IMPROVING UPON INADMISSIBLE ESTIMATORS
IN DISCRETE EXPONENTIAL FAMILIES

by

Jiunn Tzon Hwang
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-14

August 1979
CHAPTER I
INTRODUCTION

Section 1.1. Stating the Problem

Consider the problem of estimating \( \theta = (\theta_1, \ldots, \theta_p) \) under the loss function \( L_m(\theta, \hat{\theta}) = \sum_{i=1}^{p} \theta_i \hat{\theta}_i (\theta_i - a_i)^2 \), when the observations \( X_i, i=1, \ldots, p \), are independently from discrete exponential families with density \( q_i(\theta_i) t_i(x_i) \theta_i^x \). The usual estimator is typically admissible for one dimension (\( p=1 \)), but is often inadmissible for higher dimensions (\( p > 1 \)) and can hence be improved upon. In this thesis, the problem of improving upon inadmissible estimators is reduced to the study of difference inequalities. Typical difference inequalities are presented and solved. (Special cases had earlier been solved by M.L. Clevenson and J.V. Zidek (1975), J.C. Peng (1975), H.W. Hudson (1978), and K.W. Tsui and S.J. Press (1977)). Also, theorems are obtained which establish the inadmissibility of certain broad classes of estimators.

In Section 1.2, the notation and definitions are discussed. Section 1.3 gives a review of related results obtained by other statisticians. Section 1.4 summarizes the results in this thesis.
Section 1.2. Definitions and Notation

In this section, we briefly discuss the definitions and notation that are used throughout this thesis.

Let \( X_1, \ldots, X_p \) be \( p \) independent random variables, and assume the probability density of \( X_\ell \) with respect to some measure \( \mu_\ell \) is

\[
f_\ell(x_\ell | \theta_\ell), \quad \ell = 1, \ldots, p,
\]

where \( \theta = (\theta_1, \ldots, \theta_p) \) is some unknown parameter. We use the notation

\[
X_\ell \overset{\text{indep.}}{\sim} f_\ell(x_\ell | \theta_\ell), \quad \ell = 1, \ldots, p
\]

(1.2.1)

to indicate this. The measure \( \mu_\ell \) is assumed to be Lebesgue measure when \( X_\ell \) has an absolutely continuous distribution, and is taken to be the counting measure on nonnegative integers when \( X_\ell \) has a discrete distribution. For most of this thesis, it is assumed that the densities are from the discrete exponential family,

\[
f_\ell(x_\ell | \theta_\ell) = \phi_\ell(\theta_\ell) t_\ell(x_\ell) \theta_\ell^{x_\ell}, \quad x_\ell = 0, 1, \ldots
\]

(1.2.2)

where \( \theta_\ell > 0 \), and \( \phi_\ell \) belongs to some subset \( \Phi_\ell \) of \( \mathbb{R} \) (the set of real numbers) \( \ell = 1, \ldots, p \). Note that \( \theta_\ell \) is not the natural parameter of the exponential family. However, in many situations, \( \theta_\ell \) is the interesting parameter to estimate.

Some important special cases of the density in (1.2.2) are the Poisson distribution, the negative binomial distribution and the logarithmic distribution. Denote the Poisson distribution with mean \( \theta \) by \( \text{Po}(\theta) \). Also, let \( \text{NB}(r, \theta) \) denote the negative binomial distribution having the following density

\[
f(x | \theta) = \binom{r+x-1}{r-1} (1-\theta)^r \theta^x, \quad x = 0, 1, \ldots
\]

(1.2.3)
where \( 0 < \theta < 1 \) and \( r \) is a known positive integer.

It is desired to estimate \( \theta = (\theta_1, \ldots, \theta_p) \) on the basis of \( X = (X_1, \ldots, X_p) \). The parameter space is clearly \( \Theta = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_p \subset \mathbb{R}^p \). (\( \mathbb{R}^p \) is \( p \)-dimensional Euclidean space.) Let \( \mathbf{a} = (a_1, \ldots, a_p) \) be an available action (i.e. an estimate of \( \theta \)) and assume that the action space is \( \mathcal{A} \), and \( \mathbb{R}^p \supset \mathcal{A} \supset \overline{\Theta} \). (\( \overline{\Theta} \) is the closure of \( \Theta \).) When action \( \mathbf{a} \) is taken and \( \bar{\theta} \) is the true parameter value, it is assumed that a loss \( L(\bar{\theta}, \mathbf{a}) \) is incurred, where \( L(\bar{\theta}, \mathbf{a}) \) is a real valued function defined on \( \Theta \times \mathcal{A} \). Usually, we assume \( L(\bar{\theta}, \mathbf{a}) \) has the following form

\[
L_m(\bar{\theta}, \mathbf{a}) = \sum_{i=1}^{p} \theta_i^m (\bar{\theta}_i - a_i)^2,
\]

where \( m = (m_1, \ldots, m_p) \) and \( m_1, \ldots, m_p \) are integers. When \( m_i = m \), \( i = 1, \ldots, p \), \( L_m \) is denoted by \( L_m \):

\[
L_m(\Theta, \mathbf{a}) = \sum_{i=1}^{p} \theta_i^m (\bar{\theta}_i - a_i)^2.
\]

A (nonrandomized) estimator \( \hat{\theta}(X) = (\hat{\theta}_1(X), \ldots, \hat{\theta}_p(X)) \) is a function from the sample space to \( \mathcal{A} \), which estimates \( \bar{\theta} \) by \( \hat{\theta}(X) \) when \( X \) is observed. The risk function \( R(\bar{\theta}, \hat{\theta}) \) of an estimator \( \hat{\theta} \) is defined to be

\[
R(\bar{\theta}, \hat{\theta}) = \int L(\bar{\theta}, \hat{\theta}(X)) \prod_{i=1}^{p} f_i(x_i | \bar{\theta}_i) du_i(x_i)
\]

\[
= E_{\bar{\theta}} L(\bar{\theta}, \hat{\theta}(X)),
\]

where, as usual, \( E_{\bar{\theta}} \) denotes expectation. The subscript \( \bar{\theta} \) might be dropped when there is no ambiguity.

An estimator \( \hat{\theta}^* \) is defined to be as good as \( \hat{\theta} \) if

\[
R(\bar{\theta}, \hat{\theta}^*) \leq R(\bar{\theta}, \hat{\theta})
\]

for all \( \bar{\theta} \in \Theta \). The estimator \( \hat{\theta}^* \) is said to be better than \( \hat{\theta} \) (or
dominates $\hat{\delta}$) if, in addition to (1.2.6),

$$R(\hat{\theta}, \hat{\delta}^*) < R(\hat{\theta}, \hat{\delta})$$  \hspace{1cm} (1.2.7)

for some $\theta \in \Theta$. The estimator $\hat{\delta}$ is admissible if there exists no
better estimator, and is inadmissible otherwise.

For any vectors $\mathbf{g}$ and $\mathbf{h}$ and any real number $F$, define

$$\mathbf{g} + \mathbf{h} = (g_1 + h_1, \ldots, g_p + h_p),$$  \hspace{1cm} (1.2.8)

$$\mathbf{gh} = (g_1 h_1, \ldots, g_p h_p),$$  \hspace{1cm} (1.2.9)

and

$$F\mathbf{g} = (Fg_1, \ldots, Fg_p).$$  \hspace{1cm} (1.2.10)

Let $\mathbf{e}_1, \ldots, \mathbf{e}_p$ denote the unit vectors in $\mathbb{R}^p$, i.e.

$$\mathbf{e}_i = (0, \ldots, 0,1,0,\ldots,0), \quad i=1,\ldots,p$$  \hspace{1cm} (1.2.11)

$\rightarrow$ ith component

For any function $F(x)$, denote the ith partial difference of $F(x)$ by

$$\Delta_i F(x) = F(x) - F(x - e_i).$$  \hspace{1cm} (1.2.12)

Also, for any number $g$, define

$$g^+ = \begin{cases} 
    g & \text{if } g \geq 0 \\
    0 & \text{if } g < 0.
\end{cases}$$  \hspace{1cm} (1.2.13)

Section 1.3. History

In the following, some previous results concerning the problem
of improving upon standard estimators will be discussed. It would be
too difficult to list all known results, so we will only mention the
ones that are closely related to the problems considered here.
Subsection 1.3.1. Stein's Result

Let \( X_i \overset{\text{indep.}}{\sim} N(\theta_i, 1) \), \( i = 1, \ldots, p \), i.e. \( X_1, \ldots, X_p \) are independent normal random variables with means \( \theta_1, \ldots, \theta_p \) and variance 1. In Stein (1955), the problem of estimating \( \theta = (\theta_1, \ldots, \theta_p) \) based on \( \bar{X} = (X_1, \ldots, X_p) \) under the loss function \( L_0(\theta, \bar{a}) \) was considered. (Recall from (1.2.5) that

\[
L_0(\theta, \bar{a}) = \sum_{i=1}^{p} (\theta_i - a_i)^2.
\]

Stein proved the surprising result that the usual estimator \( \hat{\theta}^O(X) = \bar{X} \) is inadmissible when \( p \geq 3 \). A better estimator \( \hat{\theta}^* \) was found in James and Stein (1960), which has the form

\[
\hat{\theta}^*(X) = (1 - \frac{p-2}{p}) \bar{X}.
\]

Since then, a considerable amount of work by a number of authors (see the references) has gone into finding significant improvements upon \( \hat{\theta}^O(X) = \bar{X} \) in more general settings. For the normal distribution, the results in the most general setting obtained so far can be found in Berger, et. al. (1976) and Gleser (1979). In their paper, \( \bar{X} \) is assumed to be a multivariate normal vector with unknown mean \( \bar{\theta} \) and unknown covariance matrix \( \Sigma \). The loss function they considered was

\[
L_M(\theta, \bar{a}) = (a - \bar{\theta}) M (a - \bar{\theta})^t,
\]

where \( (a - \bar{\theta})^t \) denotes the transpose of \( (a - \bar{\theta}) \) and \( M \) is a known pxp positive definite matrix. If an estimator \( W \) of \( \bar{\theta} \) is available and \( W \) has a Wishart distribution with parameter \( \frac{1}{2} \), new estimators were obtained which dominate the usual estimator \( \hat{\theta}^O(X) = \bar{X} \) for \( p \geq 3 \). Brown (1966) has also shown for a wide class of loss functions and densities that the estimator \( \hat{\theta}^O(X) = \bar{X} \) can be
improved when \( p \geq 3 \). All the estimators mentioned above correct the usual estimator by shrinking toward the origin.

In Stein (1973), an identity (proven by integration by parts) was developed which has proven to be a powerful tool in the problem of improving upon the standard estimators. In searching for an estimator, \( \delta^*(\bar{x}) \), better than \( \delta^0(\bar{x}) \), Stein wrote \( \delta^*(\bar{x}) \) as \( \delta^0(\bar{x}) + \phi(\bar{x}) \) and used the identity to obtain the representation

\[
R(\theta, \delta^*) - R(\theta, \delta^0) = E_\theta [\mathcal{E}(\phi(\bar{x}))],
\]

where \( \mathcal{E}(\phi(\bar{x})) \) is an expression that does not involve \( \theta \). \( \mathcal{E}(\phi(\bar{x})) \) involves partial derivatives of \( \phi_i(x) \), \( i=1,\ldots,p \). (For the discrete case, \( \mathcal{E}(\phi(\bar{x})) \) will involve partial differences of \( \phi_i(x) \)). The idea of Stein was then to find \( \phi(x) \) so that \( \mathcal{E}(\phi(x)) < 0 \). If a solution exists, then for such \( \phi(x) \) and \( \delta^*(x) \),

\[
R(\theta, \delta^*) - R(\theta, \delta^0) = E_\theta [\mathcal{E}(\phi(x))] < 0,
\]

and it follows that \( \delta^* \) is better than \( \delta^0 \). The original example of Stein's illustrates this idea.

**Example 1.1.** Let \( \bar{x} = (x_1,\ldots,x_p) \), \( \theta = (\theta_1,\ldots,\theta_p) \), and

\[
x_i \overset{\text{indep.}}{\sim} N(\theta_i,1), \quad i=1,\ldots,p.
\]

Under the loss function \( L_0 \), and estimator better than the maximum likelihood estimator \( \delta^0(\bar{x}) = \bar{x} \) can be obtained by the following procedures when \( p \geq 3 \).

(i) Write the new estimator \( \delta^* \) as \( \delta^*(\bar{x}) = \delta^0(\bar{x}) + \phi(\bar{x}) \).

Now,
\[ L(\hat{\theta}, \hat{\delta}^*(X)) - L(\hat{\theta}, \hat{\delta}^0(X)) = \sum_{i=1}^{p} \{ (\theta_i - \hat{\delta}^*_i(X))^2 - (\theta_i - \hat{\delta}^0_i(X))^2 \} \]
\[ = \sum_{i=1}^{p} \{ 2\phi_i(X)(\hat{\delta}^0_i(X) - \theta_i) + \phi_i^2(X) \} \]
\[ = \sum_{i=1}^{p} \{ 2\phi_i(X)(X_i - \theta_i) + \phi_i^2(X) \} \]

It follows that
\[ R(\hat{\theta}, \hat{\delta}^*) - R(\hat{\theta}, \hat{\delta}^0) = E_{\bar{\theta}} \sum_{i=1}^{p} \{ 2\phi_i(X)(X_i - \theta_i) - \phi_i^2(X_i) \}. \tag{1.3.1} \]

(ii) An identity derived by Stein (1973) shows that, if \( \phi_1, \ldots, \phi_p \) satisfy some regularity conditions, then
\[ E_{\bar{\theta}}[(X_i - \theta_i)\phi_i(X)] = E_{\bar{\theta}}[\frac{\partial}{\partial X_i}\phi_i(X)]. \tag{1.3.2} \]

Define
\[ \mathcal{E}(\hat{\delta}(X)) = \sum_{i=1}^{p} \{ 2\frac{\partial}{\partial X_i}\phi_i(X) + \phi_i^2(X) \}. \tag{1.3.3} \]

Under the regularity conditions, it follows from (1.3.1), (1.3.2), and (1.3.3) that
\[ E_{\bar{\theta}}[\mathcal{E}(\hat{\delta}(X))] = R(\hat{\theta}, \hat{\delta}^*) - R(\hat{\theta}, \hat{\delta}^0) \]

(iii) Letting \( \phi_i(X) = -(p-2)X_i^2/\sum_{i=1}^{p} X_i^2 \), it is easy to check that \( \phi_1(X), \ldots, \phi_p(X) \) satisfy the needed regularity conditions when \( p \geq 3 \). A straightforward calculation yields
\[ \mathcal{E}(\hat{\delta}(X)) = \sum_{i=1}^{p} \left[ 2\frac{\partial}{\partial X_i}\phi_i(X) + \phi_i^2(X) \right] \]
\[ = -(p-2)^2/\sum_{i=1}^{p} X_i^2 < 0. \]

Therefore \( \hat{\delta}^* \) is better than \( \hat{\delta}^0 \).
Note in the above, that the main problem was reduced to the study of a differential inequality. The importance of the relation between such differential inequalities and inadmissibility has also been emphasized in Brown (1974). See also Stein (1965), Brown (1971 and 1974) and Berger (1976a, 1976b and 1976c).

Subsection 1.3.2. Exponential Families

Stein's phenomenon has also been observed for many other distributions. In the following, we will briefly describe some of these other cases.

(i) The Poisson distribution with the loss function

\[ L_{\lambda}(\theta, \alpha) = \sum_{i=1}^{p} \theta_i^{-1}(\theta_i - \alpha_i)^2. \]

Let \( X = (X_1, \ldots, X_p) \) and \( X_i \sim \text{indep. Po}(\theta_i), \ i=1, \ldots, p. \)

Under the loss function \( L_{\lambda}, \) Clevenson and Zidek (1975) obtained an estimator which dominates the usual estimator \( \hat{\theta}^0(X) = \bar{X}, \) when \( p \geq 2. \)

The new estimator \( \hat{\theta}^C \) has the form

\[ \hat{\theta}^C(X) = (1 - \frac{c(\sum_{j=1}^{p} X_j)}{p-1 + \sum_{j=1}^{p} X_j}) \bar{X}, \]

where \( c(\cdot) \) is any nondecreasing function such that \( 0 \leq c(\cdot) \leq 2(p-1). \)

Using the property that, conditioning on \( \sum_{j=1}^{p} X_j, \) \( X \) has a multinomial distribution, Clevenson and Zidek obtained an expression for the difference of the risk of \( \hat{\theta}^C \) and that of \( \hat{\theta}^0. \) This expression involves only \( \sum_{j=1}^{p} X_j \) and \( \sum_{j=1}^{p} \theta_j \) and was shown to be negative valued.

Thus \( \hat{\theta}^C \) is better than \( \hat{\theta}^0. \)
(ii) The Poisson distribution with the loss function

\[ L_0(\theta, a) = \sum_{i=1}^{p} (\theta_i - a_i)^2. \]

Peng (1975) considered the same problem as in (i), except under the different loss function \( L_0 \). Essentially, Peng tackled this problem, of improving upon the usual estimator \( \hat{\theta}^0(X) = \bar{X} \), by Stein's technique. An identity derived by Stein for the Poisson distribution says that, for any function \( g \) defined on \( \mathbb{R}^p \),

\[ \theta_1 E_{\theta} (g(X)) = E_{\theta} [X_1 g(X - e_1)], \tag{1.3.5} \]

where, recalling (1.2.11), \( e_i \) represents the unit vector pointing along the \( i \)th direction in \( \mathbb{R}^p \). By writing the competitor to \( \hat{\theta}^0(X) \) as \( \hat{\theta}^0(X) + \phi(X) \) and using the identity (1.3.5), the expression

\[ \& \phi(X) = \sum_{i=1}^{p} \left( 2X_i \Delta_i \phi_i(X) + \phi_i^2(X) \right) \tag{1.3.6} \]

was obtained where \( E_{\theta} [\& \phi(X)] = R(\theta, \phi^0 + \phi) - R(\theta, \phi^0) \). Recall from (1.2.12) that \( \Delta_i \phi_i(X) \) denotes \( \phi_i(X) - \phi_i(X - e_i) \). Therefore, the problem is reduced to the search for \( \phi_1, \ldots, \phi_p \) for which \( \& \phi(X) \leq 0 \). By using the interchangeability of the indices, (1.3.6) can be rewritten in another form which depends on \( X \) only through \( N_i \), the number of indices \( i \) so that \( X_i = 1 \). After a considerable effort, Peng found a solution to \( \& \phi(X) \leq 0 \) when \( p \geq 3 \) with \( E_{\theta} \& \phi(X) < 0 \) for all \( \theta \). Hence, a better estimator was obtained.

To describe Peng's estimator, recall that \( g^+ \) denotes \( \max\{g, 0\} \). Also denote the number of indices \( i \) such that \( X_i > 0 \) by \( \#(X) \). Peng's estimator, \( \hat{\theta}^p \), has the form
\[ \delta_i^p(x_i) = x_i - \frac{(\#(x_i) - 2)^+ h(x_i)}{\sum_{j=1}^p h^2(x_j)}, \quad i = 1, \ldots, p, \quad (1.3.7) \]

where \( h(x_i) \) is defined to be

\[
h(x_i) = \begin{cases} \frac{1}{k} & x_i = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}
\]

(iii) The Poisson distribution with the loss function

\[
L_m = \sum_{i=1}^p \rho_i^m (\theta_i - a_i)^2, \quad m = -2, -3, \ldots.
\]

Tsui and Press (1977) obtained estimators which dominate the usual estimator \( \hat{\theta}^0 \) for \( p \geq 2 \), under the loss function \( L_m, m \leq -2 \). Essentially, they followed Stein's technique. As in (i) and (ii), the new estimators correct \( \hat{\theta}^0 \) by shrinking toward the origin (i.e., each component of the correction term \( \hat{\theta}(X) \) is nonpositive).

Hudson (1978) generalized the technique of Stein, and applied it to improve upon uniform minimum variance unbiased estimators under the loss function \( L_0 \), when the observations are independently from the exponential (discrete and continuous) family. Here we will mainly discuss the discrete case. Suppose

\[
x_i \overset{\text{i.i.d.}}{\sim} f(x_i | \theta_i) = \varphi(\theta_i) t(x_i) \theta_i^{x_i}, \quad i = 1, \ldots, p.
\]

Denote the uniformly minimum variance unbiased estimator of \( \theta_i \) by \( a(X_i) \). (i.e., \( a(X_i) = t(X_i - 1) / t(X_i) \).) Hudson then established an identity,

\[
\theta_i E_\theta (g(X)) = E_\theta [a(X_i) g(X - e_i)],
\]

which is the generalized form of (1.3.5). By use of this identity,
the problem of improving upon the estimator \((a(X_1), \ldots, a(X_p))\) of \((\theta_1, \ldots, \theta_p)\) was reduced to the study of the following difference inequality:

\[
\sum_{i=1}^{p} \left[ 2a(x_i) \Delta_i \phi_i(x) + \phi_i^2(x) \right] \leq 0
\]  

(1.3.8)

Following an argument similar to Peng's (1978) for solving (1.3.6), a function \(\psi(x) = (\phi_1(x), \ldots, \phi_p(x))\) was shown to be a solution to (1.3.8) under the assumptions that \(a(\cdot)\) is an increasing function and that the dimension \(p\) is big enough. Therefore, a better estimator was found under the given assumptions.

