1.1 Introduction

A classical statistical problem is "which probability models adequately describe the data?" This question can be asked for descriptive purposes or as a preliminary to formal inference from the data. Particularly in the latter case, the statistician may have in mind a specific family of probability distributions and the more accurate question should be "Do the data support or impugn the hypothesis that the population distribution is a member of this family?" Most common families of distributions have distribution functions of specified functional form indexed by a (real or vector) parameter. Thus, given independent random variables $X_1, \ldots, X_n$ having common unknown distribution function, we wish to test the hypothesis

$$H : F(.) = G(.| \theta), \quad \text{for some } \theta \in \Omega,$$

where $G(.| \theta)$ is the distribution function of specified form depending on $\theta$. This is called the problem of goodness-of-fit.

The family of tests of goodness-of-fit given by K. Pearson (1900) is the oldest one. Assuming that $G$ is specified, he partitioned the line into cells $E_1, \ldots, E_k$ and based his test on the observed frequencies $(n_1, \ldots, n_k)$ of these cells. He proposed the statistic

$$X^2_p = \sum_{i=1}^{k} \left( \frac{(n_i - np_i)^2}{np_i} \right),$$

(1.1.1)
where \( \pi_i = \int g(x) \, dx, \quad i = 1, \ldots, k \) and \( \sum_{i=1}^{k} \pi_i = n \). It was proved that the asymptotic distribution of \( X_p^2 \) is \( X_{k-1}^2 \). Large values of \( X_p^2 \) is an evidence of lack-of-fit. When \( G \) depended on an unknown parameter, the statistic proposed was
\[
X_p^2(\hat{\theta}_0) = \sum_{i=1}^{k} (n_i - n \pi_i(\hat{\theta}_0))^2 / n \pi_i(\hat{\theta}_0),
\]
(1.1.2)
where \( \hat{\theta}_0 \) is an ML-estimator of \( \theta \). Fisher (1924) proved that the asymptotic distribution of \( X_p^2(\hat{\theta}_0) \) is \( X_{k-s-1}^2 \), where \( s \) is the dimension of \( \theta \). A rigorous proof of this result appears in Cramer (1946). The proof is also given by Birch (1964), Rao (1965) under different regularity conditions. Neyman (1949) defined Best Asymptotically Normal (BAN) estimators and gave some methods of constructing the estimators. BAN estimators possess the same asymptotic properties as those of the maximum likelihood (m.l.) estimators and hence they can be used in computing \( X_p^2 \) in place of m.l. estimators (Cochran, 1952).

Only since 1950's a rigorous study of chi-square statistic has been made. There are extensions of the standard results in different directions. We can group the developments into the following four categories:

1. Formulation of chi-square type distance functions and study of the asymptotic properties of the estimators obtained from these distance functions (Hellinger's distance function, Taylor's distance function, discrepancy function, etc.),
(2) Inclusion of higher order terms in the expansion of the estimators and test statistics and study of their performance (second order efficiency, power computation up to $O\left(\frac{1}{\sqrt{n}}\right)$, etc.),

(3) Use of different estimators in computing the chi-square statistic and study of their asymptotic distributions (use of m.l. estimators from the original data, use of any general root-$n$ consistent estimators, etc.),

(4) Use of random cells instead of fixed cells in determining cell frequencies.

The research work which we have presented in this thesis deals with some problems stated in (1) and (2). It is well known that to the first order approximation, the estimators and tests obtained from chi-square type distance functions are asymptotically equivalent to each other. However, the differences can be observed only if we consider higher order approximations for the asymptotic distributions of the estimators and test statistics. Our main objectives in this study are:

(i) to obtain a general expression for the expected mean square of the chi-square type estimators and to compute second order efficiency and deficiency of the estimators,

(ii) to compute the power of the tests based on chi-square type statistics up to the order $\frac{1}{\sqrt{n}}$ for simple and composite hypothesis, and

(iii) to find suitable corrections to the chi-square type statistics by making use of moments.
1.2 Distance Functions

