GOODNESS-OF-FIT TEST FOR THE FAMILY OF LOGISTIC DISTRIBUTIONS

N. AGUIRRE* AND M. NIKULIN†

Chi-squared goodness-of-fit test for the family of logistic distributions is proposed. Different methods of estimation of the unknown parameters $\theta$ of the family are compared. The problem of homogeneity is considered.

Key words: Logistic distribution; BAN estimator; method of moments; maximum likelihood method; Chernoff-Lehmann theorem; MVUE; homogeneity problem; chi-squared test.

AMS classification: 62E15, 62E20, 62F03, 62G05, 62G10

1. INTRODUCTION

Let $X_1, \ldots, X_n$ be independent identically distributed random variables and suppose that according to the hypothesis $H_0$

\[(1) \ P\{X_i \leq x\} = F(x; \theta), \quad \theta = (\theta_1, \ldots, \theta_s)^T \in \Theta \subset \mathbb{R}^s, \quad x \in \mathbb{R}^1, \]

* N. Aguirre. Mathématiques Appliquées, Université Bordeaux 1, France.
† M. Nikulin. Mathématiques Stochastiques, Université Bordeaux 2, V.A. Steklov Mathematical Institute, St. Petersburg, Russia.

- Article rebut el novembre de 1993.
- Acceptat el juliol de 1994.
where $\Theta$ is an open set. We devide the real line into $k$ intervals $I_1, \ldots, I_k$:

$$I_1 \cup \ldots \cup I_k = \mathbb{R}, \quad I_i \cap I_j = \emptyset, \quad i \neq j.$$

We shall suppose that

$$p_i(\theta) = P\{X_1 \in I_i \mid H_0\} > 0, \quad i = 1, \ldots, k.$$  

Let $\nu = (\nu_1, \ldots, \nu_k)^T$ be the vector of frequencies arising as a result of grouping the random variables $X_1, \ldots, X_n$ into the classes $I_1, \ldots, I_k$. We denote

$$X_n^2(\theta) = X_n^T(\theta)X_n(\theta) = \sum_{i=1}^{k} \frac{(\nu_i - np_i(\theta))^2}{np_i(\theta)},$$

where

$$X_n = \left( \frac{\nu_1 - np_1(\theta)}{\sqrt{np_1(\theta)}}, \ldots, \frac{\nu_k - np_k(\theta)}{\sqrt{np_k(\theta)}} \right)^T.$$  

**Theorem.** (K. Pearson, 1900)

If $\theta$ is known or given by the hypothesis $H_0$, then

$$\lim_{n \to \infty} P\{X_n^2(\theta) \geq x \mid H_0\} = P\{X_{k-1}^2 \geq x\}$$

It is known that if the value of the parameter $\theta$ is unknown and is estimated relative to the observed values of $X_1, \ldots, X_n$, the limiting distribution of the Pearson’s statistics $X_n^2(\theta_n^*)$ is given by the asymptotic properties of the estimator $\theta_n^*$ which is substituted into (3) in place of $\theta$.

Here we shall give some results concerning this problem.

2. **FISHER’S THEOREM**

Following Cramer (1946) we suppose that

1) $p_i(\theta) > c > 0$, $i = 1, \ldots, k$, ($k \geq s + 2$);

2) $\frac{\partial^2 p_i(\theta)}{\partial \theta_j \partial \theta_l}$ are continuous functions.
3) the information matrix of Fisher

\[ J = J(\theta) = \frac{1}{\sum_{i=1}^{k} p_i(\theta)} \begin{bmatrix} \frac{\partial p_i(\theta)}{\partial \theta_i} & \frac{\partial p_i(\theta)}{\partial \theta_j} \end{bmatrix}_{s \times s} = B^T(\theta)B(\theta) \]

exists, and rank \( J = s \), where

\[ B(\theta) = \frac{1}{\sqrt{p_i(\theta)}} \frac{\partial p_i(\theta)}{\partial \theta_j} \]

In this case \( nJ \) is the information matrix of Fisher of the statistic \( \nu = (\nu_1, \ldots, \nu_k)^T \).