A direct application of Hudson's result to the negative binomial case (i.e. the observations \(X_i \overset{\text{indep.}}{\sim} \text{NB}(r, \theta_i), \ i=1, \ldots, p\)) gives an estimator \(\hat{\phi}_i^H\) which dominates the usual one \((\frac{X_i}{r-1+X_i}, \ldots, \frac{X_p}{r-1+X_p})\) under the loss function \(L_0\) for \(p \geq 4\). The new estimator, \(\hat{\phi}_i^H\), has the form (componentwise)

\[
\hat{\phi}_i^H(x) = \frac{X_i}{r-1+X_i} - \frac{\#(x) - 3}{p} \sum_{j=1}^{p} h^2(X_i),
\]  

(1.3.9)

where \(\#(x)\) denotes the number of indices \(j\) for which \(x_j > 0\), and

\[
h(x_i) = \sum_{k=1}^{X_i} (r-1+k)/k, \text{ if } x_i > 0, \text{ and } h(x_i) = 0, \text{ otherwise.}
\]

Subsection 1.3.3. The Gamma Distribution

Berger (1978) obtained solutions to a general differential inequality and applied them to the gamma distribution. To demonstrate the idea, consider the special case that the observations

\[X_i/\theta_i \overset{\text{indep.}}{\sim} \chi^2_{n}, \ i=1, \ldots, p.\]
\( (X^2_n \text{ stands for chi square distribution with } n \text{ degrees of freedom.}) \)

Berger obtained a better estimator than the standard one, 
\[ \hat{\phi}^0(x) = \chi_x/(n+2) \], under the loss function \( L_{m-2} \). (When referring to this problem, \( L_{m-2} \) is used instead of \( L_m \). The reason for this is to facilitate the comparison between the results for the gamma distribution and those for the Poisson distribution that will be developed.)

In search of a better estimator \( \hat{\phi}^B \), Berger wrote \( \hat{\phi}^B \) as (componentwise)
\[ \hat{\phi}_i^B(x) = \hat{\phi}_i^0(x)(1 + \phi_1 (x)), \quad i=1, \ldots, p. \]

By use of an identity derived in Hudson (1978), the problem was reduced to the search for \( \phi_1, \ldots, \phi_p \) satisfying a key differential inequality which involves partial derivatives of \( \phi_i(x) \) of different orders. In general, the differential inequality involves many terms. However, it was shown that higher order differential terms can usually be neglected. This led to the consideration of the following inequality involving only the first order differential terms and the square term of \( \phi_i \):
\[ \xi_m = \sum_{i=1}^{p} \left( x_i^{m+1} \frac{3}{\partial x_i} \phi_1(x) + b_1 x_i^m \phi_2(x) \right) < 0, \quad (1.3.10) \]

where \( b_1, \ldots, b_p \) are positive constants depending on \( m \). Berger obtained solutions to the more general differential inequality,
\[ \xi^B(\phi(x)) = \sum_{i=1}^{p} \{ v_i (x_i) \frac{3}{v_i (x_i)} \phi_1(x) + w_i (x_i) \phi_2(x) \} < 0. \quad (1.3.11) \]

Berger's solutions to (1.3.11) are described below.

Let
\[ g_i(x_i) = \int_{t_i}^{x_i} \frac{1}{v_i(t)} \, dt \quad \text{(indefinite integral)} \]
Suppose it is possible to choose nonnegative constants, \( \beta_1, \ldots, \beta_p, \) \( d_1, \ldots, d_p, \) and \( b \) such that
\[
\frac{w_i(x)g_i^2(x_i)}{b + \sum_{j=1}^p d_j |g_j(x_j)|^{\beta_j}} \leq K < \infty, \quad i=1, \ldots, p. \tag{1.3.12}
\]
then for \( p > \max_{1 \leq j \leq p} \beta_j \) and
\[
0 < c < (p - \max_{1 \leq j \leq p} \beta_j)/pK,
\]
\[
\phi_i(x) = \frac{-cg_i(x_i)}{b + \sum_{j=1}^p d_j |g_j(x_j)|^{\beta_j}}, \quad i=1, \ldots, p \tag{1.3.13}
\]
is a solution to (1.3.11). It follows that a solution to (1.3.10) is
\[
\phi_i^{(m)}(x) = \begin{cases} 
\frac{cx_i^{-m}}{m[b + \sum_{j=1}^p b_j x_j^{-m}]} & \text{when } m \neq 0, \ p \geq 2 \\
\frac{-c \omega_n x_i}{b + \sum_{j=1}^p b_j (\omega_n x_j)^2} & \text{when } m = 0, \ p \geq 3,
\end{cases} \tag{1.3.14}
\]
for some small enough constant \( c > 0. \) Heuristically, \( \phi_0^B \) is thus dominated by \( \phi_i^B \) defined by
\[
\delta_i^B(x) = \frac{x_i}{n + 2} (1 + \phi_i^{(m)}(x)).
\]
This is actually shown to be the case for \( m=0 \) and \( p \geq 3, \) and for \( m=-1, \ 1 \) and \( p \geq 2. \)

Two surprising phenomena exhibited in the above situation are:
(i) The correction terms (i.e. \( \phi_i^{(m)}(\chi)X_i/(n+2), \ i=1,\ldots,p \)) might be positive or negative depending upon the loss function \( L_{m-2} \). For \( m > 0 \), the correction is positive. (\( \hat{o}^0 \) is pulled towards \((\infty,\ldots,\infty)\)! For \( m < 0 \), the correction terms are always negative. (\( \hat{o}^0 \) is pulled towards the origin.) For \( m = 0 \), \( \hat{o}^0 \) is pulled towards a point.

(ii) The dimension needed for inadmissibility of \( \hat{o}^0 \) depends on the loss function. In most cases, \( \hat{o}^0 \) is inadmissible if and only if \( p \geq 2 \).

Brown (1978) also observed Berger's phenomena in the problem of estimating a normal mean under the loss function

\[
L_r(\hat{o},a) = \prod_{i=1}^{p} e^{\theta_i(\hat{o}_i-a_i)^2}.
\]

It was shown that, under \( L_r \), \( \hat{o}^0 \) is inadmissible if and only if \( p \geq 2 \) when \( r \neq 0 \). For \( r = 0 \), this reduces to Stein's case, so that \( \hat{o}^0 \) is inadmissible if and only if \( p \geq 3 \). (The admissibility of \( \hat{o}^0 \) for \( p = 1 \) was established in Hodges and Lehmann (1951), and for \( p \geq 2 \) in Stein (1955).) Brown's improved estimators also pull \( \hat{o}^0 \) towards \((\infty,\ldots,\infty)\) for \( r > 0 \) and towards \((-\infty,\ldots,-\infty)\) for \( r < 0 \). For \( r = 0 \), the better estimator (the James-Stein estimator) pulls \( \hat{o}^0 \) towards a point.

Section 1.4. Summary of Results Obtained in this Thesis

So far, all the discrete cases considered here have been concerned with improving upon unbiased estimators. Since many reasonable estimators are not unbiased, it is also interesting to see if such estimators can be improved upon. To improve upon such estimators under the loss function \( L^m \) (cf. (1.2.4)), we follow the steps of Stein's technique as described in subsection 1.3.1. Instead of
reducing the problem to the study of a difference inequality, an
inequality of a more general type is encountered. (cf. (2.1.8) and
(2.1.9)). This is dealt with in Chapter II by writing a competitor,
$\delta^*$, of $\delta^0$ as
$$
\delta^*(x) = \delta^0(x) + g(x)\phi(x),
$$
where
$$
g(x) = (q_1(x),\ldots,q_p(x)), \quad \phi(x) = (\phi_1(x),\ldots,\phi_p(x)),
$$
and $g(x)\phi(x) = (q_1(x)\phi_1(x),\ldots,q_p(x)\phi_p(x))$. By choosing a suitable
$g(x)$, the inequality of the more general type is then transformed into
a difference inequality of the following form:
$$
\sum_{i=1}^{p} \left( v_i(x)\Delta_i \phi_i^2(x) + w_i(x)\phi_i(x) \right) \leq 0. \tag{1.4.1}
$$
A similar transformation was developed in Berger (1978) for the con-
tinuous case and the squared error loss. Note that for estimators
that are unbiased it will be seen to be sufficient to choose
$q_i(x) = 1, \ i=1,\ldots,p$. For such a situation the new estimator is
then $\delta^0 + \phi$ which corresponds to the earlier work.

For the special case that $v_i(x)$ depends on $x_i$ only, $i=1,\ldots,p$,
a class of nontrivial solutions to (1.4.1) is found in Chapter II.
This special case occurs, for example, when the $i$th component of
$\delta^0(x)$ depends solely on $x_i, \ i=1,\ldots,p$.

In Chapter III, typical applications of the general theorems
developed in Chapter II are given. For each specific application, a
broad class of estimators are given that dominate the standard esti-
mator. These classes of estimators include those obtained by
Cleveorson and Zidek (1975) and Tsui and Press (1977). Also, Peng
(1975) obtained an estimator which is similar to one of our estima-
tors. For the negative binomial distribution, the uniformly minimum
variance unbiased estimator is shown to be dominated by a class of estimators under the loss function $L_m$ for $p \geq 3$. (Recall that Hudson found a better estimator under the loss $L_0$ only for $p \geq 4$.)

Chapter IV contains inadmissibility results for some broad classes of estimators. By choosing appropriate $q(x)$, the theorems in Chapter II can be applied even to cases in which the $i$th component of $\hat{\delta}^0$ depends on the entire $x$. In Section 4.1, a theorem is thus developed which proves the inadmissibility of certain general types of estimators. Further in Section 4.2, another theorem is established, with the aid of $q(x)$ functions, which essentially states that if an estimator $\hat{\delta}^0$ can be improved upon by the theorems in Chapter II, and if the better estimator $\hat{\delta}^*$ pulls $\hat{\delta}^0$ towards the origin, then any other estimator $\hat{\delta}'(x)$, which has $i$th component greater than or equal to that of $\hat{\delta}^0(x)$ for sufficiently large $x$ and all $i=1,...,p$, is inadmissible. (In this sense, $\hat{\delta}^0$ is an "upper bound" for the class of admissible estimators.) It can be concluded that the estimators considered in Section 4.1 are upper bounds of the class of admissible estimators. Some interesting conclusions are that Hudson's estimator $\hat{\delta}^H$ (cf. (1.3.9)) and some of the Peng type estimators and Tsui type estimators are inadmissible. Clevenson's estimator $\hat{\delta}^C$ (cf. (1.3.4)) is also inadmissible if $c(\cdot) \leq \lambda < p-1$, for some constant $\lambda$, which proves a conjecture of Brown (1974) concerning the inadmissibility of some estimators similar to (1.3.4).

In Chapter V, some miscellaneous problems are considered. Section 5.1 exposes the special role played by discreteness in the problem of improving upon standard estimators. A theorem is given which
implies essentially that, under the loss function $L_m$, $m > 0$, it is impossible to improve upon a standard estimator in certain discrete problems by always expanding it. This explains why the first aspect of Berger's phenomena is not observed in the problem of improving upon a standard estimator of Poisson means, while the second aspect of Berger's phenomena is observed. The problem of improving upon the standard estimator of Poisson means under the loss function $L_m$, $m$ a positive integer, is also compared to the related gamma estimation problem. The similarity between the inequalities involved in these two problems, the theorem described above, and Berger's results (1978), seem to suggest that the standard estimator is admissible in this particular case.

In Section 5.2, we consider the question of whether the lack of a solution (except the zero solution) to the key difference inequality encountered, for a particular estimator $\hat{\theta}^0$ implies that $\hat{\theta}^0$ is admissible. An example is given to show that the answer to this question is negative.

In Section 5.3, an example is considered in which it is desired to estimate the unknown parameter $\bar{\theta} = (\vartheta_1, \vartheta_2, \vartheta_3)$ based on three independent observations $X_1, X_2$ and $X_3$, with $X_1 \sim \text{Po}(\vartheta_1)$, $X_2 \sim \text{N}(\vartheta_2,1)$, and $X_3/\vartheta_3 \sim X^2_m$. Under the loss function $L_m$, $m = (0,0,-1)$, an estimator dominating the standard estimator $(X_1, X_2, X_2/(n+2))$ is obtained. The implications of this example are discussed.

Some useful generalizations are given in Section 5.4.
CHAPTER II
THE DIFFERENCE INEQUALITY AND SOLUTIONS

In the first section of this chapter, it is shown how the problem of improving upon an estimator can be reduced to the study of a difference inequality. This reduction follows essentially the steps described in Example 1.1 of Subsection 1.3.1.

In Section 2.2, solutions to a fairly general type of difference inequality are given. In the remainder of this thesis, unless otherwise stated, $x_i$ and $X_i$ will denote an integer and an integer valued random variable, respectively.

Section 2.1. Derivation of the Difference Inequality

Let $X$ be a one dimensional random variable having discrete density

$$f(x|\theta) = \varphi(\theta)t(x)\theta^x, \ x=0,1,...$$  \hspace{1cm} (2.1.1)

For convenience, $t(x)$ is defined to be zero when $x < 0$. The following two lemmas are the keys to obtaining the difference inequality. For the case $m = 1$, these lemmas were proven in Hudson (1975) by using changes of variables. The same method can be used to prove the lemmas for the case where $m$ is a positive integer. When $m$ is a negative integer, a special case of the lemmas for which $X$ has Poisson distribution has been established in Tsui (1977). The proof
below follows essentially Hudson's.

Lemma 2.1. Assume that $X$ has density (2.1.1) with $t(x) > 0$ for $x = 0, 1, \ldots$. For any function $g$, defined on $R$, for which

$$E_\theta(|g(X)|) < \infty,$$

the following are true:

(1) (Hudson 1975)

$$E_\theta[\theta g(X)] = E_\theta[g(X-1) \frac{t(x-1)}{t(x)}].$$

(2)

$$E_\theta[\theta^m g(x)] = E_\theta[g(x-m) \frac{t(x-m)}{t(x)}], \quad (2.1.2)$$

for any nonnegative integer $m$.

(3) Equation (2.1.2) is true for negative integer $m$, if $g(x) = 0$ whenever $x < -m$.

Proof. From (2.1.1) and by change of variables, we have

$$E_\theta[\theta^m g(X)] = \sum_{k=0}^{\infty} \theta^{m+k} g(k) \varphi(\theta) t(k)$$

$$= \sum_{s=m}^{\infty} g(s-m) t(s-m) q(\theta) \theta^s. \quad (2.1.3)$$

If $m > 0$, then $t(s-m) = 0$ for $s < m$. Hence (2.1.3) can be written as

$$E_\theta[\theta^m g(X)] = \sum_{s=0}^{\infty} g(s-m) \frac{t(s-m)}{t(s)} \varphi(\theta) t(s) \theta^s$$

$$= E_\theta[g(x-m) \frac{t(x-m)}{t(x)}].$$

For $m < 0$, the assumption on $g(x)$ in (3) implies $g(s-m) = 0$ when $s < 0$. Thus by dropping the trivial terms of the summation in (2.1.3), it follows that
\[
E_{\varnothing}[\varnothing^m g(X)] = \sum_{s=0}^{\infty} g(s-m) \frac{t(s-m)}{t(s)} \varnothing(s) \varnothing^s
= E_{\varnothing}[g(X-m) \frac{t(X-m)}{t(X)}]. \quad \text{Q.E.D.}
\]

In Lemma 2.2, recall that \(e_i\) denotes the \(i\)th coordinate vector in \(\mathbb{R}^p\) as in (1.2.11).

Lemma 2.2. Let \(X = (X_1, \ldots, X_p)\), where \(X_i \sim \text{indep. } f_i (x_i | \varnothing_i)\) and \(f_i\) is given in (1.2.2) with \(t_i(x_i) > 0\) for \(x_i = 0, 1, \ldots,\) and \(i = 1, \ldots, p\). Then for any real-valued function \(g(X)\), defined on \(\mathbb{R}^p\), for which 
\(E_{\varnothing}(|g(X)|) < \infty\), the following equation is true for any nonnegative integer \(m\):

\[
E_{\varnothing}[\varnothing^m g(X)] = E_{\varnothing}[g(X-m e_i) \frac{t_i(X_i-m)}{t_i(X_i)}]. \quad (2.1.4)
\]

Furthermore, when \(m\) is a negative integer, (2.1.4) is true, if \(g(x_1, \ldots, x_i, \ldots, x_p) = 0\) whenever \(x_i < -m\).

Proof: By Lemma 2.1 and conditioning on \(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_p\), (2.1.4) is easily obtained.

Now consider the loss function \(L_m\), and let \(\varnothing_0 = (\varnothing_1, \ldots, \varnothing_p)\) be an estimator we hope to improve upon. Write a competitor \(\varnothing^*\) of \(\varnothing_0\) as \(\varnothing^* = \varnothing_0 + \varnothing_1 \varnothing_1\), \(\varnothing = (\varnothing_1, \ldots, \varnothing_p)\). Suppose that \(E_{\varnothing}(\varnothing_0(X))^2 < \infty\) and 
\(E_{\varnothing}[\varnothing_0^2(X)] < \infty\) for all \(\varnothing\) and \(i = 1, \ldots, p\), so that \(R(\varnothing, \varnothing_0)\) and \(R(\varnothing, \varnothing^*)\) are both finite, and

\[
R(\varnothing, \varnothing^*) - R(\varnothing, \varnothing_0)(\varnothing_0(X)) = E[L_m(\varnothing, \varnothing^*(X)) - L_m(\varnothing, \varnothing_0(X))], \quad (2.1.5)
\]

where

\[
L_m(\varnothing, \varnothing_0(X)) - L_m(\varnothing, \varnothing_0(X))
= \sum_{i=1}^{p} \varnothing_i \sum_{i=1}^{m_i} ([\varnothing_0(X) + \varnothing_i(X) - \varnothing_i]^2 - [\varnothing_0(X) - \varnothing_i]_0^2)
= \sum_{i=1}^{p} \{2 \varnothing_i(\varnothing_0(X) - \varnothing_i(\varnothing_i + \varnothing_i^2(X)). \quad (2.1.6)
\]
If \( \phi_i(x_1, \ldots, x_i, \ldots, x_p) = 0 \) when \( x_i < -m_i \), then from Lemma 2.2, we have
\[
E_\varnothing \left[ \left( 2\phi_i^m(X) - \phi_i(X) \right) \phi_i(X) + \phi_i^2(X) \right] = E_\varnothing [\xi_i(\phi_i(X))],
\] (2.1.7)
where
\[
\xi_i(\phi_i(x)) = 2\phi_i^m(x-m_i e_i) \phi_i(x-m_i e_i) \frac{t_i(x_i - m_i)}{t_i(x_i)}
- 2\phi_i(x-(m_i+1)e_i) \frac{t_i(x_i - m_i - 1)}{t_i(x_i)}
+ \phi_i^2(x-m_i e_i) \frac{t_i(x_i - m_i)}{t_i(x_i)}
\] (2.1.8)
Define
\[
\xi(\phi(x)) = \frac{1}{p} \sum_{i=1}^{p} \xi_i(\phi_i(x)).
\] (2.1.9)
From (2.1.5), (2.1.6), (2.1.7), (2.1.8) and (2.1.9), we obtain
\[
R(\theta, \delta^*) - R(\theta, \delta^0) = E_\varnothing [\xi(\phi(X))].
\] (2.1.10)
Thus \( \delta^* \) is better than \( \delta^0 \), if for all \( x \), \( \phi(x) \) is a solution to
\[
\xi(\phi(x)) < 0,
\] (2.1.11)
and for some set of \( x \) with positive probability for some \( q \), strict inequality in (2.1.11) actually holds.

It seems difficult to find a solution to (2.1.11). If, however, we write \( \phi(x) \) as \( q(x) \phi(x) \), where \( q_i(x) \geq 0 \) for all \( x \), then with appropriate choice of \( q(x) \), \( \xi_i(\phi_i(x)) \) can be reduced to an expression which involves only a partial difference term and a square term of \( \phi_i \). We thus end up with only a partial difference inequality to deal with. (This method was first introduced for a special differential inequality in Berger (1978).) Theorem 2.1 describes explicitly how this can be done.
Theorem 2.1. Let \( \hat{x} \) be as in Lemma 2.2, and let \( \hat{o}_0 \) be any estimator of \( o \) such that \( R(\theta, \hat{o}_0) < \infty \) for all \( \theta \), under the loss function \( L_m \).

Define, for all \( x \) with \( x_i \geq 0 \),

\[
\phi_i^2(x) = \frac{2t_i(x_i - m_i - 1)}{t_i(x_i)} q_i(x_i - 1) \phi_i(x_i - m_i) \phi_i(x_i - m_i e_i)
+ \frac{t_i(x_i - m_i)}{t_i(x_i)} q_i^2(x_i - m_i e_i) \phi_i^2(x_i - m_i e_i),
\] (2.1.12)

where \( q_i \) and \( \phi_i \) are functions defined on \( \mathbb{Z}^p \) (\( \mathbb{Z} \) is the set of all integers.) Denote \( \frac{1}{p} \sum_{i=1}^{p} \phi_i^2(x) \) by \( \phi^2(x) \). Under the loss function \( L_m \), the estimator \( \hat{o}^* = \hat{o}_0 + \hat{q} \phi \) dominates \( \hat{o}_0 \), if \( q \) and \( \phi \) satisfy the following four conditions:

1. \( E_{\theta}(q_i(x)\phi_i(x))^2 < \infty \), \( i = 1, \ldots, p \);
2. \( x_i < -m_i \) implies that \( \phi_i(x) = 0 \), \( i = 1, \ldots, p \);
3. \( q_i(x) \geq 0 \) and

\[
\{ \delta_i^0(x_i - m_i e_i) t_i(x_i - m_i) q_i(x_i - m_i e_i)
+ t_i(x_i - m_i - 1) q_i(x_i - (m_i + 1) e_i) \phi_i(x_i - m_i e_i) \leq 0; \] (2.1.13)

4. \( \phi^2(x) \leq 0 \), (2.1.14)

for all \( x \), with strict inequality holding on some set of \( x \) of positive probability for some \( \theta \).