Neyman (1949) introduced the method of minimizing suitable distance functions between the observed proportions \( (p_1, \ldots, p_k) \), where
\[
p_i = \frac{n_i}{n}, \quad i = 1, \ldots, k
\]
and the cell probabilities \( (\pi_1(\theta), \ldots, \pi_k(\theta)) \) to obtain BAN estimators of \( \theta \). In particular, he proved that minimization of the functions
\[
X_p^2(\theta) = \sum \left[ \frac{n_i - n \pi_i(\theta)}{n \pi_i(\theta)} \right]^2
\]
and
\[
X_n^2(\theta) = \sum \left[ \frac{n_i - n \pi_i(\theta)}{n_1} \right]^2
\]
produced BAN estimators of \( \theta \). The summation extends over \( 1, 2, \ldots, k \) unless otherwise stated. Taylor (1953) defined a class of functions
\[
D_n(p, \pi(\theta)) = n \sum \left( \frac{h(p_i) - h(\pi_i(\theta))}{\pi_i(\theta) h'(\pi_i(\theta))} \right)^2,
\]
where \( h(x) \) is strictly monotonic in \( 0 < x < 1 \). Recently, Cressie and Read (1984) defined a class of discrepancy functions
\[
I^\lambda = \frac{n}{\lambda(\lambda+1)} \sum p_i \left[ \left( \frac{p_i}{\pi_i(\theta)} \right)^\lambda - 1 \right], \quad \lambda \in \mathbb{R}
\]
The well known distance functions, namely, likelihood ratio, chi-square, Hellinger distance function are particular cases of the above distance functions. In defining the distance functions, the concept of distance is modified by the presence of weights. Minimization of \( D_n \) and \( I^\lambda \) give BAN estimators of \( \theta \). The distance functions, when evaluated at the estimates, can be considered as goodness-of-fit.
test statistics.

1.3 Summary of Results Obtained

Rao (1961, 63) introduced the concept of second order efficiency (s.o.e.) in order to differentiate BAN estimators. In Chapter II we deal with s.o.e. of BAN estimators for the multinomial distribution. A general expression for the s.o.e. of a BAN estimator is obtained which is based on its estimating equation. We have followed Rao (1963), Ghosh and Subramanyam (1974) in computing the s.o.e. It is shown that the expression for the s.o.e. has two components, the one which is independent of the estimating equation and the other which depends on the estimating equation. The second component is non-negative and reduces to zero for the m.l. estimator. Therefore, m.l. estimator has the highest s.o.e. We prove that for an estimator obtained by minimising a particular member of $D_\theta(p, \pi(\theta))$ the second component reduces to zero and hence this estimator also has the highest s.o.e. We have also studied the deficiency of different estimators.

The alternative hypothesis is not well specified when we are testing the hypothesis of goodness-of-fit. Hence the power function of the test is studied under the local alternatives. The asymptotic distribution of $X^2_p(\tilde{\theta}_n)$ and $D_n(p, \pi(\tilde{\theta}_n))$ for any BAN estimator $\tilde{\theta}_n$ under the local alternatives, is non-central chi-square with same non-centrality parameter. This is true of the well known large sample tests namely, Wald test, likelihood ratio test and the score test also. Recently, Peers (1971) computed power of these tests upto $O(\frac{1}{\sqrt{n}})$. Peers considered the case of simple hypothesis. Hayakawa (1975) and
Harris & Peers (1980) dealt with composite hypothesis. In Chapter III we have considered the case of simple hypothesis and computed the power of the tests based on $D_n(\ell, \pi_0)$. We have modified the method of Peers (1971) to suit the hypothesis of goodness-of-fit. Our results agree with those of Peers for the score test.

In Chapter IV we have extended the results of Chapter III to the case of composite hypothesis. Our results can be considered as extensions of the basic result obtained by Mitra (1958). The case of single parameter is dealt in detail. Extension to the multi-parameteric case is outlined.

The asymptotic distribution of $X_p^2$, $X_N^2$, $D_n(\ell, \pi)$ and $I^\lambda$ statistics obtained by substituting any BAN estimator for $\theta$, when the hypothesis holds, is a central chi-square distribution. Adequacy of the chi-square fit to the statistic with a moderate sample size has been discussed in Chapter V. Corrections to the statistics, based on the first two moments, have been suggested. These corrections are justified on the basis of empirical studies (Larntz, 1978, Read, 1984). Our empirical study aims at obtaining the exact levels attained by members of $D_n(\ell, \pi_0)$ and its corrected statistic at selected values of nominal level of significance. We conclude that $X_p^2$ is adequately approximated by a chi-square distribution whereas the other members of $D_n(\ell, \pi_0)$ need corrections.
1.4 Some Definitions

Let $X_1, \ldots, X_n$ be a sequence of random variables providing information on a parameter $\theta$ (one-dimensional) and let $T_n$ be an estimator of $\theta$ based on $n$ observations.

