} \left( (k_i-1)\hat{u}(T) \right). \]

Proof. Let \( E \) denote the twofold expectation. Then, writing

\[ \rho_{M,i} = \rho_{M,i} \left( (x_1+k_1-1)\hat{u}(T) \right), \]

and using Lemma 3.1 one gets,

\[ r(u,k),\theta_M = \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 \geq 1) \]

\[ = \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 \geq 1) + \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 \geq 1) \]

\[ + \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 = 0) \]

\[ = \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 = 0) \]

\[ + \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 \geq 1) \]

\[ + \sum_{i=1}^p \left( (1-u)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 = 0) \]

\[ (1-u)^{-1} \sum_{i=1}^p \left( (k_i-1)\hat{u}(T)\rho_{M,i} \right)^2 \left( (x_1+k_1-1)\rho_{M,i} \right) I(x_1 = 0) \]
\[ p(1-u) + (1-u)^{-1} \sum_{i=1}^{p} \mathbb{E}(\hat{\alpha}(T)_{i}^{\lambda} | X_{i} = u) \times (X_{i} + k_{i} - 1) + \]
\[ (1-u)^{-1} \sum_{i=1}^{p} (k_{i} - 1) u^{k_{i}^{\lambda}} \times \{(1-u)^{2} - \mathbb{E}(T_{i}^{\lambda}) \} \times \}

Finally, recall that the Bayes estimate is given in (3.2) and hence an empirical Bayes estimate for \( \hat{\alpha}_{i} \), denoted by \( \hat{\delta}_{1}^{\lambda} \), is
\[
\hat{\delta}_{1}^{\lambda}(x) = \{ \begin{array}{ll}
\{1-\hat{\alpha}(t)(x_{i} + k_{i} - 1) \text{ if } x_{i} + k_{i} - 1 > 0 & 1 \leq i \leq p \\
0 & \text{otherwise}
\end{array}
\]

Then, the Relative Saving Loss of any rule \( \hat{\delta} \) is defined as
\[
\text{RSL}(\hat{\delta}) = \frac{r(\hat{\delta}) - r(\hat{\delta}^{\lambda})}{r(\hat{\delta}) - r(\hat{\delta}^{0})}
\]
where \( \hat{\delta}^{0}(x) = x \) is as above the maximum likelihood estimate. The interpretation for the RSL is the same as in part I with the empirical Bayes estimator replacing the Bayes estimator.

We know that \( r(\hat{\delta}^{0}) = p \). Now, we compute the Bayes risk of the empirical Bayes estimator given by,
\[
r(\hat{\delta}^{\lambda}) = \sum_{i=1}^{p} \mathbb{E}[ (\hat{\delta}_{i}^{\lambda} - \hat{\delta}_{i}^{0})^{2} | X]
\]
with \( \mathbb{E} \) a twofold expectation. For \( k_{i} > 1 \), using Lemma 3.1 and the fact that \( \mathbb{E}(X_{i}^{2} / \theta_{i} | X_{i} = 0) = 0 \) we have
\[
= \mathbb{E}_{1}^{\lambda} \times \mathbb{E}(X_{i}^{2} / \theta_{i} | X_{i} = 0) = 0
\]
\[
= \mathbb{E}_{1}^{\lambda} \times [ (1-\hat{\alpha}(T))^{2} (X_{i} + k_{i} - 1)^{2} - 2 \hat{\delta}_{i}^{0} (1-\hat{\alpha}(T)) (X_{i} + k_{i} - 1) + \hat{\delta}_{i}^{0}^{2} | X]
\]
\[
= \mathbb{E}_{1}^{\lambda} \times [ (1-\hat{\alpha}(T))^{2} (X_{i} + k_{i} - 1)/ (1-u) - 2 (1-\hat{\alpha}(T)) (X_{i} + k_{i} - 1) + (1-u)(X_{i} + k_{i} - 1) + (1-u)]
\]
\[
= (1-u)^{-1} \mathbb{E}[ (\hat{\alpha}(T)-u)^{2} (X_{i} - k_{i} - 1)] + (1-u)
\]
Now, for \( k_1 < 1 \),

\[
E\theta_1^{-1}\{\delta_1^{(1)}(x) - \theta_1^{-1}\}^2 = E\theta_1^{-1}\{\theta_1^{-2}I(X_1 = 0) + (1 - \hat{u}(T))(X_1 + k_1 - 1) - \theta_1^{-2}I(X_1 > 1)\}
\]

\[
= E\{2(1 - \hat{u}(T))(X_1 + k_1 - 1)I(X_1 > 1) + (1 - \hat{u}(T))^2(X_1 + k_1 - 1)\theta_1^{-1}I(X_1 > 1)\}
\]

\[
= E[2(1 - \hat{u}(T))(X_1 + k_1 - 1) - (1 - \hat{u}(T))^2(X_1 + k_1 - 1)]I(X_1 = 0)
\]

\[
= (1 - u) + (1 - u)^{-1}E[\hat{u}(T) - u]^2(X_1 + k_1 - 1) +
\]

\[
E[(1 - u) - (1 - u)^{-1}(\hat{u}(T) - u)^2](k_1 - 1)I(X_1 = 0)
\]

\[
= (1 - u) + (1 - u)^{-1}E[\hat{u}(T) - u]^2(X_1 + k_1 - 1) +
\]

\[
u^{k_1}(k_1 - 1)E[(1 - u) - (1 - u)^{-1}(\hat{u}(T_{(1)}) - u)^2]
\]

(3.18)

Hence, using (3.17) and (3.18), the Bayes risk of \( \delta_1^{(1)} \) given in (3.16) turns out to be,

\[
r((u, k), \delta_1^{(1)}) = \frac{p}{k_1} \sum_{i=1}^{k_1} u^{k_1}(k_1 - 1)E[(1 - u) - (1 - u)^{-1}(\hat{u}(T_{(1)}) - u)^2]I(k_1 < 1)
\]

\[
+ (1 - u)^{-1}E[(\hat{u}(T) - u)^2(X_1 + k_1 - 1)] + (1 - u)
\]

\[
= \frac{p}{k_1} \sum_{i=1}^{k_1} u^{k_1}(k_1 - 1)((1 - u) - (1 - u)^{-1}E[\hat{u}(T_{(1)}) - u]^2)I(k_1 < 1) +
\]

\[
(1 - u)^{-1}E^*[\hat{u}(T) - u]^2(T + k - p) + p(1 - u)
\]

where \( E^* \) is expectation over \( T \). Recall \( k = \sum_{i=1}^{k_1} k_1 \).

For simplicity, from now on we work with the case when \( k_1 = k_2 = \ldots = k_p = 1 \), so that \( \sum_{i=1}^{k_1} k_1 = k = p \). Then,

\[
r((u, k), \delta_1^{(1)}) = (1 - u)^{-1}E^*[\hat{u}(T) - u]^2T + p(1 - u)
\]

(3.19)

Hence, plugging (3.14), (3.19), and \( r((u, k), \delta_0) = p \) in (3.15) for the case of \( k_1 = k_2 = \ldots = k_p = 1 \), the relative saving loss for \( \delta_1^{(1)} \) is
\[
\text{RSL}(u,k,\delta) = \frac{\sum_{i=1}^{p} E[(u-\hat{u}(T)\rho_{M,i}(X_{i}\hat{u}(T)))^2X_{i}]}{p(u-\hat{u}(T))^2} - E[(\hat{u}(T)-u)^2]}
\]

Note that \(\rho_{M,i}(v) = \min\{1,M(k-1)/v\}\), when \(k_1 = 1 \neq 1 \leq i \leq p\), does not depend on \(i\), so we will use \(\rho_{M}(v) = \min\{1,M/v\}\) instead of \(\rho_{M,i}(v)\).