Proof: From (2.1.8), we have
\[ \hat{\delta}_i = (2 \frac{t_i(x_i - m_i)}{t_i(x_i)} - \delta^0_i(x - m_i e_i) q_i(x - m_i e_i) \]
\[ - 2 \frac{t_i(x_i - m_i - 1)}{t_i(x_i)} q_i(x - (m_i + 1) e_i) \phi_i(x - m_i e_i) \]
\[ + 2 \frac{t_i(x_i - m_i - 1)}{t_i(x_i)} q_i(x - (m_i + 1) e_i) \Delta_i \phi_i(x - m_i e_i) \]
\[ + q_i^2(x - m_i e_i) \phi_i^2(x - m_i e_i) \frac{t_i(x_i - m_i)}{t_i(x_i)} \].

From condition (iii), we have \( \hat{\delta}_i \leq \hat{\delta}'_i \), \( i = 1, \ldots, p \). It follows that
\[ \xi(\phi) = \sum_{i=1}^{p} \hat{\delta}_i(\phi_i) \leq \sum_{i=1}^{p} \hat{\delta}'_i(\phi_i) = \xi'(\phi). \]

Together with conditions (i) and (ii), (2.1.10) and (2.1.14), this implies that
\[ R(\hat{\theta}, \hat{\delta}^*) - R(\hat{\theta}, \hat{\delta}^0) = E_{\hat{\theta}} \xi(\phi(X)) \]
\[ \leq E_{\hat{\theta}} \xi'(\phi(X)) \]
\[ \leq 0. \quad (2.1.15) \]

Also, \( E_{\hat{\theta}} \xi'(\phi(X)) < 0 \) for some \( \hat{\theta} \), so that \( \hat{\delta}^* \) is better than \( \hat{\delta}^0 \). Q.E.D.

Note 1. Conditions (i) and (ii) are easy to check. To apply the theorem, we will choose \( q_1, \ldots, q_p \), independent of \( \phi_1, \ldots, \phi_p \), so that (2.1.13) is satisfied, and then concentrate on finding solutions to (2.1.14).

Note 2. For the situations considered by Clevenson and Zidek (1977), Peng (1975), Tsui and Press (1977) and Hudson (1978), it is sufficient to choose \( q_i(x) = 1, \ i = 1, \ldots, p \). This is due to the fact that the ith component of the estimator to be improved upon is
\[ \hat{\delta}^0_i(x) = t_i(x_i - 1)/t_i(x_i) \]
which is the uniformly minimum variance unbiased estimator of $\theta_1$. Therefore, the left hand side of (2.1.13) is always zero for
$q_i(x) = 1$, no matter what $m_i$'s are.

Note 3. By iteration, it is possible to determine a $q(x)$ so that
equality in (2.1.13) is actually achieved. For the case that $\delta_i^0(x)$
depends on $x_i$ only, $\delta_i^+$ will have the following form:

$$\delta_i^+(\phi_i(x)) = v_i(x_i)\Delta_i \phi_i(x - m_i e_i) + w_i(x_i)\phi_i^2(x - m_i e_i). \quad (2.1.16)$$

For such $\delta_i^+$, nontrivial solutions to $\delta' \leq 0$ are given under certain
conditions in Section 2.2.

Note 4. In more general cases, $\delta_i^0$ depends on the entire $x$. The
choice of $q_i(x)$ by iteration, as in Note 3, will then give a $\delta_i^+$
similar to (2.1.16) except that the $v_i$ and $w_i$ will now depend on the
entire $x$. Unfortunately, solutions to $\delta' \leq 0$ for such general $\delta_i^+$
are very hard to find. If, however, it is assumed that $\phi_i(x) \leq 0$,
$i = 1, \ldots, p$, we can choose a simpler $q(x)$, independent of $\phi(x)$,
which satisfies inequality (2.1.13) (but not necessarily equality).

The simple $q(x)$ will give a $\delta_i^+$ of the form

$$\delta_i^+(\phi_i(x)) = F(x)v_i(x_i)\Delta_i \phi_i(x) + w_i(x)\phi_i^2(x), \quad (2.1.17)$$

for which nonpositive solutions $\phi_i(x)$, $i = 1, \ldots, p$, to $\delta' \leq 0$ can
be found. In Chapter IV, it will be shown in detail how such $q(x)$
can be chosen.

Note 5. In either of the situations described in Notes 3 and 4, the
following difference inequality is encountered:

$$\sum_{i=1}^{p} [F(x)v_i(x_i)\Delta_i \phi_i(x) + w_i(x)\phi_i^2(x)] \leq 0. \quad (2.1.18)$$
If \( F(x) > 0 \), dividing both sides of (2.1.18) by \( F(x) \), it shows that (2.1.18) can be reduced to

\[
\prod_{i=1}^{p} \left[ v_i(x_i) \Delta_i \phi_i(x) + w_i(x) \phi_i^2(x) \right] \leq 0, \quad (2.1.19)
\]

where \( w_i(x) \) is, of course, different. (When \( F(x) \) can equal zero, a similar reduction can be worked out with a little modification.) The inequality (2.1.19) is similar to (1.3.11), although the former is a partial difference inequality while the latter is a partial differential inequality. Solving (2.1.19) is unfortunately not so easy as solving (1.3.11). The solutions obtained in the next section, however, are similar to Berger's solution (1.3.13) to the inequality (1.3.11).

An easy corollary follows from the proof of Theorem 2.1. This corollary will be needed in section 4.2.

Corollary 2.1.1. Suppose that all the notation and conditions in Theorem 2.1 remain the same with the exception that condition (iv) is now replaced by the following condition (iv)'

(iv)' \( \delta(\phi(x)) \leq 0 \) for all \( x \). \quad (2.1.20)

Then \( \delta^* \) is as good as \( \delta^0 \).

Section 2.2. Solutions to the Difference Inequality

As pointed out in Section 2.1, the key to the problem of improving upon inadmissible estimators is to find a solution to \( \delta'(\phi) \leq 0 \), where

\[
\delta'(\phi) = \prod_{i=1}^{p} \left[ v_i(x_i) \Delta_i \phi_i(x) + w_i(x) \phi_i^2(x) \right] \leq 0. \quad (2.2.1)
\]

In this section, it is assumed that for
\[ i = 1, \ldots, p, \quad v_i(x_i) \geq 0, \quad w_i(x) \geq 0, \]
and that there exist integers \( \alpha_1, \ldots, \alpha_p \) such that
\[ v_i(x_i) > 0 \quad \text{if} \quad x_i \geq \alpha_i. \quad (2.2.2) \]

These conditions will be satisfied by problems normally encountered. (Assuming \( v_i(x_i) \) to be nonnegative does not lose very much generality, since the sign of \( v_i \) can be changed by replacing \( \phi_i \) by \( -\phi_i \) in (2.2.1).)

In the following, \( h_i \) is taken to be
\[
h_i(x_i) = \sum_{k=\alpha_i}^{x_i} \frac{1}{v_i(k)} x_i > \alpha_i
\]
\[ = 0 \quad \text{otherwise}, \quad (2.2.3) \]
and \( \#(x) \) is defined as
\[
\#_{\alpha}(x) = \text{the number of } \{ i : x_i \geq \alpha_i \}, \quad (2.2.4) \]
where \( \alpha = (\alpha_1, \ldots, \alpha_p) \).

We will interpret 0/0 as 0 in the remainder of this paper. The following theorems provide solutions to (2.2.1) under varying conditions.

Theorem 2.2. Let \( d_1(\cdot), \ldots, d_p(\cdot) \) be nondecreasing functions, defined on the set of integers, such that for \( i = 1, \ldots, p \), the following conditions are satisfied:

(i) \( d_i(x_i) > 0 \) if \( x_i \geq \alpha_i \), and \( d_i(x_i) \geq 0 \) for all \( x_i \);
(ii) there exist positive constants \( \beta_1, \ldots, \beta_p \), such that
\[
v_i(x_i)h_i(x_i, -1)\Delta_i d_i(x_i) \leq \beta_i d_i(x_i, -1), \quad (2.2.5)
\]
for all \( x_i \).
(iii) for all \( z = (x_1, \ldots, x_p) \) such that \( \#(z) > \max_{1 \leq j \leq p} \beta_j \), implies

\[
\frac{\sum_{i=1}^{p} w_i(x)h_i^2(x_i)}{\sum_{j=1}^{p} d_j(x_j)} \leq K < \infty
\]

(2.2.6)

for some \( K > 0 \).

Define \( D = \sum_{j=1}^{p} d_j(x_j) \), and assume that \( c(x) \) is a function which is nondecreasing in each coordinate and which satisfies

\[
0 \leq c(x) \leq (\#(x) - \max_{1 \leq j \leq p} \beta_j)^+ / K.
\]

(2.2.7)

If \( p > \max_{1 \leq j \leq p} \beta_j \), the function \( \phi(x) \), defined by

\[
\phi_i(x) = \frac{-c(x)h_i(x)}{D}, \quad i = 1, \ldots, p,
\]

(2.2.8)

is a solution to (2.2.1). Furthermore,

\[
\delta^i(\phi) \leq -c(x)(\#(x) - \max_{1 \leq j \leq p} \beta_j - Kc(x))^+ / D,
\]

(2.2.9)

with strict inequality for those \( x^0 = (x_1^0, \ldots, x_p^0) \) for which

\[
h_i(x_i^0 - 1)\Delta_i d_i(x_i^0) > 0
\]

(2.2.10)

for at least two \( i \)'s and \( c(x^0) \neq 0 \).

Proof. Because of the monotonicity of \( c(x) \) with respect to each coordinate,

\[
\Delta_i \phi_i(x) = \Delta_i \left( \frac{-c(x)h_i(x)}{D} \right)
\]

\[
\leq c(x)\Delta_i \left( \frac{-h_i(x)}{D} \right).
\]

Define \( D_i = d_i(x_i - 1) + \sum_{j \neq i} d_j(x_j) \). Then,
\[ \Delta_{D_i}^{-h_i(x_i)} = \frac{-\Delta h_i(x_i)}{D} + \frac{h_i(x_i-1)D_i}{DD_i}. \]

It follows that

\[ \frac{P}{\sum_{i=1}^P} v_i(x_i) \Delta_{D_i} \phi_i(x) \leq \frac{c(x_i)}{D} \left( \#_{\alpha_i}(x_i) + \frac{P}{\sum_{i=1}^P} \frac{v_i(x_i)h_i(x_i-1)\Delta_i D_i}{D_i} \right) \]

\[ \leq \frac{c(x)}{D} \left( \#_{\alpha}(x) + \frac{P}{\sum_{i=1}^P} \frac{v_i(x_i)h_i(x_i-1)\Delta_i d_i(x_i)}{D_i} \right), \quad (2.2.11) \]

where \( D' \) denotes \( \sum_{j=1}^P d_j(x_j-1) \). In the last transition, the inequality is actually strict for those \( x \) for which \( c(x) \neq 0 \) and two of the \( x_i \)'s satisfy \( h_i(x_i-1)\Delta_i d_i(x_i) > 0 \). It follows from \((2.2.11)\) and \((2.2.5)\) that

\[ \frac{P}{\sum_{i=1}^P} v_i(x_i) \Delta_{D_i} \phi_i(x) \leq \frac{c(x)}{D} \left( \max_{1 \leq j \leq P} \beta_j - \#_{\alpha}(x) \right). \quad (2.2.12) \]

By \((2.2.5)\), we have \( \frac{P}{\sum_{i=1}^P} w_i(x_i) \phi_i^2(x_i) \leq Kc^2(x)/D \), which, together with \((2.2.12)\), implies that

\[ \phi'(x) \leq \frac{c(x)(Kc(x) + \max_{1 \leq j \leq P} \beta_j - \#_{\alpha}(x))}{D}. \]

Since, by \((2.2.7)\),

\[ c(x)(Kc(x) + \max_{1 \leq j \leq P} \beta_j - \#_{\alpha}(x)) \]

\[ = -c(x)(\#_{\alpha}(x) - \max_{1 \leq j \leq P} \beta_j - Kc(x))^+, \]

\((2.2.9)\) is established. \( Q.E.D. \)

Note 1. Theorem 2.2 is true even when \( p = \max_{1 \leq j \leq P} \beta_j \). However, in this case, the \( \phi_i \) in \((2.2.8)\) are always zero, which gives an uninteresting solution to \((2.2.1)\). Therefore \( p > \max_{1 \leq j \leq P} \beta_j \) is assumed in Theorem 2.2.
Note 2. It is easy to choose the \( d_i(\cdot) \), so that they are nondecreasing, nonnegative, and satisfy (2.2.5). Indeed, the following is such a choice:

\[
d_i(x_i) = \prod_{k=\alpha_i+1}^{x_i} \left( 1 + \frac{\beta_i}{v_i(k)h_i(k-1)} \right) \quad x_i \geq \alpha_i + 1
\]

\[
= 1 \quad x_i = \alpha_i
\]

\[
= 0 \quad \text{otherwise}
\]

(2.2.13)

Unfortunately, this choice is too complicated to be useful. Hence, in the following corollaries, efforts will be made to obtain simpler \( d_i(\cdot) \) for special cases.

Note 3. From (2.2.5), it is clear that a larger \( \beta_i \) allows a larger \( d_i \), and hence (2.2.6) is easier to satisfy. But then the dimension, \( p \), required for nontrivial solutions to the inequality (2.2.1), is higher. Also, a larger \( d_i(\cdot) \) gives a larger upper bound in (2.2.9); and the corresponding new estimator will have a smaller improvement in risk. (cf (2.2.9) and (2.1.15)) For these two reasons, we will choose \( \beta_i \) as small as possible.

The following corollaries and examples illustrate the use of Theorem 2.2 in solving the difference inequalities. In each case, \( d_i \) is given explicitly. The first corollary is applicable when the \( v_i \)'s are increasing functions. It also tells more about how the choice of the \( d_i \) can be made.

Corollary 2.2.1 Let \( \lambda', v_i, w_i \) and \( \alpha_i \) denote as in (2.2.1) and (2.2.2). Suppose that \( v_1, ..., v_p \) are increasing functions and that \( \beta_1, ..., \beta_p \) are positive integers. If

\[
d_i(x_i) = b_i h_i(x_i) h_i(x_i+1) ... h_i(x_i+\beta_i-1)+b_0,
\]

(2.2.14)
for \( i = 1, \ldots, p \) and some constants \( b_0 \geq 0 \) and \( b_i > 0 \), then (2.2.5) is satisfied. If, in addition, (2.2.6) is satisfied for this choice of the \( d_i \)'s, then \( \phi_1, \ldots, \phi_p \), as in (2.2.8), is a solution to the inequality (2.2.1), provided \( p > \max_{1 \leq j \leq p} \beta_j \).

Proof: Clearly \( d_i(x_i) \) is increasing and greater than zero for \( x_i \geq \alpha_i \). Therefore, by Theorem 2.2, it is only necessary to show that (2.2.5) is satisfied. Now (2.2.5) is trivial for \( x_i \leq \alpha_i \).

For \( x_i \geq \alpha_i + 1, \beta_i \geq 2, \) we have

\[
\Delta d_i(x_i) = b_i [h_i(x_i+\beta_i-1)-h_i(x_i-1)]h_i(x_i)\ldots h_i(x_i+\beta_i-2) \\
\leq \frac{b_i^{\beta_i}}{v_i(x_i)} h_i(x_i)\ldots h_i(x_i+\beta_i-2).
\]

Consequently, for \( \beta_i \geq 2, \)

\[
v_i(x_i)h_i(x_i-1)\Delta d_i(x_i) \leq \beta_i b_i h_i(x_i-1)h_i(x_i)\ldots h_i(x_i+\beta_i-2) \\
\leq \beta_i d_i(x_i-1).
\] (2.2.15)

For \( \beta_i = 1, \) it is clear that

\[
v_i(x_i)h_i(x_i-1)\Delta d_i(x_i) = b_i h_i(x_i-1) \leq d_i(x_i-1).
\]

Q.E.D.

Note that the \( d_i \) in (2.2.14) are similar to the functions

\( b_i h_i(x_i) + b_0 \). In all the problems discussed in this thesis, the \( d_i \) can be chosen to be (or at least to be similar to) a polynomial function of \( h_i \) so that (2.2.5) is satisfied.

In applying Theorem 2.2, we assume at first that the \( d_i \) have the form \( h_i^{\beta_i} \), and choose \( \beta_i \) as small as possible so that (2.2.6) is satisfied. For such \( \beta_i \), we modify the \( d_i \) (by using the form (2.2.14) or by adding extra positive terms of the form \( h_i^{k_i}, k_i < \beta_i \)) so that
(2.2.5) is satisfied. Solutions to (2.2.1) are then given in (2.2.8).

If, no matter how large \( \beta_i \) is, (2.2.6) is never satisfied, by this theorem, it seems unlikely that solutions to (2.2.1) can be found. Indeed all admissible estimators have a difference inequality for which (2.2.6) is never satisfied for \( d_i(x) = h_i^{\beta_i}(x) \) and any \( \beta_i \).

For the case \( 0 \leq \beta_i \leq 1 \), Theorem 2.2 can be reduced to the following simple corollary.

Corollary 2.2.2 Let \( \Phi_i, \psi_i, w_i \) and \( \alpha_i \) denote the same as in (2.2.1) and (2.2.2). For any constants \( b_0, b_i, \) and \( \beta_i \) such that \( b_0 \geq 0, b_i > 0 \) and \( 0 \leq \beta_i \leq 1, i = 1, \ldots, p \), the function

\[
d_i(x) = b_0 + b_i h_i^{\beta_i}(x)
\]

(2.2.16)

satisfies (2.2.5). If, in addition, (2.2.6) is satisfied for this choice of the \( d_i \), then \( \Phi_i \), with \( \psi_i \) as in (2.2.8), is a solution to the inequality (2.2.1), provided \( p > \max_{1 \leq j \leq p} \beta_j \).

Proof: Again, it is only necessary to show that \( d_i \) satisfies (2.2.5) for \( x \geq \alpha_i + 1 \). Now \( \beta_i \leq 1 \), which, together with mean value theorem implies

\[
v_i(x) h_i(x - 1) \Delta_i d_i(x) = v_i(x) h_i(x - 1) \Delta_i (b_i h_i^{\beta_i}(x))
\]

\[
\leq \beta_i b_i v_i(x) h_i^{\beta_i}(x - 1) \Delta_i h_i(x)
\]

\[
= \beta_i b_i h_i^{\beta_i}(x - 1)
\]

\[
\leq \beta_i d_i(x - 1).
\]

Q.E.D.

Example 2.1. In Chapter III, it will be shown (See (3.1.9), (3.1.10) and (3.1.11)) that the problem that Clevenson and Zidek (1973) considered (refer to Section 1.3.2) can be reduced to the study of the
following inequality:

\[ \sum_{i=1}^{p} \left( \alpha_i \phi_i(x) + \frac{1}{2(x_i+1)} \phi_i^2(x) \right) \leq 0. \]  

(2.2.17)

Using the notation of Theorem 2.2, note that \( v_i(x_i) = 1 \), \( w_i(x) = 1/2(x_i+1) \), and \( \alpha_i = 0 \), \( i = 1, \ldots, p \).

Hence

\[ h_i(x_i) = x_i + 1 \quad x_i \geq 0 \]

\[ = 0 \quad \text{otherwise}. \]

Since, for \( \# \{x\} > 1 \), we have

\[ \sum_{i=1}^{p} \frac{1}{2} \frac{h_i^2(x_i)/(x_i+1)}{\sum_{j=1}^{p} h_j(x_j)} = \frac{1}{2}, \]

(2.2.6) is satisfied with \( \beta_i = 1 \) and \( K = \frac{1}{2} \). It follows from Corollary 2.2.2 that if \( p > 1 \), then

\[ \phi_i(x) = \frac{-c(x)(x_i+1)}{b_0 + \sum_{j=1}^{p} (x_j+1)}, \quad i = 1, \ldots, p, \]

is a solution to (2.2.17) for any nonnegative number \( b_0 \) and any function \( c(x) \) increasing in each coordinate which satisfies

\[ 0 \leq c(x) \leq 2(\# \{x\}-1)^+. \]

Note that \( \# \{x\} \) is equal to \( p \) if it can be assumed that \( x_i \geq 0, \) \( i = 1, \ldots, p \).

It happens quite often especially for the negative binomial distribution that we encounter a difference inequality of the form (2.2.1), with \( 1/v_i(x_i) \leq M_i \) for all \( x_i \geq \alpha_i \). In such a situation, and for any integer \( \beta_i, d_i \) can be chosen to be a polynomial function
of \( h_i(x_i) \), which has order \( \beta_i \) and positive coefficients, so that (2.2.5) is satisfied. However, since the applications in the following chapters involve only the case \( \beta_i = 2 \), the following corollary is restricted to deal only with this case.