Let \( \tilde{\theta}_n \) is the minimum chi-squared estimator for \( \theta \),

\[ X_n^2(\tilde{\theta}_n) = \min_{\theta \in \Theta} X_n^2(\theta), \]

or an estimator asymptotically equivalent to it. As it was shown by Cramer(1946), a root \( \tilde{\theta}_n \) of the system

\[ \sum_{i=1}^{k} \frac{\nu_i}{np_i(\theta)} \frac{\partial p_i(\theta)}{\partial \theta_j} = 0, \quad j = 1, \ldots, s. \]

is such an estimator and, under \( H_0 \) as \( n \to \infty \), the vector \( \sqrt{n}(\tilde{\theta}_n - \theta) \) satisfies the asymptotic relation

\[ \sqrt{n}(\tilde{\theta}_n - \theta) = J^{-1}(\theta)B^T(\theta)X_n(\theta) + o(1_s), \]

where \( o(1_s) \) is a random vector converging to \( 0 \), in \( P_\theta \) - probability, and hence \( \sqrt{n}(\tilde{\theta}_n - \theta) \) is asymptotically normally distributed with parameters \( 0_s \), and \( J^{-1}(\theta) \).

**Theorem.** (Fisher(1928), Cramer(1946))

If the regularity conditions of Cramer hold then

\[ \lim_{n \to \infty} P\{X_n^2(\tilde{\theta}_n) \geq x \mid H_0\} = P\{X_{k-s-1}^2 \geq x\}. \]

3. **CHERNOFF-LEHMANN'S THEOREM**

We suppose that the regularity conditions of Chernoff-Lehmann's (1954) hold:
1) \( F(x; \theta) \) has probability density \( f(x; \theta) \), where all

\[
\frac{\partial^2 f(x; \theta)}{\partial \theta_j \partial \theta_i}
\]

are continuous functions on \( \mathbb{R}^1 \times \Theta \);

2) the information matrix of Fisher

\[
I(\theta) = \| I_{ij} \|_{s \times s} = \mathbb{E}_\theta \Lambda(\theta) \Lambda^T(\theta),
\]

corresponding to one observation \( X_1 \) exists and is positive definite for any \( \theta \in \Theta \), where

\[
\Lambda(\theta) = \frac{\partial}{\partial \theta} \ln f(X_1; \theta);
\]

\((nI(\theta))\) is the amount of information of Fisher about \( \theta \) in sample \( \mathbb{X} = (X_1, \ldots, X_n)^T \).

3) differentiation with respect to parameters under the integral sign of

\[
\int f(x; \theta) dx = 1
\]

is permissible, i.e.

\[
\frac{\partial}{\partial \theta_i} \int f(x; \theta) dx = \int \frac{\partial}{\partial \theta_i} f(x; \theta) dx = 0, \quad i = 1, \ldots, s;
\]

4) the matrix \( W = [w_{ij}] \) with elements

\[
w_{ij} = \int I_j \frac{\partial}{\partial \theta_i} f(x; \theta) dx
\]

has rank \( s \);

5) the maximum likelihood estimator \( \hat{\theta}_n \) exists,

\[
L(\hat{\theta}_n) = \sup_{\theta \in \Theta} L(\theta),
\]

where

\[
L(\theta) = \prod_{i=1}^{n} f(X_i; \theta).
\]

As it is known (see for example Rao (1965)), \( \hat{\theta}_n \) is a solution of the likelihood equation

\[
\Lambda(\theta) = 0_s,
\]

320
and, under $H_0$ as $n \to \infty$, the vector $\sqrt{n}(\hat{\theta}_n - \theta)$ satisfies the asymptotic relation

\begin{equation}
\sqrt{n}(\hat{\theta}_n - \theta) = \frac{1}{\sqrt{n}} I^{-1}(\theta) \lambda(\theta) + o(1) \),
\end{equation}

from where it follows that $\sqrt{n}(\hat{\theta}_n - \theta)$ is asymptotically normally distributed with parameters 0, and $I^{-1}(\theta)$ and hence $\hat{\theta}_n$ is asymptotically efficient estimator. We say that $\hat{\theta}_n$ is a BAN estimator.