Notice that for the case at hand,

\[
P(X_1=x|T=t) = \frac{P(X_1=x,T=t)}{P(T=t)} = \frac{P(X_1=x,T-X_1=t-x)}{P(T=t)}
\]

\[
= \frac{u(1-u)}{t-x} \frac{t^x C_{t}^x}{(p+t-1)^{(t-x)} (1-u)^{t-x}} = \frac{u(t-1)...(t-(x-1))}{x!}
\]

where \(C_t^x = \frac{x(x-1)...(x-t)}{t!}\), \(C_t^0 = 1\), and \(C_t^x = 0\) when \(t < x\), for \(x=1,2,...\).

Hence, the conditional distribution of \(X_1\) given \(T\) does not depend on the coordinate \(i\). With the last remark in mind, it follows that

\[
E[(u-\hat{u}(T)\rho_{M,i}(X_{i}\hat{u}(T)))^2X_{i}^2] = E[(u-\hat{u}(T)\rho_{M,i}(X_{i}\hat{u}(T)))^2X_{i}^2]
\]

Thus,

\[
P \sum_{i=1}^{p} E[(u-\hat{u}(T)\rho_{M,i}(X_{i}\hat{u}(T)))^2X_{i}^2] = pE[(u-\hat{u}(T)\rho_{M,i}(X_{i}\hat{u}(T)))^2X_{i}^2]
\]

Furthermore,

\[
E[(\hat{u}(T)-u)^2] = \left(\sum_{t=0}^{\infty} (\hat{u}(t)-u)^2 \frac{t!(p+t-1)}{t!} u^t (1-u)^t\right)
\]

\[
= \left(\sum_{t=0}^{\infty} p(\hat{u}(t)-u)^2 \frac{(p+t-1)}{t} u^t (1-u)^t\right)
\]

\[
= \left(\sum_{t=0}^{\infty} \frac{p(1-u)}{u} (\hat{u}(t+1)-u)^2 \frac{(p+t)}{t} u^t (1-u)^t\right)
\]

where, \(E^{**}\) is expectation over \(T\) which has a negative binomial
probability function with parameters $p+1$ and $u$ instead of $p$ and $u$.

Putting (3.21) and (3.22) in (3.20), one gets

$$RSL(u, k, \delta_M) = \frac{E[(u-\hat{u}(T)\rho_M(X_1, \hat{u}(T))]^2 X_1 - (1-u) u - E^* (\hat{u}(T+1)-u)^2}{u(1-u)-(1-u) u - E^* (\hat{u}(T+1)-u)^2}$$

(3.23)

Finally, we make some simplifications of

$$E[(u-\hat{u}(T)\rho_M(X_1, \hat{u}(T))]^2 X_1).$$

Remember that in this case $c_i(t) = [M/\hat{u}(t)]$ which does not depend on $i$, for what $c(t)$ is used instead of $c_i(t)$.

If $c(t) \geq t$, then put $c(t) = t$ in what follows.

$$E[(u-\hat{u}(T)\rho_M(X_1, \hat{u}(T))]^2 X_1) = E[E[(u-\hat{u}(T)\rho_M(X_1, \hat{u}(T))]^2 X_1)|T=t]$$

$$= \sum_{t=0}^{\infty} \sum_{x=0}^{t} \left( (t-x)^{t-x} \right) \left( c(t) \right) \left( (t-x)^{y-p-2} \right)$$

(3.24)
\[(t+p-1)p = (p-1)(t-c(t)-1)/p-t)(t-c(t)+p-2) \quad (3.25)\]

Similarly, if \(c(t) < t\) then
\[
t \sum_{x=c(t)+1}^{t} x(t-x+p-2) = \sum_{x=c(t)+1}^{t} (t-(t-x))(t-x+p-2) = \sum_{y=0}^{t-c(t)-1} (t-y)(y+p-2) = \sum_{y=0}^{t-c(t)-1} [(y+p-1) - (y+p-2)] = (t-c(t)+p-2)
\]

Of course, if \(c(t) > t\) then (3.26) is zero. A final equality which we need is as follows:
\[
t \sum_{x=c(t)+1}^{t} (t-x+p-2) = \sum_{y=0}^{t-c(t)-1} (y+p-2) = \sum_{y=0}^{t-c(t)-1} [(y+p-1) - (y+p-2)]
\]

Substituting (3.25), (3.26) and (3.27) in (3.24) we obtain

\[
\frac{1}{\sum_{x=0}^{t} x(t-x+p-2)}(t-c(t)+p-2)
\]

The last expression can be plugged in (3.23), and then evaluate \(RSL\) numerically.
C. Conclusions

In the case \( k_1 = k_2 = \ldots = k = l \), where \( \sum_{i=1}^{p} k_i = p \), we have calculated \( RLS((u_k)_0, (0^M)_{0^M}) \) for three different values of \( p \) (\( p = 8 \), \( p = 5 \), and \( p = 2 \)) and for each \( p \) three different values of \( u \) (\( u = .25 \), \( u = .50 \), and \( u = .75 \)).

In doing so, we use two different estimates for \( 1-u \), namely \( 1-\frac{(k-1)/(t+k-1)}{t+k-1} \) (the minimum variance unbiased estimate) and \( 1-\frac{k/(t+k)}{t+k-1} \) (the maximum likelihood estimate). In figures 3.1-3.9, the graphs with triangles are the ones which use \( 1-\frac{(k-1)/(t+k-1)}{t+k-1} \) in the calculation of \( RLS((u_k)_0, (0^M)_{0^M}) \), while the graphs with squares are the ones which use \( 1-\frac{k/(t+k)}{t+k-1} \) in the relative saving loss calculation. Figures 3.1-3.9 and Tables 3.1-3.2 show the results.

Again, as in the case of the compromise between the maximum likelihood and the Bayes estimator (or modified Bayes estimator) the relative saving loss of the limited translation compound Bayes estimator is strictly decreasing in \( M \) as figures 3.1-3.9 show. This says that for large values of \( M \) the saving in the Bayes risk we sacrifice by using \( (0^M)_{0^M} \) instead of \( (0^1)_{0^1} \) as compared with the MLE \( (0^0)_{0^0} \) is close to zero. Therefore, the overall performance (in terms of Bayes risk) of \( (0^M)_{0^M} \) measured through the relative saving loss is better than the one of the maximum likelihood estimator.

On the other hand, the componentwise risk of \( (0^M)_{0^M} \) if bounded by \( (1+M)^2 + 2(k_1-1) \) as pointed out in equation (3.9). Although we were not able to compute the maximum componentwise risk of \( (0^1)_{0^1} \), the empirical Bayes estimator, in the light of similar results of Efron and Morris (1972) in the normal case, we conjecture that it depends on \( p \) (number of means...
we are estimating) in such a way that the maximum componentwise risk is large for $p$ large. However, the maximum risk for a component of $\hat{\theta}^M$ is bounded by a function of $M$ and $k_1$, and hence if we take $M$ small enough we can expect that $\hat{\theta}^M$ does componentwise better than $\hat{\theta}^1$ through the minimax criterion.

In summary, the proposed limited translation compound Bayes estimator performs overall and componentwise better than the maximum likelihood and empirical Bayes estimators, respectively.

D. Figures and Tables

Each figure in this part contains two graphs; both are relative saving losses against values of $M$ of the limited translation compound Bayes estimator. The plot that uses triangles, as joining points, displays the relative saving loss when we estimate $1-u$ by $1-(k-1)/(T+k-1)$. The other one uses squares and $1-u$ is estimated by $1-k/(T+k)$. There are nine figures corresponding to the combination of three values of $p$ ($p=2$, $p=5$, and $p=8$) and three values of $u$ ($u=.25$, $u=.5$, and $u=.75$). Recall, we assumed for simplicity $k_i=1$ for $i=1,...,p$.