Corollary 2.2.3 Let \( \mathcal{E}' \), \( v_i \), \( w_i \) and \( \alpha_i \) denote the same as in (2.2.1) and (2.2.2). If, for some constant \( M_i > 0 \),

\[
\frac{1}{v_i(x_i)} \leq M_i \quad \text{for all} \quad x_i \geq \alpha_i + 1, \quad (2.2.18)
\]

then

\[
d_i(x_i) = h_i^2(x_i) + b_i h_i(x_i) + b_0
\]

satisfies (2.2.5) for any constants, \( b_0, b_1, \ldots, b_p \) such that \( b_i \geq M_i \) and \( b_0 \geq 0 \). If, in addition, (2.2.6) is satisfied for this choice of the \( d_i(x_i) \), then \( \varphi_1, \ldots, \varphi_p \), as in (2.2.8), is a solution to the inequality (2.2.1), provided \( p > 2 \).

Proof: Again, it is only necessary to prove that (2.2.5) is satisfied for \( x_i \geq \alpha_i + 1 \). Now,

\[
\Lambda_i d_i(x_i) = \frac{1}{v_i(x_i)} [h_i(x_i)+h_i(x_i-1)] + \frac{b_i}{v_i(x_i)}.
\]

It follows that

\[
h_i(x_i-1)v_i(x_i)\Lambda_i d_i(x_i) = [h_i(x_i-1)h_i(x_i) + h_i^2(x_i-1)] + b_i h_i(x_i-1)
\]

\[
= 2h_i^2(x_i-1) + \frac{h_i(x_i-1)}{v_i(x_i)} + b_i h_i(x_i-1)
\]

\[
\leq 2h_i^2(x_i-1) + (M_i + b_i)h_i(x_i-1)
\]

\[
\leq 2d_i(x_i-1).
\]

Q.E.D.

Example 2.2 Assume \( X_i \) indep. NB(r, \theta_i). The uniformly minimum variance unbiased estimator of \( \theta_i \) is then \( X_i/(r-1+X_i) \). It will be shown
in the next chapter, that, for improving upon
\[
\left( \frac{x_i}{r-1+x_i}, \ldots, \frac{x_p}{r-1+x_p} \right)
\]
under the loss function \( L_0 \), the key difference inequality to solve is
\[
\frac{1}{p} \sum_{i=1}^{p} \left\{ \frac{x_i}{r-1+x_i} \Delta_i \phi_i(x) + \frac{1}{2} \phi_i^2(x) \right\} \leq 0. \tag{2.2.19}
\]
Here, \( v_i(x_i) = x_i/(r-1+x_i) \), \( w_i(x) = \frac{1}{2} \), \( a_i = 1 \) and \( \#_\alpha(x) \) denotes number of \( i \) such that \( x_i \geq 1 \). Also,
\[
h_i(x_i) = \sum_{k=1}^{\frac{k+r-1}{k}} x_i \geq 1
\]
\[
= 0 \quad \text{otherwise}.
\]
Since \( \frac{1}{v_i(x_i)} = \frac{x_i+r-1}{x_i} \leq \frac{r+1}{2} \), for \( x_i \geq 2 \), it is clear that (2.2.18) is satisfied for \( M = (r+1)/2 \). Let
\[
d_i(x_i) = h_i^2(x_i) + b h_i(x_i) + b_0,
\]
where \( b \geq M \) and \( b_0 \geq 0 \) are constants. Clearly,
\[
\frac{1}{2} \sum_{i=1}^{p} \frac{h_i^2(x_i)}{d_i(x_i)} \leq \frac{1}{2}.
\]
Hence, by corollary 2.2.3, for any function \( c(x) \) which is nondecreasing in each coordinate and which satisfies \( 0 \leq c(x) \leq 2(\#_\alpha(x)-2)^+ \), it follows that
\[
\phi_i(x) = \frac{-c(x)h_i(x_i)}{\sum_{j=1}^{p} \{ h_j^2(x_j) + b h_j(x_j) + b_0 \}}, \quad i = 1, \ldots, p,
\]
is a solution to (2.2.19), provided \( p \geq 3 \).
CHAPTER III
TYPICAL APPLICATIONS

In this chapter, the theorems in Chapter II are used to improve upon certain standard estimators under losses of the form \( L_m \). To be more precise, let \( \hat{X} \) be as in Lemma 2.2, and let

\[
\hat{\delta}^O = (\delta_1^O(X_1), \ldots, \delta_p^O(X_p)),
\]

where \( \delta_i^O(X_i) \) is the usual unbiased estimator of \( \theta_i \) (i.e. \( \delta_i^O(X_i) = t_i(X_i - 1)/t_i(X_i) \), where recall \( 0/0 \) is interpreted as \( 0 \)).

Under the loss function \( L_m \), we will develop classes of vector functions \( q(X) = (q_1(X), \ldots, q_p(X)) \) and \( \phi(X) = (\phi_1(X), \ldots, \phi_p(X)) \) which satisfy the four assumptions in Theorem 2.1. It will follow that \( \hat{\delta}^O + q\phi \) dominates \( \hat{\delta}^O \) and hence a class of better estimators will have been found.

Assumption (iii) in Theorem 2.1 indicates that for

\[
\delta_i^O(x_i) = t_i(x_i - 1)/t_i(x_i),
\]

it is sufficient to choose \( q_i(X) = 1 \), since, by plugging such \( \delta_i^O(x_i) \) into (2.1.13), it is clear that the right hand side of (2.1.13) is always zero if \( q_i(X) = 1 \), no matter which loss function \( L_m \) is assumed.

To satisfy assumption (iv) in Theorem 2.1, it is necessary to find a nontrivial solution \( \phi_1, \ldots, \phi_p \) to

\[
\sum_{i=1}^p \phi_i \delta_i^O(X_i) \leq 0,
\]

(3.1.2)
where, from (2.1.12) and the fact that $q_i(x) = 1$, $\Delta_i^i$ now has the form

$$
\Delta_i^i(x) = \frac{2t_i(x_i-m_i-1)}{t_i(x_i)} \Delta_i^i(x-m_i e_i) + \frac{t_i(x_i-m_i)}{t_i(x_i)} \phi_i^2(x-m_i e_i).
$$

(3.1.3)

In the following two sections, solutions to (3.1.2) are obtained by applying the theorems in Section 2.2. All the solutions below will be bounded, so that assumption (i) in Theorem 2.1 is automatically satisfied. Furthermore, it can be easily checked that all the $\phi_1, \ldots, \phi_p$ below satisfy assumption (ii) in Theorem 2.1. Therefore, in applying Theorem 2.1, we only describe the difference inequality and how the solutions are obtained. Also, in the following we will look for solutions to $\frac{1}{2} \phi'(\phi) \leq 0$ which is, of course, equivalent to $\phi' \phi \leq 0$.

In Section 3.1, it is assumed that the $X_i$'s are from Poisson families, while in Section 3.2, the $X_i$'s are from negative binomial families.

Section 3.1 Poisson Distributions

In this section, it is assumed that $X_i \sim \text{indep. Poisson}(\theta_i)$, $i=1, \ldots, p$, and hence that $t_i(x_i) = 1/x_i!$ (using the notation of Theorem 2.1).

Under the loss function $L_m$, where $m$ is some negative integer, $\delta^0$ can also be improved. This will follow from Theorem 3.1, in which the loss function is assumed to be $L_m$. In the remainder of the thesis, let $\gamma_{ij}$ denote Kronecker constant, i.e.
\[ \gamma_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise.} \end{cases} \]  

(3.1.4)

Furthermore, for any function \( c(x) \), we will use, "\( c(x) \not= 0 \)" to denote that \( c(x) \) is not identically zero.

**Theorem 3.1** Let \( \mathbf{m} = (m_1, \ldots, m_p) \), where the \( m_i \) are nonpositive integers, and let \( n_i \) and \( \alpha_i \) denote \(-m_i\) and \((m_i + 1)^+\) respectively. Assume that \( p > \max_{1 \leq j \leq p} (\alpha_i + 1) \). Under the loss function \( L_m \), the usual estimator \( \hat{\alpha}(x) = \hat{x} \) is inadmissible. Indeed a better estimator can be described as follows: Define

\[
\begin{align*}
  h_i(x_i) &= \begin{cases}
    x_i & x_i \geq 1 \\
    \frac{1}{k} \sum_{k=1}^{x_i} & x_i < 1
  \end{cases} \\
  & \quad \text{if } n_i = 0 \\
  0 & \text{otherwise,}
\end{align*}
\]

(3.1.5)

and

\[
\begin{align*}
  h_i(x_i) &= \begin{cases}
    \frac{1}{n_i} (x_i + 1) \ldots (x_i + n_i) & x_i \geq 1 \\
    0 & \text{otherwise,}
  \end{cases} \\
  & \quad \text{if } n_i > 0.
\end{align*}
\]

Also, define, for some constant \( b_0 \geq 0 \),

\[
\begin{align*}
  d_i(x_i) &= h_i(x_i) h_i(x_i + 1) + b_0 \quad \text{if } n_i = 0 \\
  &= \frac{1}{n_i} h_i(x_i) + b_0 \quad \text{if } n_i > 0.
\end{align*}
\]

(3.1.6)

Let \( \#(x) \) denote the number of indices \( i \) for which \( x_i \geq \alpha_i \). For any nonnegative number \( b_0 \) and any function \( c(x) \) which is nondecreasing in each coordinate and satisfies \( c(x) \not= 0 \) and

\[ 0 \leq c(x) \leq 2(\#(x) - \max_{1 \leq j \leq p} (\alpha_j + 1))^+, \]

(3.1.7)
\( \phi^0 \) is dominated by \( \phi^0(x) + \psi(x) \), where

\[
\phi_i(x) = \frac{-c(x-n_i e_i)h_i(x_i-n_i)}{\sum_{j=1}^{p} d_j(x_j-n_j \gamma_{ij})}.
\]

(3.1.8)

Proof: By (3.1.1) and (3.1.2), we have to obtain a solution \( \phi_1, \ldots, \phi_p \) to the following difference inequality:

\[
\frac{1}{2} \Delta^n (\psi(x)) = \sum_{i=1}^{p} [v_i(x_i) \Delta_i \psi_i(x) + w_i(x_i) \psi_i^2(x)] \leq 0,
\]

(3.1.9)

where

\[
\psi_i(x) = \phi_i(x_i + n_i e_i),
\]

(3.1.10)

and

\[
v_i(x_i) = x_i
\]

if \( n_i = 0 \)

\[
w_i(x_i) = \begin{cases} 1/2 & \text{if } n_i = 0 \\ \frac{x_i!}{(x_i+n_i)!} & \text{if } n_i > 0. \end{cases}
\]

(3.1.11)

or

\[
v_i(x_i) = \frac{x_i!}{(x_i+n_i)!} \quad \text{if } n_i > 0.
\]

\[
w_i(x_i) = \frac{x_i!}{2(x_i+n_i)!}.
\]

To find a solution to (3.1.9), we will use Theorem 2.2 as a guide.

Note that \( v_i(x_i) > 0 \) if \( x_i > \alpha_i \), and it follows from (3.1.5) and (3.1.11) that

\[
h_i(x_i) = \sum_{k=\alpha_i}^{x_i} \frac{1}{v_i(k)} \quad x_i > \alpha_i
\]

\[
= 0 \quad \text{otherwise}.
\]

When \( n_i = 0 \), \( v_i(x_i) \) is increasing. By Corollary 2.2.1, \( d_i(x_i) \)
satisfies (2.2.5) with $\beta_i = 2$. Also by Corollary 2.2.2, $d_i(x_i)$ satisfies (2.2.5) with $\beta_i = 1$ when $n_i > 0$. Therefore assumption (ii) of Theorem 2.2 is satisfied with $\beta_i = \alpha_i + 1$. It is also clear that

$$W_i(x_i)h_i^2(x_i) = h_i^2(x_i)/2 \leq d_i(x_i)/2 \quad \text{if } n_i = 0$$

and

$$W_i(x_i)h_i^2(x_i) = h_i^2(x_i)/2\pi n_i < d_i(x_i)/2 \quad \text{if } n_i > 0.$$

Therefore, assumption (iii) of Theorem 2.2 is satisfied, and hence a solution to (3.1.9) is $\psi_1, \ldots, \psi_p$, where

$$\psi_i(x) = \frac{-c(x)h_i(x_i)}{\sum_{j=1}^{p} d_j(x_j)} , \quad i = 1, \ldots, p, \quad (3.1.12)$$

or equivalently

$$\phi_i(x) = \frac{-c(x-x_i e_i)h_i(x_i-x_i e_i)}{\sum_{j=1}^{p} d_j(x_j)} , \quad i = 1, \ldots, p. \quad (3.1.13)$$

Also, it is clear that

$$E_{\tilde{\varphi}}^{\infty}(\psi) < 0$$

for all $\tilde{\varphi}$. Theorem 2.1, thus implies that $\tilde{\varphi}^0 + \phi$ dominates $\tilde{\varphi}^0$. Q.E.D.

Corollary 3.1.1 (Clevenson and Zidek 1975) Assume that $p \geq 2$.

Under the loss function $L_{-1}$, $\tilde{\varphi}^0(X) = X$ is inadmissible. Indeed for any constant $b_0 \geq 0$, and any function $c(x)$ which is nondecreasing in each coordinate and satisfies $c(x) \not\equiv 0$ and

$$0 \leq c(x) \leq 2(p-1), \quad (3.1.14)$$

$\tilde{\varphi} + \phi$ dominates $\tilde{\varphi}^0$, where $\phi = (\phi_1, \ldots, \phi_p)$ and

$$\phi_i(x) = \frac{-c(x-x_i e_i)X_i}{b_0 + p-1 + \sum_{j=1}^{p} X_j}. \quad (3.1.15)$$
Proof: Clearly $m_i = -1$, $n_i = 1$ and $a_i = 0$, $i = 1, \ldots, p$. Hence
\[ \#_0(x) \] is the number of indices $i$ for which $x_i \geq 0$. It is clear that $\phi_i$ in (3.1.8) is reduced to the form in (3.1.15) with a different $b_0$. To complete the proof, it is only necessary to show that the condition (3.1.7) is equivalent to (3.1.14). Now, for each $c(x)$ satisfying (3.1.14), we will design a version $c'(x)$ of $c(x)$ so that (3.1.7) is satisfied ($c'(x)$ is also nondecreasing in each coordinate and not identically zero) and $c'(x) - e_j x_i = c(x) - e_j x_i$, $i = 1, \ldots, p$ with probability one. Therefore $\phi_1, \ldots, \phi_p$ remain the same with probability one even if $c$ is replaced by $c'$ in (3.1.13). This proves that (3.1.14) is in fact equivalent to (3.1.7). Indeed let $A = \{(x_1, \ldots, x_p); x_i \geq 0\}$ and
\begin{equation}
\tag{3.1.16}
c'(x) = c(x)I_A(x).
\end{equation}
Since with probability one, $x_i \geq 0$, $i = 1, \ldots, p$, it is clear that with probability one,
\[ c(x) - e_j x_i = c'(x) - e_j x_j \quad j = 1, \ldots, p. \]
Now $c'(x)$ is nondecreasing in each coordinate and is not identically zero. Furthermore $c'(x)$ satisfies (3.1.7) hence we are done. Q.E.D.

The better estimators obtained in Clevenson and Zidek (1975) correspond to those $\delta^0 + \delta$ with $c(x)$ depending on $x$ only through
\[ \sum_{i=1}^{p} x_i. \]

The main idea in the proof in Corollary 3.2.1 will be used in other occasions. Therefore a more general lemma is established here.

Lemma 3.1. Assume that $c(x)$ is defined on $I^P$ (I is the set of all integer.) which satisfies the following conditions:
(i) $c(x) \equiv 0$;
(ii) $c(x)$ is nondecreasing in each coordinate;
(iii) $0 \leq c(x) \leq n_0(p-\beta)$, where $n_0$ and $\beta$ are some positive constants.

Let $h_j(x_j)$ be as in (2.2.3) with $\alpha = (0, \ldots, 0)$, and let $c'(x)$ denote $c(x)I_A(x)$, where $A = \{(x_1, \ldots, x_p); x_i \geq 0\}$. Then, conditions (i) and (ii) above are satisfied when $c(x)$ is replaced by $c'(x)$. Besides, the following are true:

(iii)' $0 \leq c'(x) \leq n_0(\#(x) - \beta)^+$ for all $x \in \mathbb{R}^p$;
(iv) $c(x - ne_j)h_j(x_j - n) = c'(x - ne_j)h_j(x_j - n)$, $j = 1, \ldots, p$ for all $x \in A$ and any integer $n$.

Proof: It is obvious that $c'(x)$ satisfies conditions (i) and (ii). If $x \notin A$, then $c'(x) = c(x)$ and $\#(x) = p$, condition (iii)' is thus equivalent to condition (iii) and hence is satisfied. If $x \notin A$, then $c'(x) = 0$, and condition (iii) is clearly satisfied. Therefore condition (iii)' is satisfied for all $x \in \mathbb{R}^p$. To prove condition (iv), assume that $x \notin A$. For any index $j$, if $x_j > n$, then $(x - ne_j) \notin A$ and condition (iv) is true since $c'(x - ne_j) = c(x - ne_j)$. If $x_j < n$, then $h_j(x_j - n) = 0$ and hence condition (iv) is trivially satisfied. Therefore, condition (iv) holds for all $x \in A$.

Q.E.D.

Corollary 3.1.2. (Tsui and Press 1977) Assume that $p \geq 2$. Under the loss function $L_m$, where $m = -n$ is some negative integer, $c^0(x) = x$ is inadmissible. Indeed, for any nonnegative number $b_0$ and any function $c(x)$ which is nondecreasing in each coordinate and satisfies $c(x) \equiv 0$ and

$$0 \leq c(x) \leq 2n(p-1),$$

(3.1.17)

$c^0$ is dominated by $c^0 + \psi$, where $\psi = (\phi_1, \ldots, \phi_p)$ with
\[
\phi_i(x) = \frac{-c(x - n\varepsilon_i)h(x_i - n)}{b_0 + \sum_{j=1}^{p} h(x_j - n\gamma_j)} ,
\]

and
\[
h(x) = (x+1)...(x+n) \quad \text{if} \quad x \geq 0
\]
\[
= 0 \quad \text{otherwise.}
\]

Proof: Clearly \( \alpha_i = 0, \ i = 1,...,p. \) Again the main step of the proof is to show that the following assumption about \( c(x) \),
\[
0 \leq c(x) \leq 2n(\#_\varepsilon (x) - 1)^+ ,
\]

(3.19)
can be replaced by (3.17). However, this is an immediate result of Lemma 3.1. Q.E.D.

The better estimators obtained in Tsui and Press (1977) correspond to those \( \delta^0 + \xi \) with \( c(X) \) in (3.18) depending on \( X \) only through \( \sum_{i=1}^{p} x_i. \)

For the loss function \( L_0 \), the following corollary is a direct result from Theorem 3.1.

Corollary 3.1.3. Assume \( p \geq 3 \), under the loss function \( L_0, \delta^0(x) = x \)
is inadmissible. Indeed, a better estimator can be described as follows: Denote the number of indices \( i \) for which \( x_i \geq 1 \) by \( \#_1(x) \).

Define
\[
h(x_i) = \sum_{k=1}^{x_i} \frac{1}{k} \quad x_i \geq 1
\]
\[
= 0 \quad \text{otherwise}
\]
Let \( c(x) \) be any function which is increasing in each coordinate and satisfies \( c(x) \rhd 0 \) and
\[
0 \leq c(x) \leq 2(\#_1(x) - 2)^+
\]
Then, for any constant $b_0$, the new estimator $\hat{\delta}^*$ with the $i$th component,

$$\hat{\delta}_i^*(X) = x_i - \frac{c(x)h(x_i)}{b_0 + \sum_{j=1}^{p} h(x_j)h(x_{j+1})}$$  \hspace{1cm} (3.1.20)$$

is a better than $\delta^0$.

The improved estimator obtained in Peng (1975) (refer to (1.3.7)) is very similar to $\hat{\delta}^*$ with $c(x) = (#_1(x)-2)^4$ and $b_0 = 0$.

As another interesting application of Theorem 3.1, consider the situation $p = 3$ and $m = (0,-1,2)$. By Theorem 3.2, $\delta^0(X) = X$ is inadmissible and is dominated by $\delta^0 + \phi$ with $\phi_i$ as in (3.1.8). The functions $h_1, h_2, h_3$ and $d_1, d_2, d_3$, which determine the form of the correction terms $\phi_i$, are given in (3.15) and (3.1.6). It is clear that $h_1$ and $d_1$ are similar to $h(x_i)$ and $h^2(x_i)$ in Peng's estimator (cf. (1.3.7)). Also, $h_2$ and $d_2$ have the same form as in Cleevenson's estimator (cf. (3.1.15)), and $h_3$ and $d_3$ are as in Tsui's estimator (cf. (3.1.18)). Note that the choice of $h_i$ and $d_i$ depends only upon $m_i$ (and not $m_j$, $j \neq i$). A similar property is also observed in the negative binomial case (Section 3.1), as well as in the more general case in which the densities of the $X_i$ are not of the same form (Section 5.3). In this general situation, the choice of $h_i$ and $d_i$ depends on $m_i$ and the density of $X_i$ (i.e. $t_i(x_i)$), and not on the other coordinates.