**Definition 1.1 Probability Consistency (P.C.)**

A sequence of statistics $T_n$ is said to be consistent for $\theta$ if $T_n \to \theta$ in probability.

**Definition 1.2 Fisher Consistency (F.C.)**

A statistic $T = f(S_n)$, where $S_n$ is the empirical distribution function based on $n$ i.i.d. (independent and identically distributed) observations and $f$ is a weakly continuous functional defined on the space of distribution functions is said to be Fisher consistent if $f(F_\theta) \equiv \theta$, where $F_\theta$ is the true distribution function from which the observations are drawn.

Note that F.C. $\Rightarrow$ P.C. and that F.C. refers to a restriction on the estimate for any finite $n$ and it is not just a limiting property of a sequence of statistics.

**Definition 1.3 CAN Estimator**

An estimator $T_n$ is said to be CAN (consistent asymptotically normal) estimator of $g(\theta)$ if the asymptotic distribution of $\sqrt{n}[T_n - g(\theta)]$ is normal.

A CAN estimator $T_n$ is said to be the best or efficient if the variance of the limiting distribution of $\sqrt{n}[T_n - g(\theta)]$ is minimum.
It was thought that when i.i.d. observations were considered the lower bound for the variance of the limiting distribution was \( \left[ g'(\theta) \right]^2 / i. \), where \( i. = i.(\theta) \) is the Fisher information in a single observation. But result concerning the lower bound is not true without further conditions.

**Definition 1.4 First Order Efficiency (Rao, 1961)**

A statistic \( T_n \) is said to be first order efficient (f.o.e.) if

\[
\sqrt{n} |\beta(T_n - \theta) - Z_n| \rightarrow 0
\]

in probability, where \( \beta \) is a function of \( \theta \) only, and \( Z_n = \frac{1}{n} \frac{d \log p_\theta(X_n)}{d \theta} \),

\( p_\theta(X_n) \) being the density function of the observations \( X_n = (X_1, \ldots, X_n) \).

The above condition implies that the asymptotic correlation between \( T_n \) and \( Z_n \) is unity.

The above definitions can be extended to the case when \( \theta \) is s-dimensional.

### 1.5 Estimation of the Parameter in a Multinomial Distribution

#### 1.5.1 Condition for the First Order Efficiency

Consider a multinomial distribution with \( k \) cells. The cell probabilities are denoted by \( \pi_1(\theta), \ldots, \pi_k(\theta), \pi_i(\theta) > 0 \) \( i = 1, \ldots, k \), \( \sum_{i=1}^{k} \pi_i(\theta) = 1 \), where \( \theta \) is an unknown real parameter. Let \( n_1, \ldots, n_k \) denote the observed cell frequencies with \( \sum_{i=1}^{k} n_i = n \) and \( p_i = \frac{n_i}{n} \), \( i = 1, \ldots, k \). We denote this multinomial distribution by \( M(n; \pi_1, \ldots, \pi_k) \). An estimate of \( \theta \) is a suitably chosen root of the equation

\[
f(\theta; p_1, \ldots, p_k) = 0 \quad (1.5.1)
\]
which is generally obtained by minimizing a distance function. The following assumptions are made:

(A1). The true parameter point \( \theta \) is an interior point of the parameter space and \( \pi_1(\theta) \) admit second order derivative w.r.t. \( \theta \) which is continuous.

(A2). \( f(\theta ; \pi_1(\theta), ..., \pi_k(\theta)) = 0 \) \hspace{1cm} (1.5.2)

(A3). \( f(\theta ; p_1, p_2, ..., p_k) \) has continuous derivatives upto second order in \( \theta \) as well as \( p_1, ..., p_k \).

Let us denote \( f', f_r, \) and \( f_{rs} \), the derivative \( \partial f/\partial \theta \), \( \partial f/\partial p_r \)
\( \partial^2 f/\partial p_r \partial p_s \) evaluated at \( \theta \) and \( p_r = \pi_r \) respectively. We state the following Lemma which is well established (see Rao, 1961).