Table 3.1 gives numerical values of the relative saving loss described in Figures 3.1-3.9 and for those plots using $1-(k-1)/(T+k-1)$ as an estimate of $1-u$. Table 3.2 contains similar information as Table 3.1, the only difference is the estimate for $1-u$, which in this case is $1-k/(T+k)$. 
Triangle \( \hat{\Delta}(T) = \frac{(k-1)}{(T+k-1)} \)

Square \( \hat{\Delta}(T) = \frac{k}{(T+k)} \)

\[ k_1 = k_2 = 1 \]

**FIGURE 3.1 RELATIVE SAVING LOSS OF \( \hat{\Delta}_M^{p=2, u=.25} \)**
Triangle: \( \hat{\theta}(T) = \frac{(k-1)}{T+k-1} \)

Square: \( \hat{\theta}(T) = \frac{k}{T+k} \)

\( k_1 = k_2 = 1 \)

**Figure 3.2** Relative Saving Loss of \( \hat{\rho}_M (p=2, u=.50) \)
Figure 3.3: Relative saving loss of $\rho^M_0$ ($p=2, u=.75$)

Triangle: $\hat{u}(T) = (k-1)/(T+k-1)$
Square: $\hat{u}(T) = k/(T+k)$

VALUES OF M

RELATIVE SAVING LOSS

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

0.0 0.5 1.0 1.5 2.0 2.5 3.0

FIGURE 3.3 RELATIVE SAVING LOSS OF $\rho^M_0$ ($p=2, u=.75$)
Triangle $\hat{u}(T) = \frac{(k-1)}{T+k-1}$
Square $\hat{u}(T) = \frac{k}{T+k}$

$\hat{u}_1 = \hat{u}_2 = \ldots = \hat{u}_5$

**Figure 3.4** Relative Saving Loss of $\hat{\delta}_2^M$ ($p=5, u=.25$)
Triangle \( \hat{u}(T) = \frac{k-1}{T+k-1} \)
Square \( \hat{u}(T) = \frac{k}{T+k} \)

\[ k_1 = k_2 = \ldots = k_5 \]

**FIGURE 3.5 RELATIVE SAVING LOSS OF** \( \hat{p}_M \) (p=5, u=.50)
Triangle \( \hat{u}(T) = \frac{(k-1)}{(T+k-1)} \)

Square \( \hat{u}(T) = \frac{k}{(T+k)} \)

\( k_1 = k_2 = \ldots = k_5 \)

FIGURE 3.6 RELATIVE SAVING LOSS OF \( \rho_M^p \) (\( p=5, u=.75 \))
Triangle $\hat{u}(T) = \frac{k-1}{T+k-1}$
Square $\hat{u}(T) = \frac{k}{T+k}$

$k_1 = k_2 = \ldots = k_g$

FIGURE 3.7 RELATIVE SAVING LOSS OF $\hat{\rho}_M$ (p=8, u=.25)
Triangle \( \hat{u}(T) = \frac{k-1}{T+k-1} \)

Square \( \hat{u}(T) = \frac{k}{T+k} \)

\( k_1 = k_2 = \ldots = k_8 \)

FIGURE 3.8 RELATIVE SAVING LOSS OF \( \delta^P_M \) (\( p=8, u=.50 \))
Triangle $\hat{u}(T) = \frac{(k-1)}{(T+k-1)}$

Square $\hat{u}(T) = \frac{k}{(T+k)}$

$k_1 = k_2 = \ldots = k_8$

FIGURE 3.9 RELATIVE SAVING LOSS OF $\hat{\rho}_M$ ($p=8, u=.75$)
<table>
<thead>
<tr>
<th>M</th>
<th>u</th>
<th>p=2</th>
<th>p=5</th>
<th>p=8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.25</td>
<td>.71797</td>
<td>.78527</td>
<td>.79513</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.73150</td>
<td>.79253</td>
<td>.80132</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.74040</td>
<td>.79759</td>
<td>.80405</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.34303</td>
<td>.49370</td>
<td>.51340</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.33218</td>
<td>.47182</td>
<td>.49309</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.32180</td>
<td>.46691</td>
<td>.48394</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.15174</td>
<td>.31862</td>
<td>.33541</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.11896</td>
<td>.28389</td>
<td>.30969</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.06809</td>
<td>.23869</td>
<td>.26027</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.04254</td>
<td>.20532</td>
<td>.23432</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.02212</td>
<td>.17772</td>
<td>.20360</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.00606</td>
<td>.12452</td>
<td>.14236</td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td>.00113</td>
<td>.13155</td>
<td>.16150</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.00004</td>
<td>.10769</td>
<td>.13056</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.07080</td>
<td>.09019</td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.08341</td>
<td>.10300</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.06365</td>
<td>.08415</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.03532</td>
<td>.05115</td>
</tr>
<tr>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.05208</td>
<td>.07208</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.03715</td>
<td>.05448</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.01641</td>
<td>.02799</td>
</tr>
<tr>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.03181</td>
<td>.04152</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.02054</td>
<td>.03486</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.00727</td>
<td>.01598</td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.01430</td>
<td>.02415</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.00756</td>
<td>.01746</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.00166</td>
<td>.00642</td>
</tr>
<tr>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.00412</td>
<td>.00910</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.00142</td>
<td>.00657</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.00012</td>
<td>.00168</td>
</tr>
<tr>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.00083</td>
<td>.00449</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.00014</td>
<td>.00226</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td></td>
<td>.00036</td>
</tr>
<tr>
<td>2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td>.00008</td>
<td>.00173</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>.00008</td>
<td>.00067</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td>.00008</td>
<td>.00006</td>
</tr>
<tr>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 3.2 VALUES FOR RSL \( ((u,k), \delta^p_M) \) USING \( 1-(k/t+k) \) AS ESTIMATE OF (1-u)

<table>
<thead>
<tr>
<th>M</th>
<th>u</th>
<th>( p=2 )</th>
<th>( k_1=k_2=1 )</th>
<th>( p=5 )</th>
<th>( k_1=k_2=\ldots=k_5=1 )</th>
<th>( p=8 )</th>
<th>( k_1=k_2=\ldots=k_8=1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>.25</td>
<td>.76257</td>
<td>.78478</td>
<td>.78806</td>
<td>.78866</td>
<td>.79895</td>
<td>.80284</td>
</tr>
<tr>
<td>0.3</td>
<td>.43434</td>
<td>.45202</td>
<td>.44191</td>
<td>.49973</td>
<td>.48708</td>
<td>.48074</td>
<td>.51373</td>
</tr>
<tr>
<td>0.5</td>
<td>.26304</td>
<td>.25331</td>
<td>.20442</td>
<td>.32850</td>
<td>.30001</td>
<td>.25601</td>
<td>.33914</td>
</tr>
<tr>
<td>0.7</td>
<td>.17166</td>
<td>.15169</td>
<td>.08674</td>
<td>.21672</td>
<td>.19744</td>
<td>.13730</td>
<td>.23632</td>
</tr>
<tr>
<td>0.9</td>
<td>.10155</td>
<td>.07591</td>
<td>.03473</td>
<td>.14333</td>
<td>.12636</td>
<td>.08468</td>
<td>.16350</td>
</tr>
<tr>
<td>1.1</td>
<td>.06050</td>
<td>.03263</td>
<td>.00853</td>
<td>.09514</td>
<td>.07964</td>
<td>.04612</td>
<td>.10500</td>
</tr>
<tr>
<td>1.3</td>
<td>.03034</td>
<td>.00944</td>
<td>.00104</td>
<td>.06312</td>
<td>.05069</td>
<td>.02397</td>
<td>.07408</td>
</tr>
<tr>
<td>1.5</td>
<td>.01006</td>
<td>.00112</td>
<td>.00002</td>
<td>.04195</td>
<td>.03170</td>
<td>.01249</td>
<td>.04350</td>
</tr>
<tr>
<td>1.8</td>
<td>.00008</td>
<td>0</td>
<td>0</td>
<td>.02241</td>
<td>.01466</td>
<td>.00410</td>
<td>.02734</td>
</tr>
<tr>
<td>2.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.00933</td>
<td>.00462</td>
<td>.00066</td>
<td>.01263</td>
</tr>
<tr>
<td>2.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.00349</td>
<td>.00111</td>
<td>.00007</td>
<td>.00573</td>
</tr>
<tr>
<td>3.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.00107</td>
<td>.00017</td>
<td>0</td>
<td>.00255</td>
</tr>
</tbody>
</table>
IV. COMBINED VERSUS SEPARATE EMPIRICAL BAYES ESTIMATORS