In the earlier work for the Poisson case, the proofs that the estimators presented (for example in Peng (1975), Cleevenson and Zidek (1973) and Hudson (1978)) are better than $\delta^0(X) = X$ are heavily based on the symmetry of the problems. (i.e. the $m_i$ are all equal,
and the $X_i$'s have the same type of distribution.) Theorem 3.2 shows,
however, that Stein's effect seems to be a property more basic than
symmetry. For the continuous case, this was also observed in Berger
(1978).

Section 3.2. Negative Binomial Distribution

In this section, assume that $X_i \overset{\text{indep.}}{\sim} \text{NB}(r_i, \theta_i)$, where $r_i$
is some positive integer, $i = 1, \ldots, p$, and hence that

$$t_i(x_i) = \begin{pmatrix} r_i + x_i - 1 \\ r_i - 1 \end{pmatrix}, \quad x_i = 0, 1, \ldots .$$

One of the classical estimators of $\theta = (\theta_1, \ldots, \theta_p)$ is

$$\hat{\theta}_0^0(X) = \left( \frac{X_1}{r_1 - 1 + X_1}, \ldots, \frac{X_p}{r_p - 1 + X_p} \right),$$

where $X_i/(r_i - 1 + X_i)$ is the uniformly minimum variance unbiased estimator of $\theta_i$. For $p = 1$, $\hat{\theta}_0^0$ is admissible under the square error loss.
(See Blackwell and Girshick (1954) p. 307) and is hence admissible
under any loss function of the form $L(\theta, a) = v(\theta)(a - \theta)^2$ with $v(\theta) > 0$
for all $\theta$.

Our goal here is to improve upon $\hat{\theta}_0^0$ under the loss function $L_m$. Therefore, a solution to $\mathcal{D}'(\hat{\theta}) \leq 0$ (defined in (3.1.2) and (3.1.3)).

Since $t_i(x_i) = \begin{pmatrix} r_i + x_i - 1 \\ r_i - 1 \end{pmatrix}$, the difference inequality $\frac{1}{2} \mathcal{D}'(\hat{\theta}) \leq 0$
has the following form:

$$\frac{1}{2} \mathcal{D}'(\hat{\theta}) = \sum_{i=1}^{p} \left[ v_i(x_i) \hat{\phi}_i \hat{\phi}_i (x - m_i e_i) + w_i(x_i) \hat{\phi}_i^2 (x - m_i e_i) \right] \leq 0, \quad (3.2.1)$$
where
\[ v_i(x_i) = \begin{cases} \frac{r_i + x_i - m_i - 2}{r_i - 1} & \text{if } x_i \geq (m_i + 1)^+ \\ 0 & \text{otherwise,} \end{cases} \] (3.2.2)
and
\[ w_i(x_i) = \begin{cases} \frac{r_i + x_i - m_i - 1}{r_i - 1} \times \frac{r_i + x_i - 1}{r_i - 1} & \text{if } x_i \geq m_i^+ \\ 0 & \text{otherwise.} \end{cases} \] (3.2.3)

When \( m_i = 0 \) and \( r_i = r \), \( i = 1, \ldots, p \), \( v_i(x_i) = x_i / (x_i + r_i - 1) \) and \( w_i(x_i) = 1/2 \). The corresponding difference inequality (3.2.1) was solved in Example 2.2 by applying Corollary 2.2.3. The same corollary is also applicable to the general case as seen in the following theorem.

In the theorem below, define for \( i = 1, \ldots, p \),
\[ \alpha_i = (m_i + 1)^+, \] (3.2.4)
\[ h_i(x_i) = \sum_{k=\alpha_i}^{r_i}(r_i + k - 1) / (r_i - 1) \] if \( x_i \geq \alpha_i \)
= 0 otherwise, \hspace{1cm} (3.2.5)
\[ M_i = \frac{1}{r_i} \begin{pmatrix} r_i + \alpha_i \\ r_i - 1 \end{pmatrix} \] (3.2.6)
\[ N_i = 1 \quad \text{if} \quad m_i \geq 0 \]  \tag{3.2.7} \\
\begin{pmatrix}
    r_i^{-m_i-1} \\
    r_i-1
\end{pmatrix}
\quad \text{otherwise},

and

\[ K = \max_{1 \leq i \leq p} \{ N_i \}. \tag{3.2.8} \]

**Theorem 3.2.** Assume that \( p \geq 3 \). Let \( m = (m_1, \ldots, m_p) \), with \( m_i \) being any integer. Under the loss function \( l_m \), the estimator

\[ \hat{\theta}^0(x) = \left( \frac{x_1}{r_1 + x_1^{-1}}, \ldots, \frac{x_p}{r_p + x_p^{-1}} \right) \]

is inadmissible. Indeed, a better estimator can be described as follows: Let \( \alpha(x) \) denote the number of indices \( i \) for which \( x_i \geq \alpha_i \). Furthermore, let \( b_0 \geq 0, b_j \geq M_j, \ j = 1, \ldots, p, \) and \( c(x) \) be any function that is nondecreasing in each coordinate and satisfies \( c(x) \neq 0 \) and

\[ 0 \leq c(x) \leq 2(\#_2(x) - 2)^r / K. \tag{3.2.9} \]

Then, \( \hat{\theta}^0 \) is dominated by \( \hat{\theta}^0 + \hat{\phi} \), where \( \hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_p) \) and

\[ \hat{\phi}_i(x) = \frac{-c(x + m_i e_i) h_i(x + m_i)}{b_0 + \sum_{j=1}^p \left( h_j^0(x + m_i y_{ij}) + h_j(x + m_i y_{ij}) \right)} \tag{3.2.10} \]

**Proof:** In applying Theorem 2.1, it remains only to show that \( \hat{\phi} \) is a solution to (3.2.1). Corollary 2.2.3 is applicable, since it is clear that, for \( m_i \geq 0 \),
\[
\max_{x_i \geq \alpha_i + 1} \frac{1}{v_i(x_i)} = \max_{x_i \geq \alpha_i + 1} \frac{(\gamma_i + x_i - m_i - 1) \cdots (\gamma_i + x_i - 1)}{(x_i - m_i) \cdots x_i} = \frac{1}{r_i} \left( \frac{r_i + m_i + 1}{r_i - 1} \right),
\]

and for \( m_i < 0 \) and \( x_i \geq \alpha_i \),

\[
\frac{1}{v_i(x_i)} \leq 1.
\]

It thus follows from (3.2.6) that

\[
\frac{1}{v_i(x_i)} \leq M_i \tag{3.2.11}
\]

for \( x_i \geq \alpha_i + 1 \), and so (2.2.18) is satisfied. Similarly, an upper bound on \( w_i(x_i) \) in (3.2.3) is

\[
w_i(x_i) \leq \begin{cases} 
1/2 & \text{if } m_i \geq 0 \\
\frac{(r_i - m_i - 1)}{2} & \text{otherwise},
\end{cases}
\]

or equivalently,

\[
w_i(x_i) \leq N_i / 2. \tag{3.2.12}
\]

Let

\[
\psi_i(x) = \frac{-c(x)h_i(x_i)}{D}, \tag{3.2.13}
\]

where \( D = b_0 + \sum_{j=1}^{p} \{ h_j^2(x_j) + b_j h_j(x_j) \} \). It is clear that

\[
\frac{\sum_{i=1}^{p} w_i(x_i) h_i^2(x_i)}{D} \leq \frac{1}{2} \max_{1 \leq j \leq p} \{ N_j \} \frac{\sum_{i=1}^{p} h_i^2(x_i)}{D} \leq \frac{1}{2} K,
\]
so that (2.2.6) is satisfied. Hence, by Corollary 2.2.3, 
\( \psi_i(x - m_\xi e_1) = \psi_i(x) \) satisfies (3.2.1), with strict inequality on a set of \( x \) of positive probability. Q.E.D.

The corollaries below follow immediately from Theorem 3.2.

Corollary 3.2.1 Assume that \( p \geq 3 \). Under the loss function \( L_{0,0}^0 \) is inadmissible. Indeed, let \( \#_1(x) \) denote the number of indices \( i \) for which \( x_i \geq 1 \), and define

\[
h(x_i) = \sum_{k=1}^{x_i} \frac{r_i - 1 + k}{k} \quad x_i \geq 1
\]

\[
= 0 \quad \text{otherwise.}
\]

Furthermore, let \( b_0 \geq 0, b_j \geq \frac{1 + r_j}{2} \) and \( c(x) \) be any function nondecreasing in each coordinate, which satisfies \( c(x) \parallel 0 \) and

\[
0 \leq c(x) \leq 2(\#_1(x) - 2)^+.\]

Then, \( \hat{\delta}^0 \) is dominated by \( \hat{\delta}^0 + \phi \), where \( \phi = (\phi_1, \ldots, \phi_p) \) and

\[
\phi_i(x) = \frac{-c(x)h(x_i)}{b_0 + \sum_{j=1}^p (h(x_j) + b_j h(x_j))}. \quad (3.2.14)
\]

Note that Hudson (1978) proved that \( \hat{\delta}^0 \) is inadmissible under the loss function \( L_{0,0}^0 \) when \( p \geq 4 \). His improved estimator was given in (1.3.9).

Corollary 3.2.2 Assume that \( p \geq 3 \). Under the loss function \( L_{-1,0}^0 \) is inadmissible. Indeed, let \( b_0 \geq 0, b_j \geq 1, j = 1, \ldots, p \), and \( c(x) \) be any function nondecreasing in each coordinate, which satisfies \( c(x) \parallel 0 \) and

\[
0 \leq c(x) \leq \frac{2(p - 2)}{\max\{r_1, \ldots, r_p\}}. \quad (3.2.15)
\]

Then, \( \hat{\delta}^0 \) is dominated by \( \hat{\delta}^0 + \phi \), where \( \phi = (\phi_1, \ldots, \phi_p) \) and
\[ \phi_i(x) = \frac{-c(x-e_i)x_i}{b_0 + \sum_{j=1}^{p} \{(X_j+1-\gamma_{ij})^2 + b_j(X_j+1-\gamma_{ij})\}}. \] (3.2.16)

Proof: By Theorem 3.2, \( \delta^0 + \phi \) dominates \( \delta^0 \) if \( c(x) \) satisfies
\[
0 \leq c(x) \leq 2(\#(x)-2)^+/\max\{r_1,\ldots,r_p\}.
\]
where \( \alpha = (0,\ldots,0) \). Lemma 3.1 then implies that this condition is equivalent to (3.2.15).

Q.E.D.

Corollary 3.2.3 Assume \( p \geq 3 \). Under the loss function \( L_m \), \( m = -2,-3,\ldots \), \( \delta^0 \) is inadmissible. Indeed, define
\[
n = -m
\]
\[
K = \max_{1 \leq i \leq p} \left( \frac{r_i+n_i-1}{r_i-1} \right)
\]
and, for \( i = 1,\ldots,p \),
\[
h_i(x_i) = \sum_{k=0}^{x_i} \frac{(k+1)\cdots(k+n-1)}{(k+r_i)\cdots(k+n+r_i-2)} x_i \geq 0,
\]
\[
= 0 \quad \text{otherwise.}
\]
Let \( b_0 \geq 0, b_j \geq 1, j = 1,\ldots,p \), and \( c(x) \) be any function nondecreasing in each coordinate such that \( c(x) \not\equiv 0 \) and
\[
0 \leq c(x) \leq 2(p-2)/K. \quad (3.2.17)
\]
Then \( \delta^0 \) is dominated by \( \delta^0 + \phi \), where \( \phi = (\phi_1,\ldots,\phi_p) \) and
\[
\phi_i(x) = \frac{-(cX-ne_i)h_i(x_i-ne_i)}{b_0 + \sum_{i=1}^{p} \{h_i^2(X_j-n\gamma_{ij}) + b_jh_i(X_j-n\gamma_{ij})\}}. \quad (3.2.18)
\]
Proof: An argument similar to the proof of Corollary 3.2.2 shows that (3.2.9) is equivalent to (3.2.17).

Q.E.D.
Corollary 3.2.4 Assume $p \geq 3$. Under the loss function $L_m$, $m = 1, 2, \ldots$, $\bar{\delta}^0$ is inadmissible. Indeed, let $\#(x), x=(m+1, \ldots, m+1)$, denote the number of indices $i$ such that $x_i \geq m+1$, and define

$$h_i(x_i) = \sum_{k=m+1}^{x_i} \frac{(r_i-1+k-m) \ldots (r_i-1+k)}{(k-m) \ldots k} \quad x_i \geq m+1$$

$$= 0 \quad \text{otherwise.}$$

Furthermore, let $b_0 \geq 0$, $b_j \geq \frac{(r_j+1) \ldots (m+r_j+1)}{(m+2)!}$, $j = 1, \ldots, p$, and $c(x)$ be any function, nondecreasing in each coordinate such that $c(x) \neq 0$ and

$$0 \leq c(x) \leq 2(\#(x)-2)^+.$$ 

Then $\bar{\delta}^0$ is dominated by $\bar{\delta}^0 + \psi$, where $\psi = (\psi_1, \ldots, \psi_p)$ and

$$\psi_i(x) = \frac{-c(x+me_i)h_i(x_i+m)}{b_0 + \sum_{j=1}^{p} \{h_j^2(x_j+m\gamma_{ij}) + b_j h_j(x_j+m\gamma_{ij})\}}.$$ 

(3.2.19)
CHAPTER IV

GENERAL INADMISSIBILITY RESULTS

In the last Chapter, it was shown that by choosing \( q_i(x) = 1 \) (i.e. write the new estimator \( \delta^* \) as \( \delta_i^0 + \delta_i^1 \)), the problems of improving upon the uniformly minimum variance unbiased estimators can be reduced to the study of a difference inequality of the form (2.2.1). However, since many reasonable estimators are not unbiased, a more sophisticated choice of the \( q_i \)'s is needed when one tries to improve upon such general estimators. In this chapter, it will be shown how the theorems in Chapter II can be applied to improve upon other estimators by choosing appropriate \( q_i \)'s.

As pointed out at the end of Section 2.1, it is not difficult to choose \( q_i \)'s so that assumption (iii) of Theorem 2.1 is satisfied. For instance, assume that \( X_i \overset{\text{iind}}{\overset{\text{P}}{\sim}} \text{NB}(1, \theta_i) \), \( i = 1, \ldots, p \). The uniformly minimum variance unbiased estimator of \( \theta_i \) is

\[
\delta(x_i) = \begin{cases} 
1 & x_i \geq 1 \\
0 & x_i = 0 
\end{cases}
\]

(refer to Section 3.2.) which is not an interesting estimator to improve upon. A more appealing estimator is \( \delta^G(x) = (\delta^G_1(x_1), \ldots, \delta^G_p(x_p)) \), with \( \delta^G_i(x_i) = \frac{x_i}{x_i + \epsilon_i} \), \( 0 < \epsilon_i < 1 \). For \( p = 1 \), \( \delta^G(x) \) is an admissible generalized Bayes estimator, under \( L_0 \). However, for higher dimension, \( \delta^0 \) can be improved upon by applying Theorem 2.1. To choose \( q_i \)'s so
that the inequality (2.1.13) is satisfied, it is sufficient to set

\[ q_i(x_i) = \prod_{k=1}^{x_i k + \epsilon_i} \frac{k}{k \epsilon_i} \quad x_i = 1, 2, \ldots \]

\[ = 1 \quad x_i = 0 \]

\[ = 0 \quad x_i < 0 \]

It is easy to check that the equality in (2.1.13) holds for such a choice of \( q_i \)'s. The problem is then reduced to the study of the difference in equality (2.1.14). Since \( q_i \) depends on \( x_i \) only, the difference inequality (2.1.14) is of the form (2.2.1), solutions of which are given by a generalization of Corollary 2.2.1. Better estimators are then obtained for \( p > \max_{1 \leq i \leq p} \frac{2}{1 - \epsilon_i} \). This proves that under \( L_0 \), \( \hat{\delta} \) is inadmissible for such \( p \). The detailed calculation will be reported elsewhere.

All the estimators so far considered have had \( \delta_i^0(x) \) depending on \( X \) only through \( X_i \). If, however, \( \delta_i^0(x) \) depends on the whole \( x \), it is also interesting to see if improvements can be obtained. Of course, we can choose the \( q_i \)'s so that equality (rather than inequality) in (2.1.13) is satisfied, as was done in the previous example. But then, the \( q_i \)'s depend on the whole \( x \), and the difference inequality that must be solved has the form of (1.4.1) with \( v_i \) depending on the whole \( x \). To solve such a difference inequality seems to be very hard. Therefore, in Section 4.1, the \( q_i \)'s will be chosen so that not only is the inequality (2.1.13) satisfied, but also the \( q_i \)'s lead to a difference inequality \( \sum_{i=1}^{d} \delta_i^1 < 0 \) of the type (2.1.18), to which the theorems in Section 2.2 can be applied. We thus establish a theorem which indicates the inadmissibility of many estimators obtained in Chapter III or by other statisticians.
In Section 4.2, another theorem is developed using the \( q_i \) functions, which establishes "upper bounds" on the class of admissible estimators in certain sense.

Section 4.1. Applications to the General Estimators

In this section, we will try to improve upon estimators of a general form. Before doing so, a theorem will be developed which gives solutions to the difference inequality \( \xi'(\phi) \leq 0 \), defined in (2.2.1), under much weaker assumptions. Although the solution \( \phi_1, \ldots, \phi_p \) takes zero value when the \( x_i \)'s are small, the theorem is a useful tool in proving inadmissibility.

The following lemma will be used in establishing the theorem.

Lemma 4.1 Consider the general difference inequality

\[
\xi(\phi(x)) = \sum_{i=1}^{p} v_i(x) \Delta_i \phi_i(x) + w_i(x) \phi_i^2(x) \leq 0, \tag{4.1.1}
\]

where \( v_i(x) \) and \( w_i(x) \) are nonnegative for all \( x \). Suppose that

\[
\phi^*(x) = (\phi^*_1(x), \ldots, \phi^*_p(x))
\]

is a solution to (4.1.1) and \( \phi^*_i(x) \leq 0 \) for all \( x \) and all \( i, 1 \leq i \leq p \).

For any function \( F(x), 0 \leq F(x) \leq 1 \), which is nondecreasing in each coordinate, \( F(x)\phi^*(x) \) is also a solution to (4.1.1). (Recall from (1.2.10), that \( F(x)\phi^*(x) = (F(x)\phi^*_1(x), \ldots, F(x)\phi^*_p(x)). \) Furthermore,

\[
\xi(F(x)\phi^*(x)) \leq F(x)\xi(\phi^*(x)). \tag{4.1.2}
\]

Proof: It is sufficient to show that (4.1.2) is satisfied.

Now
\[ \phi(F(x)\phi^*(x)) = \sum_{i=1}^{P} \{v_i(x)\Delta_i(F(x)\phi_i^*(x)) + w_i(x)(F(x)\phi_i^*(x))^2 \} \]

\[ \leq \sum_{i=1}^{P} \{v_i(x)F(x)\Delta_i\phi_i^*(x) + F(x)w_i(x)(\phi_i^*(x))^2 \} \]

\[ = F(x)\phi^*(x) \quad \text{Q.E.D.} \]

A direct application of Lemma 4.1 gives Lemma 4.2 which illustrates a key idea that is used in proving Theorem 4.1. In the remainder of this thesis, let \( I_A(x) \) denote the indicator function, i.e.

\[ I_A(x) = 1 \quad \text{if} \quad x \in A \]

\[ = 0 \quad \text{if} \quad x \notin A. \]

Lemma 4.2 Consider two difference inequalities

\[ \delta(x) = \sum_{i=1}^{P} \{v_i(x)\Delta_i\phi_i^2(x) + w_i(x)(\phi_i^2(x))^2 \} \leq 0, \quad (4.1.3) \]

and

\[ \delta'(x) = \sum_{i=1}^{P} \{v_i(x)\Delta_i\phi_i^2(x) + w_i(x)(\phi_i^2(x))^2 \} \leq 0. \quad (4.1.4) \]

Let \( A = \{(x_1, \ldots, x_p) : x_i \geq \alpha \} \), where \( \alpha \) is some number. Suppose that \( v_i(x) = v_i(x) \) and \( w_i(x) = w_i(x) \) for \( x \in A \) and \( i = 1, \ldots, p \). If

\[ \phi^* = (\phi_1^*, \ldots, \phi_p^*) \], with \( \phi_i^*(x) \leq 0, \quad i = 1, \ldots, p, \]

is a solution to \( (4.1.3) \) then \( \phi^*I_A \) is also a solution to \( (4.1.4) \).

Proof: For \( x \in A \), it is clear that \( \delta'(\phi^*(x)) = 0 \). For \( x \notin A \), it follows from Lemma 4.1 that \( \delta(\phi^*(x)) \leq 0 \). But since

\[ \delta(\phi^*(x)) = \delta'(\phi^*(x)) \text{ for } x \in A, \text{ it can be concluded that} \]

\[ \delta'(\phi^*(x)) \leq 0. \quad \text{Q.E.D.} \]

The implications of Lemma 4.2 are interesting. In solving the difference inequality \( (4.1.3) \) using negative \( \phi_i \)'s, Lemma 4.2 states
that the functional forms of \( v_i \) and \( w_i \) are unimportant for all small \( x_i \). To be more precise, we will say that a statement is true for sufficiently large \( x \) when there exists an \( M \) such that the statement is true for all \( x = (x_1, \ldots, x_p) \) with \( x_i \geq M, \ i = 1, \ldots, p \). Therefore, if \( v_i \) and \( w_i \) equal \( v_1^i \) and \( w_1^i \) respectively for sufficiently large \( x \), then the class of solutions to \( \mathcal{L}(\phi) \) and \( \mathcal{L}'(\phi) \) are the same for sufficiently large \( x \). In proving inadmissibility, we therefore need be concerned only with \( x \) which are sufficiently large. The next theorem makes use of this idea to obtain solutions to \( \mathcal{L}(\phi) \leq 0 \) under weak assumptions.