**Theorem.** (Chernoff-Lehmann, 1954)

If the regularity conditions 1)-5) hold then

\begin{equation}
\lim_{n \to \infty} P\left\{ \chi^2_n(\hat{\theta}_n) \geq x \mid H_0 \right\} = P\left\{ \chi^2_{k-s-1} + \sum_{i=1}^{s} \lambda_i \xi_i^2 \geq x \right\},
\end{equation}

where $\chi^2_{k-s-1}, \xi_1, \ldots, \xi_s$ are independent, $\xi_i \sim N(0, 1)$ and $\lambda_i = \lambda_i(\theta)$, $0 < \lambda_i(\theta) < 1$, $i = 1, 2, \ldots, s$, are the roots of the equation

\begin{equation}
(1 - \lambda_i) I(\theta) - J(\theta) = 0.
\end{equation}

**Remark 1.** We note here that in continuous case $\nu = (\nu_1, \ldots, \nu_k)^T$ is not sufficient statistic, and hence the matrix $I(\theta) - J(\theta)$ is positive definite.

**Remark 2.** Let us consider the density family

\begin{equation}
f(x; \theta) = h(x) e^{\theta_0 x^m + v(\theta)}, \quad x \in \mathcal{X} \subseteq \mathbb{R}^1,
\end{equation}

$\mathcal{X}$ is open in $\mathbb{R}^1$, $\mathcal{X} = \{ x : f(x; \theta) > 0 \}$, $\theta \in \Theta$.

The family (24) is very rich: it contains Poisson, normal distributions etc. It’s evident that

\begin{equation}
U_n = \left( \sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2, \ldots, \sum_{i=1}^{n} X_i^s \right)^T
\end{equation}

is completeminimal sufficient statistics for the family (24).

We suppose that

1) the support $\mathcal{X}$ does not depend on $\theta$;

321
2) the matrix of Hessen

\[ \mathbf{H}_v(\theta) = -\left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} v(\theta) \right]_{s \times s} \]

of the function \( v(\theta) \) is positive definite;

3) the moment \( a_s(\theta) = \mathbb{E}_\theta X_1^s \) exists.

In this case, using the results of Zacks (1971), it is not difficult to show (see, for example, Dzhaparidze and Nikulin (1991)) that the maximum likelihood estimator \( \hat{\theta}_n = \hat{\theta}_n(U_n) \) and the method of moments estimator \( \tilde{\theta}_n = \tilde{\theta}_n(U_n) \) of \( \theta \) coincide, i.e. \( \hat{\theta}_n = \tilde{\theta}_n \). Let

\[ a(\theta) = (a_1(\theta), \ldots, a_s(\theta))^T \quad \text{and} \quad T_n = \frac{1}{n} U_n. \]

One can verify that

\[ a(\theta) = -\frac{\partial}{\partial \theta} v(\theta), \]

and hence the likelihood equation is \( T_n = a(\theta) \), i.e. \( \hat{\theta}_n \) is root of this equation. On the other hand we have \( \mathbb{E}_\theta T_n \equiv a(\theta) \), and hence from the properties of the statistics \( U_n \) it follows that \( T_n \) is the MVUE of \( a(\theta) \), and \( \hat{\theta}_n \) is the root of the same equation \( T_n = a(\theta) \), which we used to find \( \hat{\theta}_n \). Hence \( \hat{\theta}_n = \tilde{\theta}_n \), i.e. under the conditions 1)-3) the method of moments gives for the family (24) an asymptotically efficient (BAN) estimator. We remark that in general an estimator based on the method of moments is not asymptotically efficient, and hence does not verify the Chernoff-Lehmann theorem. In “Handbook of the logistic distribution” in chap. 13, it is reported that this theorem is applied by Massaro and d’Agostino using \( \theta_n = (\bar{X}_n, s_n^2)^T \), (the moments method estimator of \( \theta = (\mathbb{E}X_1, \text{Var}X_1)^T \) for the family of the logistic distributions. But \( \hat{\theta}_n \) is not efficient and ever not asymptotically efficient for the logistic family, and hence is not BAN, since this family does not belong to the exponential family (24) and \( (\bar{X}_n, s_n^2)^T \) is not sufficient statistic in this situation. Hence, the tables of critical points, proposed by Massaro et d’Agostino in section 13.9 are not valid.

4. ROY’S EXTENSION OF THE CHERNOFF-LEHMANN THEOREM

We consider here the result of Dahiya and Gurland (1970,1972), concerning the chi-squared test of Pearson with random intervals, which is an extension of
the unpublished result of Roy (1956) and Chernoff-Lehmann's theorem. It can be found more information about Chernoff-Lehmann's theorem in the paper of LeCam, Mahan and Singh (1983).