A. Separate and Combined Empirical Bayes Estimators

In this part, we consider the problem of estimating \( p = p_1 + p_2 \) parameters \( \theta = (\theta_1, \theta_2), \theta_1 = (\theta_{11}, \theta_{12}, \ldots, \theta_{1p_1}), \theta_2 = (\theta_{21}, \theta_{22}, \ldots, \theta_{2p_2}) \). Corresponding to each \( \theta_{ij} \) there is a Poisson variable \( X_{ij} \), where

\[
X_{ij} | \theta_{ij} \sim \text{Poisson} (\theta_{ij}) \quad \text{independently} \quad j=1,2, \ldots, p_i, i=1,2 \tag{4.1}
\]

The assumption of taking only one observation from each population does not involve any loss of generality since, as said before, if we take \( n_{ij} \) observations from the \( j \)th population corresponding to the \( i \)th group (\( i=1,2 \)), then the minimal sufficient statistic for \( \theta_{ij} \) can be used instead of \( X_{ij} \) and then the parameter \( \lambda_{ij} = n_{ij} \theta_{ij} \) will play the role of \( \theta_{ij} \).

We will use weighted squared error loss functions for \( \theta_1 \) and \( \theta_2 \) separately and also for \( \theta = (\theta_1, \theta_2) \). If \( \delta = (\delta_1, \delta_2) \) is an estimate of \( \theta \), then the loss functions are

\[
L_i(\theta_i, \delta_i) = \frac{1}{P_i} \sum_{j=1}^{p_i} \left( \delta_{ij} - \theta_{ij} \right)^2 / \theta_{ij} \quad i=1,2
\]

\[
L(\theta, \delta) = \frac{1}{P} \sum_{i=1}^{2} \frac{1}{P_i} \sum_{j=1}^{p_i} \left( \delta_{ij} - \theta_{ij} \right)^2 / \theta_{ij} \tag{4.2}
\]

where \( \delta_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{ip_i}) \), \( i=1,2 \).

For any estimation rule \( \delta(x) \), the corresponding risk functions are

\[
R_i(\theta_i, \delta_i) = E_{\theta_i} L_i(\theta_i, \delta_i(x)) \quad i=1,2
\]

\[
R(\theta, \delta) = E_{\theta} L(\theta, \delta(x)) \tag{4.3}
\]
where $\mathbf{X} = (X_1, X_2)$, $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ip_i})$, $i = 1, 2$.

Our estimation rules will be of the following form

\[(\text{sep})\]
\[
\delta_{li} = \left[1 - \frac{k_i - 1}{T_i + k_i - 1}\right](\mathbf{X}_i + \mathbf{k}_i - \mathbf{1}) \quad i = 1, 2.
\]

\[(\text{comb})\]
\[
\hat{\delta}_{li} = \left[1 - \frac{k_i - 1}{T + k_i - 1}\right](\mathbf{X}_i + \mathbf{k}_i - \mathbf{1}) \quad i = 1, 2.
\]

(4.4)

where $T_i = \sum_{j=1}^{p_i} X_{ij}$, $T = T_1 + T_2$, $\mathbf{k}_i = (k_{i1}, k_{i2}, \ldots, k_{ip_i})$,

$k_i = \sum_{j=1}^{p_i} k_{ij}$, $k = k_1 + k_2$, and $\mathbf{1}$ a $p_i$-dimensional vector of ones. We will assume $p_i \geq 2$ when working with $\delta_{li}$ and $p \geq 2$ when we work with $\hat{\delta}_{li}$. The above estimators will be the separate and combined empirical Bayes estimators under the set up given in the subsequent sections.

B. Bayes Rules

Assume that $\theta_\phi = (\theta_1, \theta_2)$ has the prior distribution

$\theta_{ij} \sim \text{Gamma} \left( \frac{u_i}{1-u_i} k_{ij} \right)$ independently

$j = 1, 2, \ldots, p_i, \quad i = 1, 2$  \hspace{1cm} (4.5)

In this set up, we will constrain $k_{ij}$ to be in the interval $[1, \infty)$. For $k_{ij} \in (0, 1)$, there will be a slight difference between the true rules and the ones we will work, at the point $x_{ij} = 0$. That difference introduces more computations but nothing else.

Conditions (4.1) and (4.5) give the conditional distribution of $\theta_\phi$ given $x_\phi = x$, which is

$\theta_{ij} | x_{ij} = x_{ij} \sim \text{Gamma} \left( \left(1-u_i\right)^{-1}, x_{ij} + k_{ij} \right)$ \hspace{1cm} (4.6)
We will assume the \( k_{ij} \)'s are known, that will not be always the case for the \( u_i \)'s. If \( u_1 \) and \( u_2 \) are known, then the Bayes rule under the given loss is

\[
\delta^*_i = (\delta^*_1, \delta^*_2), \quad \text{where} \quad \delta^*_1 = (\delta^*_11, \delta^*_12, \ldots, \delta^*_1p_1)
\]

and

\[
\delta^*_ij = (1-u_i)(x_{ij}+k_{ij}-1) \tag{4.7}
\]

In the empirical Bayes case, where \( u_1 \) and \( u_2 \) are unknown, they are estimated from \( T_1 = \sum_{j=1}^{p_1} x_{1j} \) and \( T_2 = \sum_{j=1}^{p_2} x_{2j} \) respectively. The motivation for using \( T_1 \) and \( T_2 \) appears in Ghosh (1983). The marginal distribution of \( T_i \) under (4.1) and (4.5) is

\[
T_i \sim \text{Negative Binomial} \left( k_i, u_i \right) \text{ independently } i=1, 2 \tag{4.8}
\]

where \( k_i = \sum_{j=1}^{p_i} k_{ij} \). The above means

\[
P(T_i = t) = \binom{k_i+t-1}{t} u_i^t (1-u_i)^t \quad t=0, 1, \ldots
\]

Let us define the Bayes risk for the \( i^{th} \) group and for any rule

\[
\delta^*(K) = (\delta^*_1(K), \delta^*_2(K)) \text{ as follows}
\]

\[
r_i\left((u, k), \delta^*_i\right) = E_{u, k} R_i(\theta, \delta^*_i) \tag{4.9}
\]

where \( u = (u_1, u_2) \), \( k = (k_1, k_2) = (k_{11}, k_{12}, \ldots, k_{1p_1}, k_{21}, k_{22}, \ldots, k_{2p_2}) \) and \( E_{u, k} \) is expectation with respect to the distribution given by (4.5). The minimum Bayes risk is that of \( \delta^*_i \), and this is equal to

\[
r_i\left((u, k), \delta^*_i\right) = 1-u_i \tag{4.10}
\]

The above Bayes risk is compared with \( r_i((u, k), \delta^*_i) = 1 \) for the maximum likelihood estimator \( \delta^*_i = x_{1i} \). Assuming the prior distribution given in (4.5) to be true, our savings is \( 1-(1-u_i)=u_i \) by using \( \delta^*_i \) instead.
of $\delta_1$. As in Efron and Morris (1973a), we define the "relative saving loss" of $\delta_1$ for group $i$ as

$$\text{RSL}_i((u, k), \delta_1) = \frac{r_1((u, k), \delta_1) - r_1((u, k), \delta_1)}{r_1((u, k), \delta_1)}$$

$$= (r_1((u, k), \delta_1) - (1-u_1))/u_1$$ (4.11)

We can take $i=1$ when deriving properties for a specific group, since similar results will hold for $i=2$.