Lemma 1.1 Under the assumptions (A1) to (A3) there exists a root \( \hat{\theta}_n \) of the equation \( f(\theta ; p_1, ..., p_k) = 0 \) such that \( \hat{\theta}_n \to \theta \), the true value, with probability one, and for asymptotic efficiency (first order) to hold uniformly for all \( \theta \), the necessary and sufficient condition is

\[
\frac{f'_r}{f} = -\frac{1}{1_r} \cdot \frac{\pi'_r}{\pi_r},
\]

where

\[
1_r = \sum \frac{\pi_r^2}{\pi_r}.
\]

1.5.2 BAN Estimators

Neyman in his famous paper defined BAN estimators for the multinomial distributions and gave methods of constructing such estimators. We state some Lemmas and the definitions given in Neyman (1949).
Definition 1.5  A function \( \bar{g}_n \) of \( p_1, \ldots, p_k \), which is independent of \( n \), is called BAN estimator of \( \theta \) if the following conditions are satisfied:

(i) \( \bar{g}_n \rightarrow \theta \) in probability,

(ii) \( \sqrt{n}(\bar{g}_n - \theta) \rightarrow \mathcal{N}(0, \sigma^2) \),

(iii) If \( \theta^* \) is any other estimator satisfying (i) and (ii) with variance \( \sigma'^2 \) then \( \sigma'^2 \geq \sigma^2 \),

(iv) \( \bar{g}_n \) possesses continuous partial derivatives with respect to \( p_r \).

Lemma 1.2  If \( \bar{g}_n \) is a BAN estimator of \( \theta \), then it can be represented in the form

\[
\sqrt{n}(\bar{g}_n - \theta) = \sum a_r \sqrt{n}(p_r - \pi_r) + o_p(1),
\]  

where

\[
a_r = \frac{\partial \bar{g}_n}{\partial \pi_r}, \quad r = 1, \ldots, k.
\]

The following theorem implies that BAN estimator is first order efficient.

Theorem 1.1  A statistic \( \bar{g}_n \), function of \( p_1, \ldots, p_k \), is a BAN estimator of \( \theta \) if

(i) \( \bar{g}_n \) has continuous partial derivatives w.r.t. \( p_r \), \( r = 1, \ldots, k \),

(ii) \( \bar{g}_n(\pi) = \theta \),

(iii) \( a_r = \frac{\partial \bar{g}_n}{\partial p_r} \bigg|_{p_r = \pi_r} = \frac{1}{\pi_r} \frac{\pi_r'}{\pi_r} \).  

Proof:  The general representation of \( \bar{g}_n \) is

\[
\sqrt{n}(\bar{g}_n - \theta) = \sum a_r \sqrt{n}(p_r - \pi_r) + o_p(1).
\]
We assume that $\sum a_r \pi_r = 0$. Then

$$\sqrt{n}(\bar{\pi}_n - \theta) \to \text{AN}(0, \sigma^2)$$

where $\sigma^2 = \sum a_r^2 \pi_r$. The expression for the variance follows from $V(p_i) = \pi_i(1 - \pi_i)/n$ and $\text{cov}(p_i, p_j) = -\pi_i \pi_j/n$, $i \neq j$ and asymptotic normality follows from the asymptotic normality of $\sqrt{n}(p_r - \pi_r)$, $r = 1, \ldots, k$.

Using condition (11) and differentiating w.r.t $\theta$ we get

$$\frac{\partial \bar{\pi}_n}{\partial \theta} = \sum \frac{\partial \bar{\pi}_n}{\partial \pi_r} \frac{\partial \pi_r}{\partial \theta} = 1$$

which implies

$$\sum a_r \pi_r = 1.$$  \hspace{1cm} (1.5.7)

Substituting (1.5.7) in the inequality

$$\sum \pi_r \left( a_r - \frac{1}{\lambda}. \frac{\pi_r'}{\pi_r} \right)^2 \geq 0$$

we get

$$\sum a_r^2 \pi_r - \frac{1}{\lambda} \geq 0.$$  \hspace{1cm} (1.5.8)

The minimum value of (1.5.8) is attained when

$$a_r = \frac{1}{\lambda}. \frac{\pi_r'}{\pi_r}.$$ 

Hence $\bar{\pi}_n$ has the representation

$$\sqrt{n}(\bar{\pi}_n - \theta) = \frac{1}{\lambda}. \sum \frac{\pi_r'}{\pi_r} \sqrt{n}(p_r - \pi_r) + o_p(1).$$  \hspace{1cm} (1.5.9)

The log likelihood function is $\sum n_r \log \pi_r$ and hence
From Definition 1.4, (1.5.9) and (1.5.10) it follows that $\bar{\theta}_n$ is f.o.e.