Let $\theta_n^*$ is an $\sqrt{n}$ - consistent estimator for $\theta$ such that

$$\sqrt{n}(\theta_n^* - \theta) = \frac{1}{n} \sum_{j=1}^{n} v(X_j) + o(1_n), \quad (27)$$

where a function $v = (v_1, \ldots, v_s)^T$ is such that

$$E_{\theta} v(X_1) = 0, \quad \text{and} \quad \text{Var}_{\theta} v(X_1) = V$$

is finite.

As it follows from (10) and (21) $\hat{\theta}_n$ and $\tilde{\theta}_n$ satisfy (27). For each $\theta \in \Theta$ let us define a partition of the real line into $k$ classes-intervals $I_1, I_2, \ldots, I_k$ is defined with boundary points $x_0 = -\infty, x_1 = \gamma_1(\theta_n^*), \ldots, x_{k-1} = \gamma_{k-1}(\theta_n^*), x_k = +\infty$ depending on $\theta_n^*$ such that

$$I_i = I_i(\theta_n^*) = \{x : \gamma_{i-1}(\theta_n^*) \leq x < \gamma_i(\theta_n^*)\}, \quad i = 1, \ldots, k, \quad (29)$$

where $\gamma_i(\theta)$ are continuous functions having partial derivatives.

Let $\nu^* = (\nu_1^*, \ldots, \nu_k^*)^T$ be a vector of frequencies obtained as a result of grouping the observations $X_1, \ldots, X_n$ by the intervals $I_1(\theta_n^*), I_2(\theta_n^*), \ldots, I_k(\theta_n^*)$ with random boundaries, and let

$$p_i(\theta_n^*) = p_i(\theta_n^*; \theta_n^*) = P_{\theta_n^*}(X_1 \in I_i(\theta_n^*) | H_0) =$$

$$= F(\gamma_i(\theta_n^*); \theta_n^*) - F(\gamma_{i-1}(\theta_n^*); \theta_n^*). \quad (30)$$

To test $H_0$ Roy (1956) proposed to consider the statistic

$$X_n^2(\theta_n^*) = X_n^2(\theta_n^*; \theta_n^*) = \sum_{i=1}^{k} \frac{(\nu_i^* - np_i(\theta_n^*))^2}{np_i(\theta_n^*)}. \quad (31)$$

**Theorem.** (ROY(1956), DAIYA & GURLAND (1972))

If $\theta_n^*$ satisfies (27) and the Chernoff-Lehmann conditions hold then

$$\lim_{n \to \infty} P\{X_n^2(\theta_n^*) \geq x | H_0\} = P\{\sum_{i=1}^{k} \lambda_i \xi_i^2 \geq x\}, \quad (32)$$

where $\xi_1, \xi_2, \ldots, \xi_k$ are mutually independent standard normal random variables, $\xi_i \sim N(0, 1), \lambda_1 = \lambda_1(\theta), \ldots, \lambda_k = \lambda_k(\theta)$ are the characteristic roots of
the matrix $D^{-1}\Sigma$, where $D = D(\theta)$ is the diagonal matrix with the elements $p_1(\theta), \ldots, p_k(\theta)$ on the main diagonal,

$$
\Sigma = \Sigma(\theta) = D - pp^T - UTM - MTU + UTVU,
$$

$$
UT = [U_{ij}]_{k \times s}, \quad U_{ij} = U_{ij}(\theta) = \int_{\eta_{i-1}(\theta)}^{\eta_i(\theta)} \frac{\partial f(x; \theta)}{\partial \theta} dx,
$$

$$
MT = [M_{ij}]_{k \times s}, \quad M_{ij} = M_{ij}(\theta) = \int_{\eta_{i-1}(\theta)}^{\eta_i(\theta)} (\theta) v_{ij}(\theta)f(x; \theta) dx.
$$

For example, if $\theta^*_n = \hat{\theta}_n$ is the maximum likelihood estimator, which satisfies (21), then, as it was shown by Roy (1956), in (32) $k - s - 1$ of $\lambda_i$ are equal to 1, one of the $\lambda_i$ is equal to 0, and the remaining $s$ lie between 0 and 1. It is obvious that this is an extension of the Chernoff-Lehmann theorem to the case of random cell boundaries.