C. Relative Saving Loss of a Group

We start, proving the following theorem.

Theorem 4.1 For any rule $\delta_1 = (1-u_1(T_1, T_2))X_1 + k_1 - 1^{11}$

we have

$$\text{RSL}_1((u, k), \delta_1) = (k_1/p_1)\hat{E}(u_1 - \hat{u}_1(T_1, T_2))^{2}/u_1^2 +$$

$$[((k_1-p_1)/p_1 u_1 (1-u_1))\hat{E}(u_1 - \hat{u}_1(T_1, T_2))^{2}]$$ (4.12)

where $1^{11}$ is a $p_1$-dimensional vector with all the coordinates equal to one, $\hat{E}$ is expectation with respect to $T_1 \sim$ Negative Binomial $(k_1, u_1)$ and $T_2 \sim$ Negative Binomial $(k_2, u_2)$, and $\hat{E}$ is expectation with respect to $T_1 \sim$ Negative Binomial $(k_1+1, u_1)$ and $T_2 \sim$ Negative Binomial $(k_2, u_2)$.

Proof. Let $E$ denote the twofold expectation with respect to the distribution of random variables as well as the prior.

$$r_1((u, k), \delta_1) = E_p^{-1}_{j=1} p_j (1-\hat{u}_1(T_1, T_2))X_{1j} + k_1 - 1^{12}/\theta_{1j}$$

$$= p_1^{-1} E_{j=1} p_j [(1-u_1)(X_{1j} - k_1 - 1) - \theta_{1j}]^2/\theta_{1j}$$

$$+ (u_1 - \hat{u}_1(T_1, T_2))X_{1j} + k_1 - 1^{12}/\theta_{1j}$$
\[
\begin{align*}
&= p_1^{\frac{1}{2}} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ \left[ (1-u_1)(X_{ij}+k_{ij}-1)-\theta_{1j} \right]^2 + \\
&\quad 2\left[ (1-u_1)(X_{ij}+k_{ij}-1)-\theta_{1j} \right] (u_1-\hat{u}_1(T_1, T_2))(X_{ij}+k_{ij}-1) + \\
&\quad (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / \theta_{1j} \\
\end{align*}
\]

Note that,
\[
p_1^{\frac{1}{2}} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ \left[ (1-u_1)(X_{ij}+k_{ij}-1)-\theta_{1j} \right]^2 / \theta_{1j} = r_1((u_1, k), \delta_{1j}) = 1-u_1 \right. \\
\]

and
\[
\begin{align*}
\mathbb{E}[ (1-u_1)(X_{ij}+k_{ij}-1)-\theta_{1j} ] / \theta_{1j} &= \mathbb{E}\left[ \left\{ (1-u_1)(X_{ij}+k_{ij}-1)-\theta_{1j} \right\} / \theta_{1j} | \mathcal{X} \right] \\
&= 0 \\
\end{align*}
\]

Hence,
\[
\begin{align*}
r_1((u_1, k), \delta_{1j}) = (1-u_1) + p_1^{\frac{1}{2}} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / \theta_{1j} \\
\end{align*}
\]

Since,
\[
\begin{align*}
\mathbb{E}[\theta_{1j}^{-1}|\mathcal{X}=\hat{\mathcal{X}}] &= \begin{cases} \\
(1-u_1)^{-1} x_{1j}^{-1} & \text{if } k_{1j} > 1 \\
(1-u_1)^{-1} x_{1j}^{-1} & \text{if } k_{1j} = 1 \text{ and } x_{1j} > 0 \\
\end{cases} \\
\end{align*}
\]

and \(\mathbb{E}[X_{ij}^2/\theta_{1j} | I(X_{ij}=0)] = 0\) one has
\[
\begin{align*}
r_1((u_1, k), \delta_{1j}) = (1-u_1) + p_1^{\frac{1}{2}} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / \theta_{1j} \\
&= (1-u_1) + p_1^{\frac{1}{2}} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / (1-u_1) \\
&= (1-u_1) + (p_1(1-u_1))^{-1} \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / p_1 \\
\end{align*}
\]

where \(\mathbb{E}^*\) is expectation over \(\mathcal{X}\).

Hence,
\[
\begin{align*}
\text{RSL}_{1}((u_1, k), \delta_{1j}) = \mathbb{E}_{\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2} \left\{ (u_1-\hat{u}_1(T_1, T_2))^2 (X_{ij}+k_{ij}-1)^2 \right\} / p_1 u_1 (1-u_1) \\
\end{align*}
\]

Now,
\[ \bar{E}[u_1 - \hat{u}_1(T_1, T_2)]^2(T_1 + k_1 - p_1) = \bar{E}[u_1 - \hat{u}_1(T_1, T_2)]^2 T_1(p_1 u_1(1-u_1))^{-1} + (k_1 - p_1)(p_1 u_1(1-u_1))^{-1} \bar{E}[u_1 - \hat{u}_1(T_1, T_2)]^2 \]

But,

\[ \bar{E}[u_1 - \hat{u}_1(T_1, T_2)]^2 T_1 = \sum_{t_1=0}^{\infty} \sum_{t_2=0}^{\infty} [u_1 - \hat{u}_1(t_1, t_2)]^2 t_1(1+t_1)(1+t_2) \bar{u}_1(t_1, t_2)^2 \]

\[ = k_1(1-u_1)u_1^{-1} \sum_{t_1=0}^{\infty} \sum_{t_2=0}^{\infty} [u_1 - \hat{u}_1(t_1+1, t_2)]^2(1+t_1)u_1(1-u_1)^{-1}(1-u_1)^{-1} \bar{u}_1(t_1, t_2)^2 \]

\[ = k_1(1-u_1)u_1^{-1} \bar{E}[u_1 - \hat{u}_1(T_1+1, T_2)]^2 \]

which completes the proof.

For the specific case, when \( k_1 = p_1, (k_{1j} = 1, j=1,2,\ldots,p_1) \)

we have

\[ \text{RSL}_1((u,k), \delta_1) = \bar{\bar{E}}[u_1 - \hat{u}_1(T_1+1, T_2)]^2 / u_1^2 \quad (4.13) \]

which determines RSL\(_1\) in terms of how well \( \hat{u}_1(T_1+1, T_2) \) estimates \( u_1 \).

Now, we can work the problem of estimating \( u_1 \) with the loss function

\[ (\hat{u}_1 - u_1)^2 / u_1^2, \]

and in computing the risk, we use \( \bar{\bar{E}} \) instead of \( \bar{E} \).

We give now upper bounds for RSL\(_1\) when \( \delta_1 \) is either \( \delta_1^{(sep)} \) or \( \delta_1^{(comb)} \), in the latter case only when \( u_1 = u_2 = u \).