Theorem 4.1 Let \( \mathcal{L}'(\phi) \) be as in (2.2.1). Also let the assumptions about \( v_i(x_i), w_i(x) \) and \( \alpha_i \) be the same. Define \( h_i \) as in (2.2.3). Suppose, for \( i = 1, \ldots, p \), exists nonnegative constants \( b_0 \) and \( \lambda \) and positive constants \( \beta_i, U_i, K \), and \( \alpha_i, \alpha_i' \geq \alpha_i + 1 \), such that

\[
(i) \quad \frac{h(x_i)}{h(x_i-1)} \leq U_i \quad \text{for} \quad x_i \geq \alpha_i', \tag{4.1.5}
\]

\[
(ii) \quad \frac{\sum_{i=1}^{p} w_i(x)h_i^2(x_i)}{b_0 + D \lambda} \leq K \quad \text{for} \quad x_i \geq \alpha_i', \quad i = 1, \ldots, p, \tag{4.1.6}
\]

where \( D = \sum_{i=1}^{p} h_i(x_i) \).

Let \( \beta' \) denote

\[
\beta'(x_i) = \left( \frac{\beta_i - 1}{\beta_i} \right)^{\lambda} \left( \max_{1 \leq i < p} \beta_i U_i \right) \left( \max_{1 \leq i < p} \beta_i U_i \right)
\]

and define

\[
A = \{(x_1, \ldots, x_p): x_i \geq \alpha_i', \ i = 1, \ldots, p\}.
\]

Then for \( p > \beta' \), \( \phi = (\phi_1, \ldots, \phi_p) \) with
\[ \phi_i(x) = \frac{-ch_i(x_i)}{b_0 + D^\lambda} I_A(x), \]  

(4.1.7)

is a solution to \( \mathcal{L}'(\phi) \leq 0 \), providing \( c \) is any constant satisfying

\( 0 < c < (p-\beta')/K \). Furthermore

\[ x'(\phi) \leq \frac{-c(p-\beta'-cK)}{b_0 + D^\lambda} I_A(x). \]  

(4.1.8)

Proof. Note that \( p-\beta'-cK > 0 \). Therefore it is sufficient to prove that (4.1.8) holds. Let \( \psi = (\psi_1, \ldots, \psi_p) \), where

\[ \psi_i(x) = \frac{-ch_i(x_i)}{b_0 + D}. \]  

(4.1.9)

Clearly \( \phi_i = \psi_i I_A \). For \( \bar{x} \notin A \), (4.1.8) is trivial. If \( \bar{x} \in A \), we will show that

\[ x'(\bar{x}) \leq \frac{-c(p-\beta'-cK)}{b_0 + D^\lambda}. \]  

(4.1.10)

Together with Lemma 4.1, this will show that (4.1.8) holds. For \( \bar{x} \in A \), let

\[ D_j = h_1^i(x_i - 1) + \sum_{j \neq i}^{p} h_1^i(x_j) \]

and

\[ D' = \sum_{j=1}^{p} h_j^i(x_j - 1). \]

Now it is clear that
\[ \sum_{i=1}^{p} v_i(x_i) \Delta_i \psi_i(x) = -c \sum_{i=1}^{p} v_i(x_i) \Delta_i \left[ \frac{h_i(x_i)}{b_0 + D^\lambda} \right] \]

\[ = -c \sum_{i=1}^{p} v_i(x_i) \left[ \frac{\Delta_i h_i(x_i)}{b_0 + D^\lambda} - \frac{h_i(x_i - 1) \Delta_i D^\lambda}{(b_0 + D^\lambda)(b_0 + D_i^\lambda)} \right] \]

\[ = c \left[ -\frac{p}{b_0 + D^\lambda} + \frac{1}{b_0 + D^\lambda} \sum_{i=1}^{p} \frac{v_i(x_i) h_i(x_i - 1) \Delta_i D^\lambda}{b_0 + D_i^\lambda} \right]. \]

It follows that

\[ \sum_{i=1}^{p} v_i(x_i) \Delta_i \psi_i(x) \leq c \left[ -\frac{p}{b_0 + D^\lambda} + \frac{1}{b_0 + D^\lambda} \sum_{i=1}^{p} \frac{v_i(x_i) h_i(x_i - 1) \Delta_i D^\lambda}{b_0 + (D_i')^\lambda} \right], \text{ for } x \in \mathcal{A}. \]

(4.1.11)

If it can be proven that for all \( x \in \mathcal{A} \),

\[ \sum_{i=1}^{p} \frac{v_i(x_i) h_i(x_i - 1) \Delta_i D^\lambda}{b_0 + (D_i')^\lambda} \leq \beta', \]  

(4.1.12)

then, together with (4.1.11) and (4.1.6), (4.1.10) follows.

Now, for \( \lambda \geq 1 \), applying mean value theorem, we have

\[ v_i(x_i) h_i(x_i - 1) \Delta_i D^\lambda \leq D^\lambda - 1 v_i(x_i) h_i(x_i - 1) \Delta_i h_i(x_i). \]  

(4.1.13)

If \( \beta_i \geq 1 \),

\[ v_i(x_i) h_i(x_i - 1) \Delta_i h_i(x_i) \leq \beta_i h_i(x_i) h_i(x_i - 1) \]

\[ \leq \beta_i^{\beta_i - 1} h_i(x_i - 1), \]  

(4.1.14)

where the last inequality follows from assumption (i). Similarly for \( \beta_i < 1 \),
\[ v_i(x_i)h_i(x_i-1) + \beta_i^+(x_i) \leq \beta_i^+(x_i-1). \]  \hfill (4.1.15)

Therefore (4.1.14) and (4.1.15) imply that

\[ v_i(x_i)h_i(x_i-1) + \beta_i^+(x_i) \leq \beta_i^+(x_i-1) \leq \beta_i^+(x_i) \leq \beta_i^+(x_i-1). \]  \hfill (4.1.16)

Thus, by (4.1.13) and (4.1.16), it is clear that for \( \lambda \geq 1 \),

\[ v_i(x_i)h_i(x_i-1)\Delta_iD^\lambda \leq \lambda \beta_i^+(x_i) \leq \beta_i^+(x_i-1)D^\lambda. \]  \hfill (4.1.17)

For \( \lambda < 1 \), (4.1.13) again holds if \( D^\lambda \) is replaced by \( (D')^{\lambda-1} \).

Together with (4.1.16), this implies that

\[ v_i(x_i)h_i(x_i-1)\Delta_iD^\lambda \leq \lambda \beta_i^+(x_i) \leq \beta_i^+(x_i-1)(D')^{\lambda-1}. \]  \hfill (4.1.18)

Hence, (4.1.17) and (4.1.18) give

\[ v_i(x_i)h_i(x_i-1)\Delta_iD^\lambda \leq \lambda \beta_i^+(x_i) \leq \beta_i^+(x_i-1)(D')^{\lambda-1}(D/D')(\lambda-1)^+. \]  \hfill (4.1.19)

Summing over all \( i \), we get

\[ \sum_{i=1}^p v_i(x_i)h_i(x_i-1)\Delta_iD^\lambda \leq \sum_{1 \leq i \leq p} \beta_i^+(x_i) \leq \sum_{1 \leq i \leq p} \beta_i^+(x_i-1)(D')^{\lambda-1}(D/D')(\lambda-1)^+. \]

By assumption (i),

\[ D/D' \leq \max_{1 \leq i \leq p} U_i, \]  \hfill (4.1.20)

which, together with (4.1.19), implies (4.1.12).

The above theorem appears complicated at first glance, especially because of \( \beta^+ \). In most situations, however, the limit of \( h_i(x_i+1)/h_i(x_i) \) is one as \( x_i \to \infty \). Then \( \beta^+ \) can be chosen as close to \( \lambda \max_{1 \leq j \leq p} \beta_j \) as one wishes. Solutions to \( \beta^+(\phi) \leq 0 \) can hence be found.
if $p > \lambda \max_{1 \leq j \leq p} \beta_j$ and (4.1.6) is satisfied for some nonnegative constant $b_0$ and some positive constants $K$, $\beta_i$, and $\lambda$. This is stated in the following corollary.

Corollary 4.1.1. Let $\mathcal{A}$, $v_i, w_i, \alpha_i, h_i$ and $D$ be as in Theorem 4.1.

Assume that $h_i(x_i)/h_i(x_i-1)$ approaches 1 as $x_i \rightarrow \infty$, and that (4.1.6) is satisfied for some nonnegative constant $b_0$ and some positive constants $K, \beta_i$, and $\lambda$. Suppose $p > \lambda \max_{1 \leq j \leq p} \beta_j$, then there is a solution to $\mathcal{A}'(\hat{\phi}(x)) \leq 0$ with strict inequality for sufficiently large $x$.

Indeed, let $U > 1$ be any number such that

$$p > \lambda \left( \max_{1 \leq i \leq p} \beta_i \right)^+ \left( \max_{1 \leq i \leq p} \beta_i (\lambda-1)^+ \right).$$

(4.1.21)

Denote the expression on the right hand side of the inequality by $\beta'$. Furthermore let $\alpha_i' = \alpha_i + 1$ be such that $h_i(x_i)/h_i(x_i-1) < U$ for all $x_i \geq \alpha_i'$. For any constant $c$, $0 < c < (p-\beta')/K$ and $A = \{(x_1, \ldots, x_p): x_i \geq \alpha_i\}$, $\phi_1, \ldots, \phi_p$ given in (4.1.7) is a solution to $\mathcal{A}'(\hat{\phi}(x)) \leq 0$ with strict inequality holding for sufficiently large $x$.

Proof. There certainly exists $U > 1$, such that (4.1.21) is satisfied, since $\beta'$ approaches $\lambda \max_{1 \leq i \leq p} \beta_i$ as $U \rightarrow 1$, and $p > \lambda \max_{1 \leq i \leq p} \beta_i$. Theorem 4.1, then completes the proof. Q.E.D.

An application of Corollary 4.1.1 will be seen in the next theorem.

Now assume that $\mathcal{A}$, $f_i(x_i|\theta_i)$ and $t_i$ are as in Lemma 2.2. For any function $g(x)$ defined on $\mathbb{R}$, let $\Delta g(x)$ denote $g(x)-g(x-1)$. Under the loss function $L_m$, the following theorem describes how Corollary 4.1.1 and Theorem 2.1 can be applied to improve upon a complicated
estimator \( \hat{\delta}_i^0(X) \) of the form (componentwise)

\[
\delta_i^0(X) = \frac{t_i(X_i-1)}{t_i(X_i)} s_i(X_i+m_i) \left( 1 - \frac{s_0 B_i H_i^{-1} (x_i+m_i) \Delta_i H_i (x_i+m_i) \cdots H_j(j+m_i) \gamma_{ij}}{p B_j H_j(j+m_i) \gamma_{ij}} \right) +
\]

(4.1.22)

where \( s_0, B_i \) are some constants and \( s_i \) and \( H_i \) are some functions defined on the set of all integers.

One of the major problems encountered in the proof of Theorem 4.2, is that of choosing the \( q_i \) functions so that (2.1.13) is satisfied and the difference inequality \( \hat{\delta}_i^0(X) \leq 0 \) (see (2.1.14)) will be of form (2.1.18). Here we will describe a heuristic argument which guides us to such a choice. The \( q_i \)'s which satisfy

\[
q_i(x_i+1)e_i = q_i'(x_i)Q(x)i.
\]

will be considered, since these \( q_i \)'s will lead us to the difference inequality of the form (2.1.18). Assuming that \( \phi_i(x) \leq 0, i=1, \ldots, p \), it can be seen that (2.1.13) will be satisfied for such \( q_i \) if

\[
\frac{q_i'(x_i+1)}{q_i'(x_i)} \geq \frac{1}{s_i(x_i)}
\]

and

\[
\frac{Q(x+e_i)}{Q(x)} > \left( 1 - \frac{s_0 B_i H_i^{-1} \Delta_i H_i(x_i)}{p B_j H_j(j+\gamma_{ij})} \right)^+ \quad (4.1.24)
\]

It is relatively easy to construct \( q_i \) of the desired form which satisfy (4.1.23). To choose \( Q \), let

\[
D_0 = \sum_{j=1}^p B_j H_j(j+\gamma_{ij}).
\]

(4.1.25)
Under the conditions in Theorem 4.2, (4.1.24) is approximately equivalent to
\[
\frac{\Delta_i Q(x + e_i)}{Q(x)} \geq \frac{\varepsilon_0 \Delta_i D_0}{D_0}
\]  

which suggests seeking a solution to the differential equation
\[
\frac{dQ}{Q} = \frac{\varepsilon_0 dD_0}{D_0}.
\]

The suggested choice of \( Q \) is thus \( D_0 \). In Theorem 4.2, \( Q \) is chosen to be \( D_0 \), where \( \varepsilon_1 > \varepsilon_0 \), and it is shown that (4.1.24) is satisfied for sufficiently large \( x \). Modifying \( Q \) so that it takes zero value when the \( x_i \) are small, \( Q \) is thus shown to satisfy (4.1.24).

In the following, define, for any function \( g \),
\[
\Pi_{\mu=0}^{x} g(u) = g(0)...g(x) \quad x = 0, 1, ...
\]
\[
= 1 \quad x = -1, -2, ...
\]

Theorem 4.2 Let \( X_i, f_i(x_i | \theta_i) \) and \( t_i \) be as in Lemma 2.2. Consider the loss function \( L_m, m = (m_1, ..., m_p) \), and \( \delta^0 \) defined in (4.1.22) with \( B_i > 0 \) for all \( i \). Assume that for \( i = 1, ..., p \),

(i) \( \mathbb{E}_0 (\delta^0_i(x))^2 < \infty \),

(ii) \( H_i(x_i) \rightarrow \infty \) as \( x_i \rightarrow \infty \), \( H_i(x_i) \geq 0 \) for all \( x_i \) and \( H_i(x_i) > 0 \) for \( x_i > 0 \).

(iii) \( \frac{H_i(x_i+1)}{H_i(x_i)} \rightarrow 1 \) as \( x_i \rightarrow \infty \),

(iv) for \( x_i > 0 \) \( H_i(x_i) \) is strictly increasing and

\[
\frac{\Delta H_i(x_i+1)}{\Delta H_i(x_i)} \rightarrow 1 \quad \text{as} \quad x_i \rightarrow \infty,
\]

and
(v) \( s_i(x_i) \geq 0 \) for all \( x_i \) and \( s_i(x_i) > 0 \) for \( x_i > 0 \).

Define
\[
h_i(x_i) = \begin{cases} 
\frac{x_i^{k-1}}{t_i^{k-1}(k-m_i-1)} \prod_{u=0}^{k-1} s_i(u) & \text{if } x_i \geq (m_i+1)^+ \\
0 & \text{otherwise} 
\end{cases} 
\]
(4.1.27)

Let \( D_0(x) \) be as in (4.1.25) and \( D_1(x) \) denote \( \frac{p}{i=1} \beta_i h_i(x_i) \), for some positive numbers \( \beta_1, \ldots, \beta_p \). If without ambiguity, \( D_0(x) \) and \( D_1(x) \) will be denoted \( D_0 \) and \( D_1 \). Furthermore define, for some constant \( \beta_i > 0 \)
\[
q_i^0(x) = D_0^1(x+(m_i+1)e_i) = \prod_{u=0}^{x_i+m_i} \frac{1}{s_i(u)} 
\]
and
\[
\phi_i^0(x) = \frac{h_i(x_i+m_{i})}{D_1^1(x+m_i e_i)} 
\]

Assume that the following conditions hold:

(vi) \( E(q_i^0(x)\phi_i^0(x))^2 < \infty \) \( i = 1, \ldots, p \);

(vii) \( h_i(x_i)/h_i(x_i-1) \rightarrow 1 \) as \( x_i \rightarrow \infty \); (4.1.28)

(viii) For some constants \( \lambda \geq 0 \) and \( K > 0 \),
\[
\frac{p}{i=1} h_i^2(x_i) \left[ \frac{t_i(x_i-m_i)}{t_i(x_i)} \left[ \prod_{u=0}^{x_i} \frac{1}{s_i(u)} \right] \right]^2 D_0^1(x)/(0^1 D_1^1) \leq K < \infty. 
\]
(4.1.29)

then if \( p > \lambda \max_{1 \leq i \leq p} \beta_i \), \( \delta^0 \) is inadmissible. Indeed better estimators can be described as following: Let \( A_\alpha = (x_1, \ldots, x_p) : x_i \geq \alpha, i = 1, \ldots, p \) for some number \( \alpha \). For some constants \( \alpha^0 > 1, \alpha^* > 1 \) and \( c > 0 \),

\[
\phi_i^0(x) = q_i^0(x)I_{A_\alpha^0}(x+(m_i+1)e_i) 
\]
(4.1.30)

\[
\phi_i^0(x) = -c\phi_i^0(x)I_{A_\alpha^*}(x+m_i e_i). 
\]
(4.1.31)
Then for some positive constants $c$, small enough, and $\alpha^0$ and $\alpha^*$, big enough, $\delta^0_\iota$ is dominated by $\delta^*$ with $\delta^\iota_\iota = \delta^0_\iota + q_\iota(x)\phi_\iota(x), \iota = 1, \ldots, p$.

Proof: We will verify the conditions of Theorem 2.1. To verify condition (ii) of Theorem 2.1, assume for a moment that $\phi_i(x) \leq 0$, $i = 1, \ldots, p$. (This will be seen to be true later). We first show that there exists $\alpha^0 > 1$ such that $q_\iota$ in (4.1.30) satisfies (2.1.13).

Note that

$$q_\iota(x - (m_\iota + 1)e_\iota) = \frac{\delta_\iota}{D_0} \prod_{\mu=0}^{\delta_\iota} \frac{1}{s_\iota(u)} I_{\alpha_\iota 0}(x).$$  \hspace{1cm} (4.1.32)

It is sufficient to show that

$$\delta^0_\iota(x - m_\iota e_\iota)q_\iota(x - m_\iota e_\iota - t_\iota(x - m_\iota - 1)q_\iota(x - (m_\iota + 1)e_\iota) \geq 0 \hspace{1cm} (4.1.33)$$

plugging $\delta^0_\iota$ and $q_\iota$ into (4.1.33), it is clear that (4.1.33) will follow from the inequality

$$\left(1 - \frac{\delta_\iota}{\alpha_0} \frac{B_\iota H_\iota(x_\iota) \Delta H_\iota(x_\iota)}{D_0} \right) \prod_{\mu=0}^{\delta_\iota} \frac{1}{s_\iota(u)} I_{\alpha_\iota 0}(x + e_\iota) \geq \frac{\delta_\iota}{D_0} \prod_{\mu=0}^{\delta_\iota} \frac{1}{s_\iota(u)} I_{\alpha_\iota 0}(x).$$  \hspace{1cm} (4.1.34)

Since

$$I_{\alpha_\iota 0}(x + e_\iota) \geq I_{\alpha_\iota 0}(x),$$

it is sufficient to choose $\alpha^0$ such that

$$\frac{\delta_\iota}{D_0} \prod_{\mu=0}^{\delta_\iota} \frac{1}{s_\iota(u)} I_{\alpha_\iota 0}(x) \geq \frac{\delta_\iota}{D_0} I_{\alpha_\iota 0}(x).$$  \hspace{1cm} (4.1.35)

Note that, for sufficiently large $x$,

$$1 - \frac{\delta_\iota}{\alpha_0} \frac{B_\iota H_\iota(x_\iota) \Delta H_\iota(x_\iota)}{D_0} > 0. \hspace{1cm} (4.1.36)$$
This follows from the observations that

\[
0 \leq \frac{B_i^{-1}}{\prod_{j=1}^{i} \frac{H_j(x_j)}{H_i(x_i)}} \leq \frac{\Delta H_i(x_i)}{H_i(x_i)} = 1 - \frac{H_i(x_i - 1)}{H_i(x_i)}
\]

and (by condition (iii))

\[
\lim_{x_i \to \infty} [1 - \frac{H_i(x_i - 1)}{H_i(x_i)}] = 0.
\]

Now let \( \alpha^1 > 1 \) be the number such that \( x \in A_{\alpha^1} \), implies that

\[
B_i^{-1}(x_i) - \frac{\varepsilon_0 B_i H_i^{-1}(x_i) \Delta H_i(x_i)}{1 - \varepsilon_0 B_i H_i^{-1}(x_i) \Delta H_i(x_i)} > 0.
\]

Clearly, (4.1.36) holds for \( x \in A_{\alpha^1} \). To choose \( \alpha^0 \) so that (4.1.34) is satisfied, it is only necessary to choose \( \alpha^0 \geq \alpha^1 \) so that, for \( x \in A_{\alpha^0} \) and \( i = 1, \ldots, p \),

\[
\frac{D_0(x+e_i)}{D_0} \geq \frac{B_i^{-1}}{1 - \frac{\varepsilon_0 B_i H_i^{-1}(x_i) \Delta H_i(x_i)}{D_0}}
\]

or equivalently

\[
\frac{\Delta_i D_0(x+e_i)}{D_0} \geq \frac{\varepsilon_0 B_i H_i^{-1}(x_i) \Delta_i H_i(x_i)}{D_0 - \frac{\varepsilon_0 B_i H_i^{-1}(x_i) \Delta H_i(x_i)}{D_0}}.
\]

By mean value theorem, there exists a \( D^*_i, D_0 < D^*_i < D_0(x+e_i) \), so that

\[
\Delta_i D_0(x+e_i) = \varepsilon_0 (D^*_i) \Delta_i H_i(x_i + 1).
\]
Hence (4.1.38) is equivalent to
\[
\varepsilon_1 (D_i^*) \frac{{\varepsilon_7 - 1}}{{\Delta H_i^1(x_i + 1)/D_0}} \geq \frac{{\varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}}{D_0 - \varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}
\] (4.1.40)

Now we must separately consider three cases (a) \( \varepsilon_1 \geq 0 \geq \varepsilon_0 \),
(b) \( \varepsilon_1 > \varepsilon_0 > 0 \) and (c) \( 0 > \varepsilon_1 > \varepsilon_0 \). Case (a) is trivial, since by choosing \( \alpha^0 \) equal to \( \alpha^1 \), the left hand side of (4.1.38) is nonnegative and the right hand side is nonpositive.