**Remark 3.** As it was shown by Dahiya and Gurland (1972) if the density function $f(x; \theta)$ of $X_i$ belongs to a location and scale family

$$
f(x; \theta) = \frac{1}{\sqrt{2\pi}\theta_2} f \left( \frac{x - \theta_1}{\sqrt{\theta_2}} \right), \quad \theta = (\theta_1, \theta_2)^T, \quad |\theta_1| < \infty, \theta_2 > 0,
$$

then it is possible to choose the grouping intervals in order to have the asymptotic distribution of $X^2_R$ independent of $\theta$. For example, let suppose that we test hypothesis $H_0$ according to which $X_i$ follows the normal distribution $N(\theta_1, \theta_2)$,

$$
E\{X_i \mid H_0\} = \theta_1, \quad \text{Var}\{X_i \mid H_0\} = \theta_2,
$$

and let $\theta_n = (\bar{X}_n, s^2_n)^T$, where

$$
\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad s^2_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.
$$

$\theta_n$ is the method of moments estimator for $\theta = (\theta_1, \theta_2)^T$. Since $f(x; \theta)$ belongs to the exponential family (24) of order 2, $s = 2$, the method of moments gives as a result the maximum likelihood estimator, $\theta_n = \hat{\theta}_n$.

If in (29) we choose

$$
\gamma_1(\theta) = \theta_1 + c_1 \sqrt{\theta_2},
$$

and hence $\gamma_1(\theta_n) = \bar{X}_n + c_1 s_n$, then as it was shown by Gurland and Dahiya (1972) (see also Watson (1957),(1958)), the statistic $X^2_R$ is distributed, under $H_0$, in the limit as $n \rightarrow \infty$ as

$$
\chi^2_{k-3} + \lambda_1 \xi^2_1 + \lambda_2 \xi^2_2,
$$

324
where \(\lambda_1\) and \(\lambda_2\) do not depend on \(\theta\). Using some results of Watson (1957), Dahiya and Gurland tabulated the distribution of Roy’s statistic \(X_R^2\) for \(k = 3, 4, \ldots, 15\), for the significance level \(\alpha\) equal to 0.1, 0.05, 0.01, in the case when the constants \(c_i\) in (37) are chosen so that \(p_i(\theta_n) = 1/k\).

**Remark 4.** It is important to remark that when \(\theta\) is unknown and we have to estimate it, the limit distribution of the Pearson’s statistic \(X_n^2(\theta_n^*)\) changes in general in accordance with asymptotical properties of the estimator \(\theta_n^*\) we shall use.

For this reason it has looked reasonable to have a statistic which limit distribution is well-known when we apply the maximum likelihood estimator or anyone BAN estimator. In the papers of Nikulin (1973) (see also, for example, Rao and Robson (1974), Moore and Spruill (1975)), is exposed how to construct a chi-squared test for a continuous distribution (in particular, for the normal distribution and for distributions with shift and scale parameters), based on the statistic \(Y_n^2(\theta_n^*)\), by using any BAN estimator \(\theta_n^*\) of \(\theta\). For example we can take \(\theta_n^* = \hat{\theta}_n\), where \(\hat{\theta}_n\) is the maximum likelihood estimator. We note that the technique of chi-squared tests for the exponential family of distributions of rank one, \(s=1\), and some applications of MVUE’s were exposed by Nikulin and Voinov (1989). Another modification \(W_n^2(\theta_n^*)\), which limit distribution is stable with respect to any statistical method of estimation, providing \(\sqrt{n}\) - consistent estimator \(\theta_n^*\) was proposed by Dzhabaridze and Nikulin (1974). We shall apply the statistic \(Y_n^2\) to test the hypothesis \(H_0\) according to which the distribution of \(X_1\) belongs to the family of the logistic distributions. This topic is studied also by Dudley (1976), Drost (1988).