First,

\[ \bar{\bar{E}}[u_1 - \hat{u}_1(T_1+1, T_2)]^2 = \text{Var}(\hat{u}_1(T_1+1, T_2)) + \bar{\bar{E}}[\hat{u}_1(T_1+1, T_2)]^2 \]

\[ = \bar{u}_1^2(T_1+1, T_2) - 2u_1 \bar{\bar{E}}[\hat{u}_1(T_1+1, T_2)] + u_1^2 \]
Now, \( \hat{u}_1(T_1, T_2) = (k_1-1)(T_1+k_1-1)^{-1} \) for \( \nu_1 \). Hence,

\[
\hat{u}_1^2(T_1+1, T_2) = \mathbb{E}(k_1-1)^2(T_1+k_1)^{-2} = \sum_{t_1=0}^{\infty} (k_1-1)^2(t_1+k_1)^{-2} \frac{t_1+k_1}{t_1} \frac{k_1+1}{t_1} \frac{1}{1-u_1^1 t_1} \\
\leq (k_1-1)^2 \sum_{t_1=0}^{\infty} (t_1+k_1)^{-1}(t_1+k_1)^{-1} \frac{t_1+k_1}{t_1} \frac{k_1+1}{t_1} \frac{1}{1-u_1^1} \\
= (k_1-1)^2 \frac{k_1}{t_1} \frac{k_1}{t_1} \frac{1}{1-u_1^1} = (k_1-1)^2 \frac{k_1}{t_1} \frac{1}{1-u_1^1} \quad (4.14)
\]

Next,

\[
\hat{u}_1^2(T_1+1, T_2) = \mathbb{E}(k_1-1)(T_1+k_1)^{-1} = \sum_{t_1=0}^{\infty} (k_1-1)(t_1+k_1)^{-1} \frac{t_1+k_1}{t_1} \frac{k_1+1}{t_1} \frac{1}{1-u_1^1} \\
= (k_1-1)^2 \frac{k_1}{t_1} \frac{1}{1-u_1^1} = (k_1-1)^2 \frac{k_1}{t_1} \frac{1}{1-u_1^1} \quad (4.15)
\]

From (4.14) and (4.15) we get for \( \hat{\delta}_1^{(\text{sep})} \) that

\[
\mathbb{E}[u_1 - \hat{u}_1(T_1+1, T_2)]^2 \leq (k_1-1)k_1 u_1^2 + 2(k_1-1)^2 k_1 u_1^2 + u_1^2 \\
= u_1^2 [1-(k_1-1)k_1^2] = u_1^2 k_1^{-1} \quad (4.16)
\]

When \( \hat{u}_1 \) is the estimator corresponding to \( \hat{\delta}_1^{(\text{sep})} \), similar computations when \( k_1 > 2 \) yield

\[
\mathbb{E}[u_1 - \hat{u}_1(T_1, T_2)]^2 \leq u_1^2 (k_1-2)^{-1} \quad (4.17)
\]

Hence, (4.16) and (4.17) in (4.12) gives

\[
\text{RSL}_1((u, k), \hat{\delta}_1^{(\text{sep})}) \leq p_1^{-1} + (k_1-2)^{-1} u_1/(1-u_1) \quad (4.18)
\]

When \( k_1=2 \), since we assume \( p_1 \geq 2 \) and \( k_1 \geq p_1 \) we must have \( p_1=2 \), and then the second term in (4.18) becomes zero.

Next, consider the case when \( u_1 = u_2 = u \). In this case \( T \sim \text{Negative Binomial} (k, u) \). For the rule \( \hat{\delta}_1^{(\text{comb})} \), the corresponding estimate for \( u = u_1 \) is \( \hat{u}_1(T_1, T_2) = (k-1)(T+k-1)^{-1} \); therefore, an upper bound for
RSL\(_1((u,k),\delta^{(\text{comb})})\) can be given using (4.16) and (4.17) but with \(u\) and \(k\) playing the role of \(u_1\) and \(k_1\). This is

\[
RSL\(_1((u,k),\delta^{(\text{comb})})\) \leq k_1^{-1} p_1^{-1} + (k_1 - p_1) p_1^{-1} (k_2 - 1) (u/(1-u))
\]  

(4.19)

Putting \(u_1 = u\) in (4.18) one can see that the upper bound is bigger than the corresponding one in (4.19). This fact does not say that 

\[
RSL\(_1((u,k),\delta^{(\text{comb})})\) \leq RSL\(_1((u,k),\delta^{(\text{sep})})\)
\]

when \(u_1 = u_2\) but at least gives a clue to think that will be the case. Moreover, it may be the case that the last observation holds when \(u_1\) and \(u_2\) are close enough.

D. Total Bayes Risk When \(p_1=p_2\)

In section C, we studied the relative saving loss for a group. Now, we study the total relative saving loss (RSL) for all \(p=p_1+p_2\) components. The total RSL is defined as in (4.11), i.e.,

\[
RSL((u,k),\delta) = \frac{r((u,k'),\delta_k) - r((u,k),\delta_{k1})}{r((u,k'),\delta_k) - r((u,k),\delta_{k1})}
\]  

(4.20)

We consider the symmetric case where \(p_1=p_2\). The following theorem gives the total RSL as a convex combination of the partial RSLs of the two groups.

Theorem 4.2 For any rule \(\delta=(\delta_1,\delta_2)\) the total RSL is given by

\[
RSL((u,k),\delta) = RSL\(_1((u,k),\delta_{k1}) + vRSL\(_2((u,k),\delta_{k2})\) / (1+v)
\]  

(4.21)

where \(v=u_2/u_1\).

Proof. We have \(r((u,k'),\delta_k)=1\). Further, since \(p_1=p_2\),

\[
r((u,k),\delta_{k1}) = p_{1}^{2} \sum_{i=1}^{2} E_{i}^{*}(\delta_{k1}^{*}(x) - \theta_{1j})^2 / \theta_{ij}
\]  

\[
= p_{1}^{2} \sum_{i=1}^{2} p_{1} F_{i}((u,k),\delta_{k1}^{*})
\]
\[ r[(u, k), \delta_o (\nu)] = \frac{1}{2} u_1 + u_2 = \frac{1}{2} u_1 (1+\nu). \] (4.22)

Then, \( r[(u, k), \delta^o_o (\nu)] = \frac{1}{2} (u_1 + u_2) = \frac{1}{2} u_1 (1+\nu). \) (4.23)

Finally,
\[
\begin{align*}
& r[(u, k), \delta_o (\nu)] = \frac{1}{2} \sum_{i=1}^{2} r_1 [(u, k), \delta_i (\nu)] = \frac{1}{2} \sum_{i=1}^{2} r_1 [(u, k), \delta_i (\nu)] \\
& = \frac{1}{2} \sum_{i=1}^{2} [u_1 \text{RSL}_1 [(u, k), \delta_i (\nu)] + (1-u_1)] \\
& = \frac{1}{2} u_1 \text{RSL}_1 [(u, k), \delta_1 (\nu)] + v \text{RSL}_2 [(u, k), \delta_2 (\nu)] + \frac{1}{2} \sum_{i=1}^{2} (1-u_1) \\
& \text{using (4.11) in the penultimate step.}
\end{align*}
\] (4.24)

Substituting (4.22)-(4.24) in (4.20) one gets (4.21).

If, it happens that \( \text{RSL}_1 [(u, k), \delta_1 (\nu)] \leq \text{RSL}_1 [(u, k), \delta_1 (\nu)] \) when \( u_1 \) and \( u_2 \) are close enough, then the same will happen to the total RSL between \( \delta_1 (\nu) = (\delta_1 (\nu), \delta_2 (\nu)) \) and \( \delta_2 (\nu) = (\delta_1 (\nu), \delta_2 (\nu)) \) because of (4.21).

E. Conclusions

We carried out numerical calculations, but once again, for the case where the \( k_{i,j} = 1, i=1,2, j=1,\ldots,p \), which leads to \( k_1 = p_1 \) and \( k_2 = p_2 \). Then, (4.12) reduces to

\[ \text{RSL}_1 [(u, k), \delta_1 (\nu)] = \frac{\hat{v}(u_1 - \delta_1 (T_1 + T_2))^2}{u_1^2} \]

In this case and under the additional assumption that \( u_1 = u_2 \), we compute \( \text{RSL}_1 [(u, k), \delta_1 (\nu)] \) and \( \text{RSL}_1 [(u, k), \delta_1 (\nu)] \) for a fixed \( k = 8 \) and three different values of \( k_1 = 2, 4, 6 \). Figures 4.1-4.3 and Table 4.1 show the dominance of \( \delta_1 (\nu) \) over \( \delta_1 (\nu) \) under a Bayes criterion given through the relative saving loss of group one. This is what one expects when
the parameters $u_1$ and $u_2$ are close.