For case (b), (4.1.40) is equivalent to
\[
R_i \geq 1,
\]
where
\[
R_i = \frac{{\varepsilon_1 (D_i^*) \frac{{\varepsilon_7 - 1}}{{\Delta H_i^1(x_i + 1)/D_0}}}}{{\varepsilon_0 \frac{{\varepsilon_7 - 1}}{{\Delta H_i^1(x_i + 1)/D_0}}}} \frac{{\varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}}{D_0 - \varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}
\] (4.1.41)

Clearly,
\[
R_i = \frac{{\varepsilon_1 \left( \frac{{D_i^* - \varepsilon_1 H_i^1(x_i + 1)}}{D_i^*} \right)}}{\varepsilon_0 \left( \frac{{D_i^* - \varepsilon_1 H_i^1(x_i + 1)}}{D_i^*} \right)} \frac{{\varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}}{D_0 - \varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)} \frac{{\Delta H_i^1(x_i + 1)}}{B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}
\]
\[
\geq \frac{{\varepsilon_1 \left( \frac{{D_i^* - \varepsilon_1 H_i^1(x_i + 1)}}{D_i^*} \right)}}{\varepsilon_0 \left( \frac{{D_i^* - \varepsilon_1 H_i^1(x_i + 1)}}{D_i^*} \right)} \frac{{\varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}}{D_0 - \varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)} \frac{{\Delta H_i^1(x_i + 1)}}{B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}
\]
\[
\geq \frac{{\varepsilon_1 \left( \frac{{D_0 (1-\varepsilon_1)}}{D_0 (x+e_i)} \right)}}{\varepsilon_0 \left( \frac{{D_0 (1-\varepsilon_1)}}{D_0 (x+e_i)} \right)} \frac{{\varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}}{D_0 - \varepsilon_0 B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)} \frac{{\Delta H_i^1(x_i + 1)}}{B_i^1 H_i^1(x_i) \Delta H_i^1(x_i)}
\]
\[
\times \left( \frac{{H_i^1(x_i)}}{H_i^1(x_i + 1)} \right)^{1-\varepsilon_1} \frac{{\Delta H_i^1(x_i + 1)}}{\Delta H_i^1(x_i)}
\]
\[
\geq \left( \frac{H_i(x_i)}{H_i(x_i + 1)} \right)^{1-\epsilon_i} + \frac{B_i(x_i - x_0) B_i^{-1}}{H_i(x_i) - B_i x_i + 1} \frac{\Delta H_i(x_i)}{H_i(x_i)}
\]

\[
x \left( \frac{H_i(x_i)}{H_i(x_i + 1)} \right)^{1-\epsilon_i} \frac{\Delta H_i(x_i + 1)}{\Delta H_i(x_i)}
\]

By conditions (iii) and (iv), the last expression approaches \( \frac{\epsilon_i}{\epsilon_0} > 1 \) as \( x_i \to \infty \). Therefore there exists \( a^0 > a^1 \), such that \( R_i \geq 1 \), \( i = 1, \ldots, p \) if \( \chi \in A_q \).

For case (c), (4.1.40) is equivalent to

\[
R_i < 1.
\]

Now from (4.1.41),

\[
R_i \leq \frac{\epsilon_i}{\epsilon_0} \frac{D_0 - \epsilon_0 B_i H_i}{D_0} \frac{B_i^{-1}}{x_i} \frac{\Delta H_i(x_i)}{H_i(x_i)} \frac{H_i(x_i + 1)}{B_i - 1} \frac{\Delta_i H_i(x_i + 1)}{\Delta_i H_i(x_i)}.
\]

Again by conditions (iii) and (iv), the last expression approaches \( \frac{\epsilon_i}{\epsilon_0} < 1 \) as \( x_i \to \infty \). Therefore there exist \( a^0 > a^1 \) such that \( \chi \in A_q \) implies that \( R_i \leq 1 \) for all \( i \). In conclusion (4.1.37), and hence (4.1.33) are satisfied if \( q_i \) is defined as in (4.1.30). For convenience in the following, it is also assumed that \( a^0 > (m_i + 1)^+ \), \( i = 1, \ldots, p \).

To satisfy condition (iv) of Theorem 2.1, it is necessary to solve the inequality

\[
\epsilon_i = \frac{p}{i=1} \epsilon_i \leq 0 \quad \text{or} \quad \frac{1}{2} \sum_{i=1}^{p} \epsilon_i \leq 0.
\]

where
\[ s'_i = \frac{2t_i (x_i - m_i - 1)}{t_i(x_i)} q_i (x - (m_i + 1)e_i) \Delta_i \phi_i (x - m_i e_i) \]

\[ + \frac{t_i (x_i - m_i)^2}{t_i(x_i)} q_i (x - m_i e_i) \phi_i^2 (x - m_i e_i). \]

Clearly,

\[ \frac{1}{2} s' = D_0^{\hat{\gamma}_1} I_{\alpha_0}^{s''}. \quad (4.1.43) \]

where

\[ s'' = \sum_{i=1}^{p} \frac{t_i (x_i - m_i - 1)}{t_i(x_i)} \frac{1}{s(u)} \Delta_i \phi_i (x - m_i e_i) \]

\[ + \sum_{i=1}^{p} \frac{2 z_i}{2} D_0^{\hat{\gamma}_1 \frac{1}{2}} \frac{t_i (x_i - m_i)}{t_i(x_i)} \frac{1}{s(u)} \phi_i^2 (x - m_i e_i). \quad (4.1.44) \]

Of course, a solution to \( I_{\alpha_0} s'' \leq 0 \) is also a solution to \( s' \leq 0 \).

Note that

\[ 1 \leq \left[ \frac{D_0 (x + e_i)}{D_0 (x)} \right] \leq \left[ \frac{H_i (x_i + 1)}{B_i (x_i)} \right], \]

for all \( x \) such that \( H_i (x_i) \neq 0 \). By condition (iii), there exists \( K_1 \) so that

\[ \frac{B_i}{H_i (x_i + 1)} \leq K_1 \]

for all \( x_i > 0 \).

Thus there exists \( K_2 \) so that if \( x_i > 0, \ i = 1, \ldots, p \), then

\[ \frac{2 s_i}{D_0 (x_i + e_i) / D_0} \leq K_2 D_0. \quad (4.1.45) \]
Therefore, by (4.1.44) and (4.1.45),

$$I_{A^0_D} g'' = I_{A^0_D} g''$$

where

$$g'' = \sum_{i=1}^{p} \frac{t_i(x_i-m_i-1)}{t_i(x_i)} \frac{x_i^{-1}}{u=0} \Delta_i^u \phi_i(x_m e_i)$$

$$+ \sum_{i=1}^{p} \frac{K_2}{2} D_0 \frac{\phi_i(x_m e_i)}{D_1^x} \frac{x_i^{-1}}{u=0} \frac{1}{s^2(u)} \phi_i(x_m e_i).$$

A solution to $g'' \leq 0$ is certainly a solution to $g' \leq 0$. Let $h_i$ be as in (4.1.27). By Corollary 4.1.1, (4.1.28) and (4.1.29), there exists a solution to $g'' \leq 0$; namely $\phi_1, \ldots, \phi_p$, where

$$\phi_i(x_m e_i) = \frac{-c h_i(x_i)}{D_1^x} I_{A^0_D}(x)$$

for some constants $c > 0$ and $\alpha^* > \alpha^0$. Furthermore, for such $\phi_i$'s,

$$g'' \phi(x_m e_i) < 0$$

for sufficiently large $x$. By Theorem 2.1, the proof is now complete.

Q.E.D.

Corollary 4.2.1. Assume $X_i \sim \text{indep. Po}(\theta_i), i = 1, \ldots, p$. Under the loss function $L_1$, the estimator $\hat{\delta}^C$ of Clevenson's type given component-wise by

$$\hat{\delta}^C(X) = \left(1 - \frac{\theta_0}{p-1+i} \right) X_i$$

(4.1.46)

is inadmissible if $\theta_0 < p-1$.

Proof: To apply Theorem 4.2, let $\hat{\delta}_i^0$ be as in (4.1.22) with $m_i = -1$, $s_i(x_i) = 1$, $B_i = 1$ and $H_i(x_i) = 1 + x_i$ if $x_i \geq 0$ and $H_i(x_i) = 0$ if $x_i < 1$. Clearly $\hat{\delta}_i^0 = \hat{\delta}_i^C$ and conditions (i) through (v) of Theorem 4.2...
are satisfied. Let $h(x_i) = \sum_{k=0}^{x_i} 1 = x_i + 1$, as in (4.1.27). Obviously condition (vii) of Theorem 4.2 is satisfied. Let $\varepsilon_1$ be any number such that if fit $\varepsilon_0 < \varepsilon_1 < p - 1$. Furthermore let

$$D_0 = \frac{1}{\sum_{j=1}^{p} h_j(x_j)} = \frac{1}{\sum_{j=1}^{p} h_j(x_j)} = D_1.$$

To check condition (viii), note that

$$\frac{1}{\sum_{i=1}^{p} h_i^2(x_i)} \frac{t_i(x_i^{x_i+1})}{t_i(x_i)} \frac{D_0^\lambda}{D_1^\lambda}$$

$$= \frac{x_i^{x_i+1}}{D_0^\lambda}$$

$$\leq K,$$

for some constant $K$ if $\lambda = (x_i^{x_i+1})^+$. Hence condition (viii) is satisfied with $\lambda = (x_i^{x_i+1})^+$. Since $\varepsilon_1 < p - 1$, $(x_i^{x_i+1})^+ < p$ and consequently, $p > \lambda$. Finally, by (4.1.30) and (4.1.31), $q_i \phi_i$ is bounded. Hence $E q_i^2(x) \phi_i^2(x) < \infty$ and condition (vi) is satisfied. By Theorem 4.2, $\theta^C$ is thus inadmissible.

Corollary 4.2.2. Assume $X_i \mathcal{I} \text{N} \text{D}_{\text{D}(\theta_i)}, i = 1, \ldots, p$. Under the loss function $L_0$, the estimator $\hat{\theta}_P$ of Peng's type similar to (1.3.7) given componentwise by

$$\hat{\theta}_P(x) = \left( x_i - \frac{h(X_i)}{\sum_{j=1}^{p} h_j^2(x_j)} \right)^+, \quad (4.1.47)$$

is inadmissible if $\varepsilon < p - 2$. Recall that $h(\cdot)$ was defined by

$$h(x_i) = \sum_{K=1}^{x_i} \frac{1}{K} \quad x_i = 1, 2, \ldots$$

$$= 0 \quad x_i < 1.$$
Proof: To apply Theorem 4.2, here let \( \delta_i^0 \) be as in 4.1.22 with \( m_i = 0, s_i(x_i) = 1, B_i = 2 \) and \( H_i(x_i) = h(x_i) \). Clearly \( \delta_i^0 = \delta_i^0 \) if \( \varepsilon_0 = \varepsilon/2 \). It is straightforward to show that condition (i) through (v) of Theorem 4.2 are satisfied. Let \( h_i(x_i) \) be defined as in (4.1.27), then \( h_i(x_i) = h(x_i) \). Again condition (vii) is satisfied. Furthermore let \( b_i = 2 \) and

\[
D_0 = \prod_{j=1}^p h^2(x_j) = D_1.
\]

To check condition (viii), let \( \varepsilon_1 \) be any number so that \( \varepsilon_0 < \varepsilon_1 < \frac{p-2}{2} \).

Now

\[
\prod_{i=1}^p h^2(x_i) \frac{t_i(x_i-m_i)}{t_i(x_i)} \frac{\varepsilon_1}{D_0} = \frac{\varepsilon_1+1}{D_0} \leq K
\]

for some constant \( K \) if \( \lambda = (\varepsilon_1+1)^+ \). Note \( p > 2\lambda \). Finally by (4.1.30) and (4.1.31), \( q_i(x)\phi_i(x) \) is bounded, hence Eq. (4.1.26) is satisfied, and condition (vi) is satisfied. By Theorem 4.2, the proof is complete. Q.E.D.

The proofs of the following two corollaries are similar to those of the previous corollaries, and so are omitted.

Corollary 4.2.3. Assume \( X_i \overset{\text{indep.}}{\sim} \text{Po}(\theta_i) \) \( i = 1, \ldots, p \). Under the loss function \( L_m, m = -n \) and \( n \) being a positive integer, the estimator \( \delta_i^T \) of Tsui's type given componentwise by

\[
\delta_i^T(x) = \left( x_i - \frac{\varepsilon h(x_i-n)}{\prod_{j=1}^p h(x_j-nY_{ij})} \right)^+,
\]

is inadmissible if \( \varepsilon < n(p-1) \). Recall that \( h(\cdot) \) was defined as

\[
h(x_i) = (x_i+1) \cdots (x_i+n) \quad \text{if} \quad x_i \geq 0
\]

\[
= 0 \quad \text{otherwise}.
\]
Corollary 4.2.4. Assume that \( X_i \overset{\text{indep.}}{\sim} \text{NB}(r_i, \theta_i), i = 1, \ldots, p \). Let \( \hat{\theta}_{\text{NB}}^i \) be the estimator of \( \theta \) given (componentwise) by

\[
\hat{\theta}_{\text{NB}}^i(x) = \left( \frac{X_i}{r_i - 1 + X_i} - \frac{2h(X_i)}{\sum_{j=1}^{p} h^2(X_j)} \right)^+
\]

where

\[
h(x_i) = \sum_{k=1}^{x_i} \frac{r_i - 1 + k}{k} \quad x_i \geq 1
\]

\[
h(x_i) = 0 \quad \text{otherwise.}
\]

Under the loss function \( L_0 \), \( \hat{\theta}_{\text{NB}}^i \) is inadmissible if \( \varepsilon < p-2 \).

Note that all the inadmissible estimators stated in the above corollaries are dominated by a corresponding \( \hat{\theta}^* \) given in Theorem 4.2. The correction terms \( q_i(\hat{\theta})_i^2(\hat{\theta}) \) (See (4.1.30), (4.1.31)) are nonpositive and are obtained by applying Theorem 2.1. These facts will be used in the next section.

Section 4.2. "Upper Bounds" on the Class of Admissible Estimators

As seen in the last section, to improve upon an estimator of a general form, quite complicated calculations are generally involved. In this section, a theorem is developed by which a broad class of estimators can be shown to be inadmissible. This theorem also shows what is meant by an "upper bound" on the class of admissible estimators. Again, we will use \( I_A(x) \) to denote the indicator function, i.e.

\[
I_A(x) = 1 \quad \text{if } x \in A
\]

\[
= 0 \quad \text{otherwise.}
\]

Theorem 4.3 Let \( \hat{\theta} \) be as in Lemma 2.2. Consider two estimators \( \hat{\theta}^0 \) and \( \hat{\theta}^1 \) of \( \hat{\theta} \), both of which have nonnegative components. Assume
E[\delta_{i}^{1}(X)]^2 < \infty \text{ for } i = 1, \ldots, p. \text{ Let } A = \{(x_1, \ldots, x_p) : x_i \geq \alpha, i = 1, \ldots, p\} \text{ for some number } \alpha. \text{ Under the loss function } L_m, \text{ suppose the following two conditions hold:}

(i) \delta_{i}^{1}(x) \geq \delta_{i}^{0}(x) \text{ for all } x \in A \text{ and } i = 1, \ldots, p;

(ii) \delta_{i}^{0}(x) \text{ can be improved by the procedure of Theorem 2.1 with the correction terms } q_i(X)\phi_i(X) \text{ being nonpositive.}

Then, the estimator \delta^{*} with ith component

\[ \delta_{i}^{*}(X) = \delta_{i}^{1}(X) + I_{S}(X+m_{i}e_i)q_i(X)\phi_i(X), \]

is as good as \delta_{i}^{1}, \text{ where } S = \{x : x_i \geq \alpha + m_i, i = 1, \ldots, p\}.

Proof: \text{ We will use Corollary 2.1.1. Thus, assuming conditions (i), (ii) and (iii) of Theorem 2.1 and condition (iv)' of Corollary 2.1.1, we need to show that these four conditions are also satisfied with } \delta_{i}^{0} \text{ and } \phi_i \text{ being replaced by } \delta_{i}^{1}(x) \text{ and } \phi_i(x)I_{S}(x+m_{i}e_i), \text{ respectively. (}q_i \text{ remains unchanged).}

Clearly, conditions (i) and (ii) of Theorem 2.1 are satisfied.

Note that condition (iii) in this situation has the form

\[ (\delta_{i}^{1}(x-m_{i}e_i) - t_i(x_{i-1}e_i)q_i(x-m_{i}e_i) - t_i(x_{i-1}e_i)q_i(x-(m_{i}+1)e_i)I_{S}(x)) \leq 0 \] (4.2.1)

To verify (4.2.1), observe that

\[ x \in S \Rightarrow (x-m_{i}e_i) \in A. \]

Consequently, condition (i) of this theorem implies that

\[ \delta_{i}^{1}(x-m_{i}e_i) \geq \delta_{i}^{0}(x-m_{i}e_i), \quad i = 1, \ldots, p. \] (4.2.2)

By (2.1.13), (4.2.2), and condition (ii) of this theorem, (4.2.1) is established. Furthermore, from condition (iv)' of Corollary 2.1.1, we have
Lemma 4.1 implies that \( \phi_i(x)I_S(x+m_i\xi_i), \ i = 1, \ldots, p, \) also satisfies (4.2.3). Therefore, condition (iv) of Corollary 2.1.1 is satisfied.

Q.E.D.

Corollary 4.3.1 Let the notation and assumptions be as in Theorem 4.3. If \( I_S(x+m_i\xi_i)q_i(x)\phi_i(x) \) is not identically zero, then \( \delta^1 \) is inadmissible.

Proof: Clearly \( \delta^* \) is as good as \( \delta^1 \), and \( \delta^* \vdash \delta^1 \). Because \( L_{\delta^1} \) is a strictly convex loss,

\[
R(\theta, \frac{1}{2}\delta^* + \frac{1}{2}\delta^1) < \frac{1}{2} R(\theta, \delta^*) + \frac{1}{2} R(\theta, \delta^1) \leq R(\theta, \delta^1).
\]

Q.E.D.

If the correction term \( q_i(x)\phi_i(x) \) is nonzero for all sufficiently large \( x \), Corollary 4.3.1 implies that any estimator \( \delta^1 \), with \( \delta^1(x) \geq \delta^0(x) \) for sufficiently large \( x \), is inadmissible. In this sense, \( \delta^0 \) is an upper bound on the class of admissible estimators. The following corollaries follow immediately from Theorem 4.2 and Corollary 4.3.1.

Corollary 4.3.2. Suppose that the assumptions in Theorem 4.2 hold. Under the loss function \( L_{\delta^0} \), \( \delta^0 \) is an upper bound on the class of admissible estimators, if \( p > \lambda \max_{1 \leq j \leq p} \beta_j \).

Corollary 4.3.3. Assume \( X_i \text{iid} \text{Po}(\theta_i), i = 1, \ldots, p \). Under the loss function \( L_{\delta^c} \) (given in (4.1.46)), with \( \kappa_0 < p-1 \) is an upper bound on the class of admissible estimators.
Corollary 4.3.4. Assume $X_i \overset{\text{indep}}{\sim} \text{Po}(\theta_i)$, $i = 1, \ldots, p$. Under the loss function $L_0$, $\hat{\delta}_i^p$ (given in (4.1.47)), with $\varepsilon < p-2$, is an upper bound on the class of admissible estimators.

Corollary 4.3.5. Assume $X_i \overset{\text{indep}}{\sim} \text{Po}(\theta_i)$, $i = 1, \ldots, p$. Under the loss function $L_{-n}$, $n = 1, 2, \ldots$, $\hat{\delta}_i^T$ (given in (4.1.48)), with $\varepsilon < n(p-1)$, is an upper bound on the class of admissible estimators.

Corollary 4.3.6. Assume $X_i \overset{\text{indep}}{\sim} \text{NB}(r_i, \theta_i)$, $i = 1, \ldots, p$. Under the loss function $L_0$, $\hat{\delta}_i^{\text{NB}}$ (given in (4.1.49)), with $\varepsilon < p-2$, is an upper bound on the class of admissible estimators.