5. LOGISTIC DISTRIBUTION AND THE CHI-SQUARED GOODNESS-OF-FIT TEST

Let \(X = (X_1, \ldots, X_n)^T\) be a random sample, i.e. \(X_1, \ldots, X_n\) are independent identically distributed random variables. In this section we consider the problem of testing the hypothesis \(H_0\) that the distribution function of \(X_1\) belongs to the family of logistic distributions \(G\left(\frac{x-\mu}{\sigma}\right)\) depending on the shift parameter \(\mu\) and the scale parameter \(\sigma\):

\[
(39) \quad \mathbf{P}\{X_1 \leq x \mid H_0\} = G\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{1 + \exp\left(-\frac{x}{\sqrt{3}}\left(\frac{x-\mu}{\sigma}\right)\right)}, \quad x \in \mathbb{R},
\]

\[
\mu = \mathbf{E}\{X_1 \mid H_0\}, \quad |\mu| < \infty, \quad \sigma^2 = \mathbf{Var}X_1, \quad \sigma > 0.
\]
Under $H_0$ the density function of $X_i$ is

$$
\frac{1}{\sigma} g \left( \frac{x - \mu}{\sigma} \right) = G' \left( \frac{x - \mu}{\sigma} \right) = \frac{\pi}{\sqrt{3}\sigma} \frac{\exp \left( -\frac{\pi}{\sqrt{3}} \frac{x - \mu}{\sigma} \right)}{\left[ 1 + \exp \left( -\frac{\pi}{\sqrt{3}} \frac{x - \mu}{\sigma} \right) \right]^2}, \quad x \in \mathbb{R}^1.
$$

We remark that $g(x)$ is symmetric.

We point out that “Handbook of the logistic distribution” edited by Balakrishnan (1992) was published recently about the theory, the methodology and some applications of the family of logistic distributions, see also, Oliver (1964), Pearl and Reed (1920), Reed and Berkson (1929).

We denote $\theta = (\mu, \sigma^2)^T$, and let $\hat{\theta}_n = (\hat{\mu}_n, \hat{\sigma}_n^2)^T$ be the maximum of likelihood estimator of $\theta$. Since there is no any other sufficient statistic for $\theta$ than the trivial one $X = (X_1, \ldots, X_n)^T$, the maximum likelihood equation has no explicit root. Balakrishnan and Cohen (1990) proposed an approximate solution of the maximum likelihood equations based on a “type II censored sample” of Harter and Moore (1967) (see also Grizzle (1961)). They proved that this approximate solution gives an asymptotically efficient estimator, i.e. asymptotically equivalent to $\hat{\theta}_n$.

Let $\hat{\theta}_n$ be such an estimator. As it follows from (21) the limit covariance matrix of the random vector $\sqrt{n}(\hat{\theta}_n - \theta)$ will be $\Gamma^{-1}$, where

$$
\mathbf{I} = \frac{1}{\sigma^2} ||I_{ij}||_{2\times 2} = \frac{1}{9\sigma^2} \begin{bmatrix}
\pi^2 & 0 \\
0 & \pi^2 + 3
\end{bmatrix},
$$

$I_{11} = \int_{-\infty}^{+\infty} \left[ \frac{g'(x)}{g(x)} \right]^2 g(x)dx = \frac{\pi^2}{9}$,

$I_{22} = \int_{-\infty}^{+\infty} x^2 \left[ \frac{g'(x)}{g(x)} \right]^2 g(x)dx - 1 = \frac{\pi^2 + 3}{9}$,

and since $g(x)$ is symmetric

$$I_{12} = I_{21} = \int_{-\infty}^{+\infty} x \left[ \frac{g'(x)}{g(x)} \right]^2 g(x)dx = 0.$$

Let us fix the vector $p = (p_1, p_2, \ldots, p_k)^T$ of positive probabilities such that

$$p_1 = \ldots = p_k = 1/k,$$

and let

$$y_i = G^{-1}(p_1 + \cdots + p_i) = \frac{\sqrt{3}}{\pi} \ln \left( \frac{i}{k - i} \right), \quad i = 1, \ldots, k-1, \quad y_0 = -\infty, \quad y_k = +\infty.$$