1.5.3 Generalised Minimum Chi-square Estimators

For the multinomial distribution minimum chi-square and modified chi-square estimators, obtained by minimizing (1.2.1) and (1.2.2) have the same asymptotic properties as m.l. estimator and hence these are BAN estimators (Neyman, 1949). The function $\chi^2_p$ or $\chi^2_n$ can be considered as measures of distance between $\pi_\cdot = (p_1, \ldots, p_k)$ and $\pi = (\pi_1, \ldots, \pi_k)$. In general a distance function can be defined as follows:

**Definition 1.6** $\Delta(\pi, \bar{\pi})$ is a measure of distance between $\pi$ and $\bar{\pi}$ if it satisfies

1. $\Delta(\pi, \bar{\pi}) = 0$ for $\pi = \bar{\pi}$,
2. $\Delta(\pi, \bar{\pi}) > 0$ for $\pi \neq \bar{\pi}$,
3. $\Delta(\pi, \bar{\pi})$ is continuous and possesses partial derivatives up to second order with respect to $p_r$ and $\theta$.

**Lemma 1.3** The value of $\bar{\theta}_n$, function of $p_1, \ldots, p_k$, which minimises $\Delta(\pi, \pi(\theta))$ is a BAN estimator of $\theta$ if $\Delta(\pi, \bar{\pi})$ is such that

$$Z_n = \frac{1}{n} \left( \sum \frac{\log p \left( x_n \right)}{\pi} \right) = \frac{1}{n} \sum \frac{n_r}{\pi_r} \pi_r'$$

$$= \sum \left( \frac{p_r - \pi_r}{\pi_r} \right) \pi_r'. \quad (1.5.10)$$
\[ \frac{\partial^2 \Delta(p, \pi)}{\partial \theta \partial p_r} \bigg|_{p_r = \pi_r} = C \frac{\pi_r'}{\pi_r} \]  \hspace{1cm} (1.5.11)

and

\[ \frac{\partial^2 \Delta(p, \pi)}{\partial^2 \theta} \bigg|_{p_r = \pi_r} = -C \sum \frac{\pi_r^2}{\pi_r}, \]  \hspace{1cm} (1.5.12)

where \( C \) is a constant.

**Proof**: It is required to show that condition (1.5.5) is satisfied.

The estimating equation is

\[ \psi = \frac{\partial \Delta(p, \pi)}{\partial \theta} = 0 \]

In order to find \( \partial \pi_n / \partial p_r \), we substitute \( \tilde{\pi}_n \) and differentiate with respect to \( p_r \). This results in equation of the form

\[ \frac{\partial \psi}{\partial p_r} + \frac{\partial \psi}{\partial \pi_n} \cdot \frac{\partial \pi_n}{\partial p_r} = 0. \]

Solving for \( \partial \pi_n / \partial p_r \) at \( p_r = \pi_r \), we get

\[ \frac{\partial \pi_n}{\partial p_r} \bigg|_{p_r = \pi_r} = -\frac{\partial \psi}{\partial p_r} / \frac{\partial \pi_n}{\partial \pi_n} \bigg|_{p_r = \pi_r} \]  \hspace{1cm} (1.5.13)

If

\[ \frac{\partial \psi}{\partial p_r} \bigg|_{p_r = \pi_r} = C \frac{\pi_r'}{\pi_r} \quad \text{and} \quad \frac{\partial \pi_n}{\partial \pi_n} \bigg|_{p_r = \pi_r} = -C \sum \frac{\pi_r^2}{\pi_r}, \]

then,
Hence the lemma.

Lemma 1.4 If

\[ D_n(p, \pi) = n \left( \frac{h(p_r) - h(\pi_r)}{\pi_r \left[ h'(\pi_r) \right]^2} \right)^2 \]

where \( h(x) \) is a monotone function of \( x \) for \( 0 < x < 1 \) and possesses continuous derivatives up to third order, then \( D_n(p, \pi) \) satisfies the conditions of Lemma 1.3.