We also compute when $k=8$ the relative saving loss of the combined and separate estimators for group one and for several values of $u_1$ and $u_2$. Figures 4.4-4.6 and Tables 4.2-4.4 give the results. The combinations we used for $k_1$ and $k_2$ were $k_1=2$ and $k_2=6$, $k_1=4$ and $k_2=4$, and $k_1=6$ and $k_2=2$. Looking at Tables 4.2-4.4, for each of the three combinations of $k_1$ and $k_2$, we observe that for a fixed value of $u_1$, say $u_{10}$, there is a neighborhood around $(u_{10},u_{10})$ along the line $u_1=u_{10}$ such that $RSL_1\left((u_1,k),\delta_{1,1}^{(\text{comb})}\right) \leq RSL_1\left((u_1,k),\delta_{1,1}^{(\text{sep})}\right)$. This says that if we have some information about the parameters $u_1$ and $u_2$ from the data we may have some preference as to use one of the above two estimators. In other words, the closer $u_1$ and $u_2$, the better will be the use of the combined estimator.

For the total RSL, and at least for the symmetric case $p_1=p_2$, we can make the same considerations as above using our results in (4.21).

Finally, one can think of generalizing the idea of combining possible related estimation problems when we have $m$ groups instead of two. We conjecture that the same basic result will hold, i.e., the combined estimator will do better so long as the unknown parameters $u_1,\ldots,u_m$ are concentrated around one specific point, showing the way in which the problems (apparently independent) are related.

F. Figures and Tables

Figures 4.1-4.3 are relative saving losses of both the separate and the combined estimator for group one under the assumption that
$u_1 = u_2 = u$. For these figures, we fixed the value of $k = k_1 + k_2$ to be equal to 8, and varied only the values for $k_1$ and $k_2$, so that $k_1$ took the values 2, 4, and 6. Table 4.1 gives the numerical results of Figures 4.1-4.3.

In Figures 4.4-4.6, we let $u_1$ and $u_2$ take values in the interval $(0, 1)$ with the purpose of calculating the relative saving loss of the combined estimation for group one. As before, we fixed $k = 8$ and then $k_1$ took the values 2, 4, 6. Each figure represents a value of $k_1$. Tables 4.2, 4.3, and 4.4 give the numerical values of Figures 4.4, 4.5, and 4.6, respectively.

In doing the calculation for the above figures, we used the assumption made before on the $k_{ij}$'s, i.e., $k_{ij} = 1$, for $i = 1, 2$ and $j = 1, \ldots, p_1$. 
Fig. 4.1  RELATIVE SAVING LOSS FOR $\delta_1^{(sep)}$ AND $\delta_1^{(comb)}$ ($k_1=2, k=8$)
Figure 4.2 Relative Saving Loss for $\delta_{1}^{(\text{sep})}$ and $\delta_{1}^{(\text{comb})}$ (\(k_{1}=4, k=8\))
FIGURE 4.3 RELATIVE SAVING LOSS FOR \( R_{11}^{(\text{sep})} \) AND \( R_{11}^{(\text{comb})} \) \((k_1=6, k=8)\)
FIGURE 4.4 RELATIVE SAVING LOSS FOR THE COMBINED ESTIMATOR \((k_1=2, k=8)\)
FIGURE 4.5 RELATIVE SAVING LOSS FOR THE COMBINED ESTIMATOR ($k_1=4, k=8$)
FIGURE 4.6 RELATIVE SAVING LOSS FOR THE COMBINED ESTIMATOR ($k_1=6, k=8$)
TABLE 4.1 VALUES OF $RSL_1((u,k)_{x},\delta_{1}^{(sep)})$ AND $RSL_1((u,k)_{x},\delta_{1}^{(comb)})$ WHEN $u_1 = u_2 = u$ AND $k = 8.$

($k_{ij} = 1 $ for all $i$ all $j$)

<table>
<thead>
<tr>
<th>$u$</th>
<th>$k_1 = 2, k_2 = 6$</th>
<th>$k_1 = 4, k_2 = 4$</th>
<th>$k_1 = 6, k_2 = 2$</th>
<th>$k = k_1 + k_2 = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.34518</td>
<td>0.10041</td>
<td>0.05033</td>
<td>0.04555</td>
</tr>
<tr>
<td>0.2</td>
<td>0.37279</td>
<td>0.18536</td>
<td>0.11105</td>
<td>0.09208</td>
</tr>
<tr>
<td>0.3</td>
<td>0.34572</td>
<td>0.16837</td>
<td>0.11311</td>
<td>0.08556</td>
</tr>
<tr>
<td>0.4</td>
<td>0.32428</td>
<td>0.14857</td>
<td>0.09861</td>
<td>0.07406</td>
</tr>
<tr>
<td>0.5</td>
<td>0.30685</td>
<td>0.13084</td>
<td>0.08489</td>
<td>0.06311</td>
</tr>
<tr>
<td>0.6</td>
<td>0.29220</td>
<td>0.11480</td>
<td>0.07206</td>
<td>0.05272</td>
</tr>
<tr>
<td>0.7</td>
<td>0.27960</td>
<td>0.10014</td>
<td>0.06003</td>
<td>0.04282</td>
</tr>
<tr>
<td>0.8</td>
<td>0.26857</td>
<td>0.08664</td>
<td>0.04868</td>
<td>0.03336</td>
</tr>
<tr>
<td>0.9</td>
<td>0.25878</td>
<td>0.07414</td>
<td>0.03796</td>
<td>0.02431</td>
</tr>
</tbody>
</table>
TABLE 4.2 VALUES FOR RELATIVE SAVING LOSS OF $\delta_{1}^{(\text{comb})}$ ($k_1=2,k_2=6$).

VALUES ABOVE OF $u_1$ ARE $\text{RSL}_{1}(u_k,\delta_{1}^{(\text{sep})})$ AND $k_{1j}=1$ FOR ALL $i$ AND $j$.

<table>
<thead>
<tr>
<th>$u_2$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.04555</td>
<td>0.06620</td>
<td>0.13111</td>
<td>0.18324</td>
<td>0.22291</td>
<td>0.25353</td>
<td>0.27771</td>
<td>0.29722</td>
<td>0.31327</td>
</tr>
<tr>
<td>0.2</td>
<td>0.32667</td>
<td>0.09208</td>
<td>0.15192</td>
<td>0.23861</td>
<td>0.31679</td>
<td>0.38273</td>
<td>0.43784</td>
<td>0.48412</td>
<td>0.52336</td>
</tr>
<tr>
<td>0.3</td>
<td>0.76633</td>
<td>0.14663</td>
<td>0.08556</td>
<td>0.12560</td>
<td>0.18598</td>
<td>0.24667</td>
<td>0.30239</td>
<td>0.35212</td>
<td>0.39613</td>
</tr>
<tr>
<td>0.4</td>
<td>1.24485</td>
<td>0.28950</td>
<td>0.09575</td>
<td>0.07406</td>
<td>0.10331</td>
<td>0.14793</td>
<td>0.19559</td>
<td>0.24177</td>
<td>0.28487</td>
</tr>
<tr>
<td>0.5</td>
<td>1.71487</td>
<td>0.47724</td>
<td>0.15681</td>
<td>0.07005</td>
<td>0.06311</td>
<td>0.08667</td>
<td>0.12185</td>
<td>0.16067</td>
<td>0.19960</td>
</tr>
<tr>
<td>0.6</td>
<td>2.16067</td>
<td>0.68613</td>
<td>0.24874</td>
<td>0.09769</td>
<td>0.05273</td>
<td>0.05272</td>
<td>0.07287</td>
<td>0.10198</td>
<td>0.13463</td>
</tr>
<tr>
<td>0.7</td>
<td>2.57789</td>
<td>0.90361</td>
<td>0.35964</td>
<td>0.14688</td>
<td>0.06377</td>
<td>0.03910</td>
<td>0.04282</td>
<td>0.06075</td>
<td>0.08572</td>
</tr>
<tr>
<td>0.8</td>
<td>2.96644</td>
<td>1.12269</td>
<td>0.48212</td>
<td>0.21098</td>
<td>0.09052</td>
<td>0.04091</td>
<td>0.02747</td>
<td>0.03336</td>
<td>0.04975</td>
</tr>
<tr>
<td>0.9</td>
<td>3.32789</td>
<td>1.33942</td>
<td>0.61139</td>
<td>0.28545</td>
<td>0.12889</td>
<td>0.05458</td>
<td>0.02373</td>
<td>0.01708</td>
<td>0.02431</td>
</tr>
</tbody>
</table>
TABLE 4.3 VALUES FOR RELATIVE SAVING LOSS OF $\delta_{\text{comb}}^{(\text{comb})} (k_1=4, k_2=4)$.