(43)
Further, let \( \nu = (\nu_1, \ldots, \nu_k)^T \) be the frequency vector arising from grouping \( X_1, \ldots, X_n \) over the intervals with random ends

\[ (-\infty, z_1], (z_1, z_2], \ldots, (z_{k-1}, +\infty), \quad \text{where} \quad z_i = z_i(\theta_n) = \mu_n + \sigma_n y_i, \]

and let

\[ a = (a_1, \ldots, a_k)^T, \quad b = (b_1, \ldots, b_k)^T, \quad W^T = -\frac{1}{\sigma} \| a \cdot b \|, \]

where for \( i = 1, 2, \ldots, k \)

\[ a_i = g(y_i) - g(y_{i-1}) = \frac{\pi}{k^2 \sqrt{3}} (k - 2i + 1), \]

\[ b_i = y_i g(y_i) - y_{i-1} g(y_{i-1}) = \frac{1}{k^2} \left[ (i-1)(k-i+1) \ln \frac{k-i+1}{i-1} - i(k-i) \ln \frac{k-i}{i} \right], \]

\[ \alpha(\nu) = k \sum_{i=1}^{k} a_i \nu_i = \frac{\pi}{\sqrt{3} k} \left[ (k+1)n - 2 \sum_{i=1}^{k} i \nu_i \right], \]

\[ \beta(\nu) = k \sum_{i=1}^{k} b_i \nu_i = \frac{1}{k} \sum_{i=1}^{k-1} (\nu_{i+1} - \nu_i) i(k-i) \ln \frac{k-i}{i}, \]

\[ \lambda_1 = I_{11} - k \sum_{i=1}^{k} a_i^2 = \frac{\pi^2}{9k^2}, \quad \lambda_2 = I_{22} - k \sum_{i=1}^{k} b_i^2. \]

Since \( g \) is symmetric we have \( a_1 + a_2 + \cdots + a_k = b_1 + b_2 + \cdots + b_k = 0 \). Let

\[ B = D - p^T p - W^T I^{-1} W, \]

where \( D \) is the diagonal matrix with the elements \( 1/k \) on the main diagonal. The matrix \( B \) does not depend on \( \theta \), and \( \text{rank} B = k - 1 \), i.e. the matrix \( B \) is singular, while the matrix \( \tilde{B} \), obtained as a result of deleting the last row and column in \( B \), has an inverse

\[ \tilde{B}^{-1} = A + A \tilde{W}^T (I - \tilde{W} A \tilde{W}^T)^{-1} \tilde{W} A, \]

where \( A = \tilde{D}^{-1} + 11^T / p_k \), \( \tilde{D}^{-1} \) is a diagonal matrix with elements \( \frac{1}{\tilde{p}_1}, \ldots, \frac{1}{\tilde{p}_{k-1}} \) on the main diagonal, \( 1 = 1_{k-1} \) is the vector of dimension \( (k-1) \), all elements of which are equal to 1, \( \tilde{W} \) is a matrix obtained from \( W \) by deleting the last column. Since the vector \( \tilde{\nu} = (\nu_1, \ldots, \nu_{k-1})^T \) is asymptotically normally distributed with parameters

327
(51) \( \mathbf{E} \tilde{\nu} = n\bar{p} + O(1) \) and \( \mathbf{E}(\tilde{\nu} - n\bar{p})^T(\tilde{\nu} - n\bar{p}) = n\tilde{\mathbf{B}} + O(1), \)

where \( \bar{p} = (p_1, \ldots, p_{k-1})^T \), we obtain the next result.

Theorem 1.

The statistic

\[
Y_n^2 = \frac{1}{n}(\tilde{\nu} - n\bar{p})^T \tilde{\mathbf{B}}^{-1}(\tilde{\nu} - n\bar{p}) = X_n^2 + \frac{\lambda_1 \beta^2(\nu) + \lambda_2 \alpha^2(\nu)}{n\lambda_1 \lambda_2}
\]

has, as \( n \to \infty \), chi-squared limit distribution with \((k - 1)\) degrees of freedom, where

\[
X_n^2 = \sum_{i=1}^{k} \frac{(\nu_i - np_i)^2}{np_i} = \frac{k}{n} \sum_{i=1}^{k} \nu_i^2 - n.
\]