There are many possible applications of these corollaries. Only two examples will be given here.

Example 4.1. Assume $X_i \overset{\text{indep}}{\sim} \text{Po}(\theta_i)$, $i = 1, \ldots, p$. Under the loss function $L_{-1}$, the estimator $\hat{\delta}_i^*$, with

$$\hat{\delta}_i^*(x) = \left( \frac{\sum_{j=1}^{p} X_j}{\beta + p - 1 + \sum_{j=1}^{p} X_j} \right) X_i,$$

was conjectured to be inadmissible for $\beta < 0$ in Brown (1974). To prove this, let $\varepsilon_0$ be any number such that $\beta + p - 1 < \varepsilon_0 < p - 1$. For such $\varepsilon_0$, compare $\delta_i^C(x)$ (given in (4.1.46)) to $\hat{\delta}_i^*(x)$. Clearly, for sufficiently large $x$,

$$\delta_i^*(x) \geq \delta_i^C(x) \quad (4.2.4)$$

$$\Rightarrow \left( 1 - \frac{\beta + p - 1}{\beta + p - 1 + \sum_{j=1}^{p} X_j} \right) X_i \geq \left( 1 - \frac{\varepsilon_0}{p - 1 + \sum_{j=1}^{p} X_j} \right) X_i$$

$$\Rightarrow \frac{\beta + p - 1}{\beta + p - 1 + \sum_{j=1}^{p} X_j} \leq \frac{\varepsilon_0}{p - 1 + \sum_{j=1}^{p} X_j}.$$
Since the last inequality holds for sufficiently large \( x \), Corollary 4.3.3 and (4.2.4) imply that \( \delta^* \) is inadmissible. This proves Brown's conjecture.

Example 4.2. Assume \( X_i \overset{\text{iid}}{\sim} \text{NB}(r, \theta_i), i = 1, \ldots, p \). Recall from (1.3.9) that Hudson's estimator \( \delta^H_i \) is

\[
\delta^H_i(x) = \frac{X_i}{r-1+x_i} - \frac{(\#(X)-3)^+ h(x_i)}{\sum_{j=1}^{p} h^2(x_j)}, \quad i = 1, \ldots, p,
\]

where \( \#(X) \) is the number of indices \( j \) for which \( x_j > 1 \), and

\[
h(x_i) = \sum_{k=1}^{x_i} \frac{(r-1+k)}{k} \quad \text{if} \quad x_i > 0, \quad \text{while} \quad h(x_i) = 0 \quad \text{otherwise}.
\]

We claim that for \( p > 3, \delta^H_i \) is inadmissible. To see this, let \( \delta^\text{NB}_i \) be as in Corollary 4.3.6 with \( \varepsilon = p-2.5 \). For sufficiently large \( x \),

\[
\delta^\text{NB}_i(x) \leq \delta^H_i(x) \tag{4.2.5}
\]

\[
= \frac{X_i}{r-1+x_i} - \frac{\varepsilon h(x_i)}{\sum_{j=1}^{p} h^2(x_j)} \leq \frac{X_i}{r-1+x_i} - \frac{(p-3)h(x_i)}{\sum_{j=1}^{p} h^2(x_j)}
\]

\[
= \frac{\varepsilon h(x_i)}{\sum_{j=1}^{p} h^2(x_j)} \geq \frac{(p-3)h(x_i)}{\sum_{j=1}^{p} h^2(x_j)}.
\]

Clearly the last inequality holds. Hence by (4.2.5) and Corollary 4.3.6, \( \delta^H_i \) is inadmissible if \( p \geq 3 \).
CHAPTER V
OTHER RELATED PROBLEMS

In this chapter, some miscellaneous problems which relate to improving upon estimators will be considered. In Section 5.1, the problems of improving upon standard estimators for the parameters of Poisson and Chi-square distributions are compared. The comparison reveals the role played by discreteness.

In Section 5.2, an admissibility problem is discussed. An example is given in Section 5.3, which deals with the simultaneous estimation problem based on three observations having distributions of completely different forms. In Section 5.4, three generalizations are discussed.

Section 5.1 Comparison of the Poisson Case and the Chi-square Case.

Assume that \( X_i \overset{\text{indep.}}{\sim} \text{P}\_0(\alpha_i) \), \( i=1,\ldots,p \). Estimators better than \( \delta^0(\overline{x})=\overline{x} \) under the loss function, \( L_{\overline{m}_i} \), \( m_i \leq 0 \), \( i=1,\ldots,p \), were presented in Section 3.1. But no results have been given for the loss function, \( L_{\overline{m}_i} \), with positive \( m_i \)'s.

To discuss this problem in detail, let us consider the loss
function \( l_m \), where \( m \) is an integer. Again, in trying to improve upon \( \hat{\delta}^0 \), write a competitor as \( \hat{\delta}^* = \hat{\delta}^0 + \phi \) with \( \phi = (\phi_1, \ldots, \phi_p) \). From Theorem 2.1 and (3.1.2), we have \( R(\hat{\delta}^0, \hat{\delta}^*) - R(\hat{\delta}^0, \hat{\delta}^0) = E_\theta \delta^l(\phi) \)

where

\[
\delta^l_m(\phi) = \begin{cases} 
\sum_{i=1}^{p} \left( 2(x_i - m) \ldots x_i \Delta_i \psi_i(x) + (x_i - m + 1) \ldots x_i \psi_i^2(x) \right) & \text{if } m > 0 \quad (5.1.1) \\
\sum_{i=1}^{p} \left( 2x_i \Delta_i \psi_i(x) + \psi_i^2(x) \right) & \text{if } m = 0 \quad (5.1.2) \\
\sum_{i=1}^{p} \left( 2 \Delta_i \psi_i(x) \right) \left( \frac{1}{x_i + 1} \ldots \frac{1}{x_i - m - 1} \right) + \frac{\psi_i^2(x)}{x_i + 1} \ldots \frac{1}{x_i - m - 1} & \text{if } m < 0 \quad (5.1.3)
\end{cases}
\]

and \( \psi_i(x) = \phi_i(x - \phi_i) \). For the case \( m > 0 \), the theorems in Section 2.2 do not yield any nontrivial solutions to \( \delta^l_m(\phi) \leq 0 \).

The problems of the existence of a nontrivial solution to (5.1.1) and of the admissibility of \( \hat{\delta}^0 \) are not yet answered. To gain some insight, however, we compare the difference inequality

\[
\delta^l_m(\phi) \leq 0 \quad (5.1.4)
\]

(\( \delta^l_m \) was given in (5.1.1) through (5.1.3)) to the differential inequality (1.3.10). For convenience, (1.3.10) is restated here:

\[
\delta^l_m(\phi) = \sum_{i=1}^{p} \left( x_i^{m+1} \frac{\partial}{\partial x_i} \phi_i(x) + b_i x_i^2 \right) \leq 0. \quad (1.3.10)
\]

The differential inequality was encountered in trying to improve upon the estimator \( \hat{\delta}^l_2(x/n+2) \) under the loss function \( l_{m-2} \), where \( x_i/\theta_i \overset{\text{iid}}{\sim} \chi_n^2 \), \( i = 1, \ldots, p \) (cf. Section 1.3.3). Inspecting \( \delta^l_m \) and \( \delta^l_m \), we see that, the difference inequality (5.1.4) for the Poisson case is analogous to the differential inequality (1.3.10) for the chi-square case. (The constants \( b_1, \ldots, b_p \) are not significant in determining the form of a solution to (1.3.10).) For \( m \leq 0 \), as seen
in (3.1.15), (3.1.18), (3.1.20) and (1.3.14), the solutions to \( \xi^*_m < 0 \) and \( \xi_m < 0 \) are very similar. The improved estimators given by these solutions all correct the standard estimator by shrinking toward a point. For \( m > 0 \), the solutions to \( \xi_m < 0 \) are all nonnegative and therefore indicate that the new estimator corrects the standard one by pulling away from \( (0, \ldots, 0) \). Due to the similarity between \( \xi_m \) and \( \xi^*_m \), the solution to \( \xi_m < 0 \) seems to suggest that the solutions to \( \xi^*_m < 0 \) (if any) are also nonnegative, and hence that better estimators pull \( \hat{\theta}^0 \) away from \( (0, \ldots, 0) \). However, a theorem can be established in discrete cases which asserts that if \( \hat{\theta}^0 + \hat{\phi} \) is as good as \( \hat{\theta}^0 \), and \( \hat{\phi} = (\phi_1, \ldots, \phi_p) \) is such that \( \phi_i(\chi) > 0 \) all \( \chi \), then all the \( \phi_i \)'s must be zero. In other words, \( \hat{\theta}^0 \) can not be improved by pulling away from the origin. Therefore, these facts seem to suggest the admissibility of \( \hat{\theta}^0 \) under the loss function \( L_m, m > 0 \).

Note that this also explains why the first aspect of Berger's phenomena (i.e. the correction terms changes sign according to the loss function) does not occur in the discrete case. But, clearly, the second aspect is certainly observed (i.e. the dimension needed for inadmissibility of \( \hat{\theta}^0 \) depends on the loss function as shown in Peng (1975), Cleveron and Zidek (1975) and Tsui and Press (1977).). In fact a general theorem asserts that any estimator of \( \theta \) can not be improved by positive correction terms if \( \chi \) is distributed as in Lemma 2.2. This will be reported elsewhere.
Section 5.2. An Admissibility Problem

The main idea used in Chapters II, III, IV to prove inadmissibility has been solving an appropriate difference inequality.

Once a nontrivial solution, which satisfies the regularity conditions, is obtained, the estimator is known to be inadmissible.

It is therefore natural to ask whether the lack of a solution (except the zero solution) to the difference inequality corresponding to a particular \( \delta^0 \), implies that \( \delta^0 \) is admissible. The following example indicates that the conjecture is false for \( p = 1 \).

Example 5.1. Let \( X \) be a one-dimensional random variable having logarithmic distribution, i.e.

\[
P(X = x) = \frac{1}{-\log(1-\theta)} \frac{\theta^x}{x}, \quad x = 1, 2, \ldots,
\]

for some unknown parameter \( \theta \), \( 0 < \theta < 1 \). It is clear that the unbiased estimator \( \delta^0(X) \),

\[
\delta^0(X) = \begin{cases} 
\frac{x}{x-1} & \text{if } x \geq 2 \\
0 & \text{if } x = 1,
\end{cases}
\]

is inadmissible, since it estimates \( \theta \) by some number greater than 1. Thus \( \delta^0(X) \) can certainly be improved if \( X/X-1 \) is replaced by 1 when \( X \geq 2 \). However we will consider the problem of improving upon \( \delta^0 \) by using Theorem 2.1. Under square error loss, the difference inequality (See (2.1.12) and (2.1.14) with \( q_1 = 1 \) ) has the form

\[
\delta' (\phi) = 2v(x)\Delta\phi(x) + \phi^2(x) \leq 0 \tag{5.2.1}
\]

where

\[
v(x) = \begin{cases} 
\frac{x}{x-1} & \text{if } x \geq 2 \\
0 & \text{if } x = 1.
\end{cases}
\]
The following lemma will show that the only solution to (5.2.1) is 
\( \phi(x) = 0 \). Therefore the lack of a nontrivial solution to the difference inequality does not necessarily imply admissibility.

**Lemma 5.1.** If \( \phi(x) \) satisfies (5.2.1), then \( \phi(x) = 0, \ x = 1, 2, \ldots \). 

**Proof:** First, we will show that \( \phi(\cdot) \) is bounded. Clearly \( \phi(\cdot) \) is nonincreasing function, since \( \Delta \phi(x) \) must be nonpositive. Also for \( x = 1 \), (5.2.1) becomes \( \phi^2(1) \leq 0 \), implying \( \phi(1) = 0 \). Thus \( \phi(x) \leq 0 \), for \( x = 1, 2, \ldots \). Since \( v(x) \leq 2 \) for all nonnegative integer \( x \), it follows that

\[
0 \geq \phi'(\phi) \\
\geq 4\phi(x) + \phi^2(x) \\
\geq 4\phi(x) + \phi^2(x)
\]

(5.2.2)

which, since \( \phi(x) \leq 0 \), implies that

\[-4 \leq \phi(x) \leq 0.\]

Next, let \( \varepsilon \) be the limit of \( \phi(x) \) as \( x \to \infty \). (This limit exists, since \( \phi \) is bounded and nonincreasing.) We then complete the proof by showing that \( \varepsilon = 0 \). Now clearly

\[-4 \leq \varepsilon \leq 0.\]

Suppose that \( \varepsilon < 0 \), then there exists some \( N > 0 \) such that \( \phi(x) < 0 \) for all \( x > N \). By (5.2.2),

\[
\frac{4(\phi(x) - \phi(x-1))}{\phi(x)} + \phi(x) \geq 0
\]

(5.2.3)

Letting \( x \) go to infinity, (5.2.3) implies that \( \varepsilon \geq 0 \), which is a contradiction.

Q.E.D.
Section 5.3. An Example of an Improved Simultaneous Estimator
Based on Discrete and Continuous Observations

The theorems in Chapter II were designed for the case when the observations \( X_1, \ldots, X_p \) are independently from discrete exponential families. Of course, the most common situations occur when the distributions of \( X_1, \ldots, X_p \) have the same form, as considered in Chapter III. However, it is also interesting to observe the following example which deals with an estimation problem based on three independent random variables, \( X_1, X_2, \) and \( X_3 \) having distributions of completely different forms.

Example 5.1. Assume that \( X_1, X_2, \) and \( X_3 \) are independent random variables; \( X_1 \sim \text{Po}(\theta_1), \) \( X_2 \sim \text{N}(\theta_2, 1) \) and \( X_3 \sim \theta_3 X_n^2. \) It is desired to estimate \((\theta_1, \theta_2, \theta_3)\) under the loss function \( L_m, \) \( m = (0, 0, -1). \) i.e.

\[
L_m(\hat{\theta}, \theta) = L_1(\theta_1, a_1) + L_2(\theta_2, a_2) + L_3(\theta_3, a_3)
\]

where \( L_1(\theta_1, a_1) = (\theta_1 - a_1)^2, \) \( L_2(\theta_2, a_2) = (\theta_2 - a_2)^2 \) and \( L_3(\theta_3, a_3) = \theta_3^{-1}(\theta_3 - a_3)^2. \) A standard estimator is

\[
\bar{\delta}^0(x) = (\bar{\delta}_1^0(x_1), \bar{\delta}_2^0(x_2), \bar{\delta}_3^0(x_3))
\]

where

\[
\bar{\delta}_1^0(x_i) = x_i, \quad i = 1, 2,
\]

and

\[
\bar{\delta}_3^0(x_3) = x_3/n + 2.
\]

It is known (see Hodges and Lehmann (1951)) that for each coordinate treated separately, \( \bar{\delta}_i^0(x_i) \) is an admissible estimator for \( \theta_i \) under the loss function \( L_i. \) However, \( \bar{\delta}^0 \) is inadmissible under \( L_m \) and can be improved by an argument similar to Stein's technique. (Stein (1973).)
Write a competitor as \( \hat{\delta}^* = (\hat{\delta}^*_1, \hat{\delta}^*_2, \hat{\delta}^*_3) \) with
\[
\hat{\delta}^*_i(x) = x_i + \phi_i(x) \quad i = 1, 2,
\]
and
\[
\hat{\delta}^*_3(x) = \frac{x_3}{n+3} (1 + \phi_3(x)).
\]
(5.3.1)

Under certain regularity conditions on \( \phi_3 \), the identity
\[
\theta^{-1} \left( E_0 (\hat{\delta}^*_3(x) - \theta_3)^2 - E_0 \left( \frac{x_3}{n+2} - \theta_3 \right)^2 \right)
\]
\[
= E_0 \left( \frac{4x_3^2}{(n+2)^2} \frac{\partial \phi_3(x)}{\partial x_3} + \frac{x_3}{n+2} \phi_3(x) + \frac{4x_3^2}{(n+2)^2} \phi_3(x) \frac{\partial \phi_3(x)}{\partial x_3} \right) \quad (5.3.3)
\]
is derived in Berger (1978). If \( \phi_3(x) \geq 0 \) and \( \frac{\partial \phi_3(x)}{\partial x_3} < 0 \), the expression on the right hand side of (5.3.3) is bounded above by
\[
E_0 \left( \frac{4x_3^2}{(n+2)^2} \frac{\partial \phi_3(x)}{\partial x_3} + \frac{x_3}{n+2} \phi_3(x) \right) \quad (5.3.4)
\]

Then together with (1.3.2) and (1.3.5), this implies
\[
R(0, \hat{\delta}^*) - R(0, \delta_0^0) \leq E_0 \left\{ 2x_3 \Delta_3 \phi_1(x) + \phi_3(x) + \frac{\partial \phi_3(x)}{\partial x_3} + \frac{x_3}{n+2} \phi_3(x) \right\}
\]
\[
+ 4 \frac{x_3^2}{(n+2)^2} \frac{\partial \phi_3(x)}{\partial x_3} + \frac{x_3}{n+2} \phi_3(x) \}
\]

under certain regularity conditions on \( \phi_1 \). Let
\[
\mathcal{L}'(\phi(x)) = x_1 \Delta_1 \phi_1(x) + \frac{\partial \phi_2(x)}{\partial x_2} + 2x_3^2 \frac{\partial \phi_3(x)}{\partial x_3}
\]
\[
+ \phi_1^2(x) + \phi_2^2(x) + \frac{x_3}{n+2} \phi_3^2(x)
\]

A solution to \( \mathcal{L}'(\phi) \leq 0 \) can be found by an argument similar to those used in Chapter II. Indeed it can be described as following:
Let
\[ h_1(x_1) = \begin{cases} \frac{1}{k} & \text{if } x_1 = 1, 2, \ldots \\ 0 & \text{if } x_1 \leq 0 \end{cases} \]
\[ h_2(x_2) = x_2 \quad x_2 \in \mathbb{R} \]
and
\[ h_3 = \frac{-(n+2)^2}{2} x_3^{-1} \quad x_3 \in \mathbb{R} \]
and, for some \( b > 0 \), define
\[ D = b + h_1(x_1)h_1(x_1+1) + h_2^2(x_2) + h_3(x_3) \]

Furthermore, let
\[ c(k) = \begin{cases} 0 & k = 0 \\ 2/n+2 & k = 1, 2, \ldots \end{cases} \]

Then
\[ \phi_i(x) = \frac{-c(x_1)h_1(x_1)}{D} \quad i = 1, 2, 3, \quad (5.3.6) \]

is a solution to \( \mathbb{E}(\phi(x)) \leq 0 \) with \( \mathbb{E}(\nabla \phi(x)) < 0 \), for all \( \theta \). This solution also satisfies the required regularity conditions, and hence under the loss function \( L_m \), \( \tilde{\delta}^0 \) is dominated by \( \tilde{\delta}^* \) defined (componentwise) as in (5.3.2) where \( \phi_1, \phi_2 \) and \( \phi_3 \) are given in (5.3.6).

The implications of the above example are interesting. First, although the distributions of \( X_1, X_2 \) and \( X_3 \) are very different and each \( \delta_i \) is an admissible estimator of \( \theta_i \) based on \( X_i \), \( \delta^0 \) is inadmissible under \( L_m \). Therefore Stein's phenomena seems to be
very general. Second, the improved estimator corrects $\hat{\theta}_0$ very differently in each coordinate. For $i = 1, 2, 3$, $h_1$, $h_2$, and $h_3$ (so are $d_1$, $d_2$, and $d_3$) are determined independently of each other. (i.e. $h_i$ depends only on the distribution of $X_i$ and $L_i$.) However, the $\varphi_i$'s are obtained by combining these $h_i$'s and $d_i$'s in a definite way.

Section 5.4. Other Generalizations

There are many other possible generalizations of the results of this research. Only three of them are discussed here.

(a) All the results can be easily extended to the loss functions of the form $L(\theta, a) = \sum_{i=1}^{p} z_i \theta_i (\theta_i - a_i)^2$, where $z_1, \ldots, z_p$ are some positive constants. There are two ways to deal with such a loss function.

(i) Include these constants $z_1, \ldots, z_p$ in the difference inequality and solve it. Clearly, the difference inequality can be solved by using theorems in Section 2.2, if and only if the difference inequality corresponding to the loss function $L_m(i.e. z_i = 1, i = 1, \ldots, p)$ can be solved.

(ii) Apply the results in Berger (1977a), in which the problem is decomposed into $p$ subproblems under the loss function

$$\sum_{j=1}^{p} m_j^2 (\theta_j - a_j)^2, \quad j = 1, \ldots, p.$$  

Improved estimators can be found for the original problems, once improved estimators are found under at least one of the subproblems.
(b) The idea of an "upper bound" on the admissible class (cf. Section 4.2) has certainly an analog in the continuous case. However, the concept is more complicated than in the discrete case, since the correction terms, frequently encountered, are not necessarily of the same sign.

(c) All the distributions of $X_1, \ldots, X_p$ considered in this work were assumed to be as in (1.2.2) with $t_i(x_i) > 0$ if and only if $X_i = 0, 1, \ldots$. For the case $t_i(x_i) > 0$ if and only if $x_i = a_i, a_i + 1, \ldots$, for some integer $a_i$, a simple transformation

$$X'_i = X_i - a_i$$

will make our results applicable to the estimation problem based on $X'_1, \ldots, X'_p$. 