Proof: By differentiating \( D_n(p, \pi) \) and evaluating at \( p_r = \pi_r \), we get

\[ \frac{\partial^2 D_n(p, \pi)}{\partial p_r \partial \theta} = -2n \frac{\pi_r'}{\pi_r} \quad (1.5.14) \]

and

\[ \frac{\partial^2 D_n(p, \pi)}{\partial \theta^2} = 2n \left( \frac{\pi_r'}{\pi_r} \right)^2 \quad (1.5.15) \]

The expressions (1.5.14) and (1.5.15) imply that \( D_n(p, \pi) \) belongs to the class of distance functions which provides BAN estimators. For any suitable choice of the function \( h \), we get a member of the class. The following distance function is obtained as a particular case.

(1) The Freeman-Tukey distance function is obtained for \( h(x) = \sqrt{x} \)

and is given by
15

\[ 4n \sum (\sqrt{p_r} - \sqrt{\pi_r})^2. \] (1.5.16)

(iii) The minimum chi-square distance function is obtained for \( h(x) = x \)

(iii) If \( h(x) = \log x \), then we get distance function

\[ n \sum \pi_r (\log p_r - \log \pi_r)^2. \] (1.5.17)

(iv) If \( h(x) = x^n \), \( n \in \mathbb{R} \), then we get the distance function

\[ n \sum \frac{(p_r^n - \pi_r^n)^2}{\pi_r^n (n^2 \pi_r^n - 1)}. \] (1.5.18)

1.6 The Multiparameter Case

In this section we list some of the results previously obtained, when the multinomial probabilities \( \pi_i \)'s, \( i = 1, \ldots, k \), are functions of the parameters \( \theta_i \), \( \theta_s \). We make the following assumptions.

(B1). The true parameter point \( \bar{\theta} = (\theta_1, \ldots, \theta_s) \) is an interior point of the parameter space \( \Omega \).

(B2). \( \pi_i(\theta_1, \ldots, \theta_s) > 0 \) for all \( i \).

(B3). \( \pi_i(\theta_1, \ldots, \theta_s) \) admit continuous second order partial derivatives with respect to \( \theta_j \)'s \( j = 1, \ldots, s \).

(B4). The matrix \( \frac{\partial^2 \pi(\bar{\theta})}{\partial \bar{\theta}^2} \) is of full rank \( s \).

Definition 1.7 A function \( \bar{\theta}_n \) of \( p_1, \ldots, p_k \) which is independent of \( n \), is called BAN estimator of \( \bar{\theta} \), if the following conditions are satisfied:

(i) \( \bar{\theta}_n \rightarrow \bar{\theta} \) in probability,
\[(11) \quad \sqrt{n} (\overline{\theta}_n - \theta) + \mathcal{N}(0, I_n), \]

\[(111) \quad \text{If} \quad \theta^*_n \quad \text{is any other estimator satisfying (1) and (11) with covariance matrix} \quad \Sigma^* \quad \text{then} \quad \Sigma^* - \Sigma \quad \text{is positive semidefinite.} \]

\[(iv) \quad \overline{\theta}_n \quad \text{possesses continuous partial derivatives with respect to} \quad p_r. \]

**Lemma 1.5** If \( \overline{\theta}_n \) is a BAN estimator, then it can be expanded in the form

\[
\sqrt{n} (\overline{\theta}_n - \theta) = (A'A)^{-1} A' D^{-1/2} \sqrt{n}(p - \pi) + o_p(1) \quad (1.6.1)
\]

where \( A \) is the \( k \times s \) matrix whose \((i, j)\)th element is \( \pi_{ij} - \pi_{ij} \) and \( D = \text{diag}(\pi_1, \ldots, \pi_k) \). Note that \( A'A \) is the information matrix. For the multiparameter case, the generalizations of Lemmas 1.3 and 1.4 follow in a similar way as the case of single parameter. For details refer Taylor (1953).

### 1.7 The \( \chi^2 \) and the Generalised \( \chi^2 \) Tests

In this section we state some standard results on \( \chi^2 \) goodness-of-fit tests. The proofs are found in standard books [Cramér, 1946; Rao, 1965; Bishop, Fienberg and Holland, 1975; Serfling, 1980].