VALUES ABOVE OF $u_1$ ARE $\text{RSL}_{\text{1}} (u, k, \delta)_{\text{1}} (\text{sep})$ AND $k_{ij}=1$ FOR ALL $i$ AND $j$.

<table>
<thead>
<tr>
<th>$u_2$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.04555</td>
<td>0.11641</td>
<td>0.19871</td>
<td>0.26789</td>
<td>0.32435</td>
<td>0.37040</td>
<td>0.40839</td>
<td>0.44017</td>
<td>0.46709</td>
</tr>
<tr>
<td>0.2</td>
<td>0.16510</td>
<td>0.09208</td>
<td>0.13121</td>
<td>0.18793</td>
<td>0.23532</td>
<td>0.29813</td>
<td>0.34537</td>
<td>0.38730</td>
<td>0.42452</td>
</tr>
<tr>
<td>0.3</td>
<td>0.29259</td>
<td>0.11056</td>
<td>0.08556</td>
<td>0.10596</td>
<td>0.14173</td>
<td>0.18165</td>
<td>0.22138</td>
<td>0.25923</td>
<td>0.29460</td>
</tr>
<tr>
<td>0.4</td>
<td>0.40328</td>
<td>0.15706</td>
<td>0.08531</td>
<td>0.07406</td>
<td>0.07406</td>
<td>0.11359</td>
<td>0.14320</td>
<td>0.17408</td>
<td>0.20470</td>
</tr>
<tr>
<td>0.5</td>
<td>0.49567</td>
<td>0.21105</td>
<td>0.10487</td>
<td>0.06779</td>
<td>0.06311</td>
<td>0.07436</td>
<td>0.09361</td>
<td>0.11680</td>
<td>0.14168</td>
</tr>
<tr>
<td>0.6</td>
<td>0.57283</td>
<td>0.26483</td>
<td>0.13307</td>
<td>0.07509</td>
<td>0.05410</td>
<td>0.05272</td>
<td>0.06219</td>
<td>0.07775</td>
<td>0.09668</td>
</tr>
<tr>
<td>0.7</td>
<td>0.63788</td>
<td>0.31572</td>
<td>0.16472</td>
<td>0.08979</td>
<td>0.05491</td>
<td>0.04233</td>
<td>0.04282</td>
<td>0.05116</td>
<td>0.06427</td>
</tr>
<tr>
<td>0.8</td>
<td>0.69334</td>
<td>0.36289</td>
<td>0.19726</td>
<td>0.10852</td>
<td>0.06179</td>
<td>0.03937</td>
<td>0.03172</td>
<td>0.03336</td>
<td>0.04094</td>
</tr>
<tr>
<td>0.9</td>
<td>0.74114</td>
<td>0.40625</td>
<td>0.22940</td>
<td>0.12932</td>
<td>0.07248</td>
<td>0.04140</td>
<td>0.02643</td>
<td>0.02191</td>
<td>0.02431</td>
</tr>
</tbody>
</table>
TABLE 4.4 VALUES FOR RELATIVE SAVING LOSS OF $\delta_{(\text{comb})_{k_1=6,k_2=2}}$.

VALUES ABOVE OF $u_1$ ARE $\text{RSL}_{1}$ $[(u,k)_{(\text{sep})_{\text{m}}}]$ AND $k_{ij}=1$ FOR ALL $i$ AND $j$.

<table>
<thead>
<tr>
<th>$u_2$</th>
<th>$u_1$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>.04555</td>
<td>.11329</td>
<td>.17732</td>
<td>.22689</td>
<td>.27227</td>
<td>.31299</td>
<td>.34933</td>
<td>.38179</td>
<td>.41090</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>.08154</td>
<td>.09208</td>
<td>.10569</td>
<td>.12700</td>
<td>.15273</td>
<td>.17974</td>
<td>.20653</td>
<td>.23242</td>
<td>.25711</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>.10923</td>
<td>.08653</td>
<td>.08556</td>
<td>.08836</td>
<td>.09917</td>
<td>.11414</td>
<td>.13124</td>
<td>.14929</td>
<td>.16765</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>.12850</td>
<td>.09573</td>
<td>.08237</td>
<td>.07406</td>
<td>.07479</td>
<td>.08103</td>
<td>.09063</td>
<td>.10227</td>
<td>.11515</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>.14251</td>
<td>.10554</td>
<td>.08506</td>
<td>.06940</td>
<td>.06311</td>
<td>.06295</td>
<td>.06680</td>
<td>.07335</td>
<td>.08171</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>.15312</td>
<td>.11454</td>
<td>.08969</td>
<td>.06904</td>
<td>.05770</td>
<td>.05272</td>
<td>.05211</td>
<td>.05459</td>
<td>.05926</td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>.16142</td>
<td>.12249</td>
<td>.09486</td>
<td>.07071</td>
<td>.05562</td>
<td>.04695</td>
<td>.04282</td>
<td>.04199</td>
<td>.04362</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>.16809</td>
<td>.12944</td>
<td>.10001</td>
<td>.07333</td>
<td>.05541</td>
<td>.04384</td>
<td>.03688</td>
<td>.03336</td>
<td>.03244</td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>.17357</td>
<td>.13552</td>
<td>.10491</td>
<td>.07638</td>
<td>.05626</td>
<td>.04238</td>
<td>.03312</td>
<td>.02736</td>
<td>.02431</td>
<td></td>
</tr>
</tbody>
</table>
V. REFERENCES

Baranchik, A.J. 1970. A family of estimators of the mean of a


Department of Statistics, Purdue University.

Blyth, C.R. 1951. On minimax statistical decision procedure and

Brown, L.D. 1966. On the admissibility of invariant estimators
of one or more location parameters. Ann. Math. Statist. 37:
1087-1136.


Pages 205-230 in S.S. Gupta and J. Berger, eds. Proc. 3rd Int. Conf.

Clevenson, M.L., and Zidek, J.V. 1975. Simultaneous estimation of
70:698-705.

Dey, D.K. 1981. On the choice of coordinates in simultaneous esti-
mation of normal means under misspecification of normal priors.
Technical Report No. 176. Department of Statistics, University of
Kentucky.

Efron, B., and Morris, C. 1971. Limiting the risk of Bayes and
Statist. Assoc. 66:807-815.

Efron, B., and Morris, C. 1972. Limiting the risk of Bayes and
empirical Bayes estimators-Part 2: the empirical Bayes case.

Efron, B., and Morris, C. 1973a. Combining possibly related esti-
35:379-421.


VI. ACKNOWLEDGMENTS

I am grateful to Professor Malay Ghosh for his suggestions, advice and valuable assistance on this work. Also, I would like to thank him for the inappreciable friendship he offered me during these years.

Special thanks to my wife Olga Lucrecia and my children Ramón and Carlos for their love and companionship. They made this the best time ever for me.

Finally, I thank Mrs. Patricia French for the repeated typings required to bring the manuscript to its final form.