The well known Pearson \( \chi^2 \) statistic for testing the goodness-of-fit is given by

\[
\chi^2_p(\overline{\theta}_n) = n \sum \frac{[p_i - \pi_1(\overline{\theta}_n)]^2}{\pi_1(\overline{\theta}_n)}, \quad (1.7.1)
\]

where \( \overline{\theta}_n \) is any BAN estimator of \( \theta \). When \( \pi_i's \quad i = 1, \ldots, k, \) are completely specified by the hypothesis \( H_0: \pi_i = \pi_{0i}, \quad i = 1, \ldots, k \), the
Pearson chi-square statistic is given by

\[ \chi^2_p = n \sum \frac{(p_i - \pi_{0i})^2}{\pi_{0i}}. \]  

(1.7.2)

**Theorem 1.2.** Under the hypothesis \( H_0 : \pi = \pi_0 \), \( \chi^2_p \) has an asymptotic chi-square distribution with \( k-1 \) d.f.

**Theorem 1.3.** Under the contiguous alternatives \( H^*_1 : \pi_0 + C/\sqrt{n} \), where \( C = (C_1, \ldots, C_k) \) is a set of numbers such that \( \sum_{i=1}^{k} C_i = 0 \), the asymptotic distribution of \( \chi^2_p \) is a non-central chi-square with non-centrality parameter \( C D^{-1} C \) and degrees of freedom \( k-1 \). When \( \pi_i's, i = 1, \ldots, k \), are functions of \( \theta_1, \ldots, \theta_s \), we assume that regularity conditions (B1) to (B5) of the previous section to hold.

**Theorem 1.4.** Under the hypothesis \( H_0 : \pi = \pi(\theta) \), the asymptotic distribution of \( \chi^2_p(\theta) \) is chi-square with \( k-s-1 \) d.f.

**Theorem 1.5.** (Mitra, 1958). Under the contiguous alternatives \( H_1 : \pi = \pi(\theta) + C/\sqrt{n} \), the asymptotic distribution of \( \chi^2_p(\theta) \) is non-central chi-square with non-centrality parameter

\[ C D^{-1/2} \pi(\theta) \left[ I - A(A'A)^{-1}A' \right] D^{-1/2} \pi(\theta) \sim C \]

and degrees of freedom \( k-s-1 \).

We now consider the chi-square type statistic

\[ D_n(p, \pi(\theta)) = n \sum \frac{(h(p_i) - h(\pi_i(\theta)))^2}{\pi_i(\theta)[h(\pi_i(\theta))]^2}. \]

(1.7.3)

Each member of this family indicates the discrepancy between observed and expected frequencies and hence it can be considered as a goodness-
of-fit statistic. Consider the simple hypothesis $H_0: \pi = \pi_0$. Using
Taylor's expansion for $h(p_i)$, it is easy to observe

$$D_n(\mathbf{p}, \pi_0) = n \sum \frac{[h(p_i) - h(\pi_{01})]^2}{\pi_{01} [h'(\pi_{01})]^2}$$

$$= n \sum \frac{(p_i - \pi_{01})^2}{\pi_{01}} + o_p(1), \quad (1.7.4)$$

since $p_i = \pi_{01} + o_p(n^{-1/2})$, $i = 1, \ldots, k$. From (1.7.4) and Theorem 1.2,
it follows that the asymptotic distribution of $D_n(\mathbf{p}, \pi_0)$ is chi-square
with $k-1$ d.f.. When the hypothesis is $\pi = \pi(\theta)$, on similar lines,
using the facts $p_i = \pi_1(\theta) + o_p(n^{-1/2})$ and $\pi_{1}(\theta) = \pi_1(\theta) + o_p(n^{-1/2})$
we have

$$D_n(\mathbf{p}, \pi(\theta)) = n \sum \frac{[h(p_i) - h(\pi_1(\hat{\theta}_n))]^2}{\pi_1(\hat{\theta}_n) [h'(\pi_1(\hat{\theta}_n))]^2}$$

$$= n \sum \frac{[p_i - \pi_1(\hat{\theta}_n)]^2}{\pi_1(\hat{\theta}_n)} + o_p(1). \quad (1.7.5)$$

From the above relationship, it follows that $D_n(\mathbf{p}, \pi(\theta))$ has an
asymptotic chi-square distribution with $k-1$ d.f.. The relationships
(1.7.4) and (1.7.5) hold even under the contiguous alternatives
$H_1: \pi = \pi(\theta) + C/\sqrt{n}$. Hence the asymptotic distribution of $D_n(\mathbf{p}, \pi(\theta))$
follow from Theorems 1.3 and 1.5.