Remark 5. We consider the hypothesis \( H_\eta \) according to which \( X_i \) follows \( G(\frac{x-\mu}{\sigma}, \eta) \), where \( G(x, \eta) \) is continuous, \( |x| < \infty, \eta \in \mathbf{H} \subset \mathbb{R}^1 \), \( G(x, 0) = G(x) \), and \( \eta = 0 \) is a limit point of \( \mathbf{H} \). Let us assume also, that

\[
\frac{\partial}{\partial x} G(x, \eta) = g(x, \eta) \quad \text{and} \quad \frac{\partial}{\partial \eta} g(x, \eta) |_{\eta=0} = \Psi(x),
\]

exist, where \( g(x, 0) = g(x) = G'(x) \). In this case if \( \frac{\partial^2 g(x, \eta)}{\partial \eta^2} \) exists and is continuous for all \( x \) in the neighbourhood of the \( \eta = 0 \), then

\[
P\{z_{i-1} < X_i \leq z_i \mid H_\eta\} = p_i + \eta c_i + o(\eta),
\]

where

\[
c_i = \int_{z_{i-1}}^{z_i} \Psi(x) dx, \quad i = 1, \ldots, k,
\]

and finally, in the limit as \( n \to \infty \) the statistic \( Y_n^2 \) has noncentral chi-squared distribution with \((k - 1)\) degrees of freedom and with non-centrality parameter \( \lambda \):

\[
\lim_{n \to \infty} P\{Y_n^2 \geq x \mid H_\eta\} = P\{\chi_{k-1}^2(\lambda) \geq x\},
\]

where

\[
\lambda = \sum_{i=1}^{k} \frac{c_i^2}{p_i} + \frac{\lambda_2 \alpha^2(c) + \lambda_1 \beta^2(c)}{\lambda_1 \lambda_2}, \quad c = (c_1, c_2, \ldots, c_k)^T,
\]

\( p, \alpha(c), \beta(c), \lambda_1, \lambda_2 \) are given by (42),(46),(47),(48) respectively.
6. EXAMPLE (ABOUT A CHOICE OF INTERVALS)

1. Simple hypotheses. Suppose that we want to test the simple hypothesis
   \( H_0 \) according to which
   \[
   P\{X_1 \leq x \mid H_0\} = G(x)
   \]
   against the simple hypothesis \( H_1 \):
   \[
   P\{X_1 \leq x \mid H_1\} = \Phi(x)
   \]
   (\( \Phi(x) \) being the standard normal distribution function).

   It would be possible to use the Neymann and Pearson test, yet it would
   entail large calculations. We shall then try to adapt a chi-squared test.

   Let \( (X_1, \ldots, X_n)^T \) be a sample of mutually independent identically distributed
   random variables with \( E(X_1) = 0, \ Var(X_1) = 1 \). Before to construct a chi-
   square test for testing \( H_0 \) against \( H_1 \) we shall do one remark on the cells' choice.
   As the test only compares the respective frequencies on the cells, it is worthwhile
   to choose those when both density curves are the most distant, that is to
   say those got by their junctions.

   ![Graph](image)

   The representative curves of the density function \( g(x) \) of \( L(0, 1) \) distribution
   and the density function \( \varphi(x) \) of the standard normal \( N(0, 1) \) distribution have
   four symmetric points of intersection:
   \[
   x_1 = -x_6 = -\infty, \quad x_2 = -x_5, \quad x_3 = -x_4,
   \]
   where \( \varphi(x_i) = g(x_i) \). Let
   \[
   I_i = \{x_{i-1} < x \leq x_i\}, \quad i = 1, 2, 3, 4, 5,
   \]
   329
be the such definite intervals of grouping the data. We can even improve the power of the test by considering the cells only differentiated by the relative positions of the two curves:

\[ J_1 = I_1 \cup I_2 \cup I_5 = \{ x : g(x) > \varphi(x) \} \text{, } g(x) \text{ higher that } \varphi(x) \text{,} \]
\[ J_2 = I_2 \cup I_4 = \{ x : \varphi(x) > g(x) \} \text{, } \varphi(x) \text{ higher that } g(x) \text{.} \]

Let us remark that if we consider, for example, an interval \(-2 < x \leq 0\) for grouping the data with one point of intersection \(x_3\) on the inside, as shown on the schema, it is clear that in this case the two probabilities

\[ \mathbf{P}\{ -2 < X_i \leq 0 \mid H_0 \} \text{ and } \mathbf{P}\{ -2 < X_i \leq 0 \mid H_1 \} \]

will be approximately equal, and it will be difficult to decide which of both hypotheses is true and hence the power of the test using such interval will be low. A test using the intervals \(\{ -2 < x \leq x_3 \}\) and \(\{ x_3 < x \leq 0 \}\) (or \(x_3 < x \leq x_4\)) will obviously be more powerful. From this remark we obtain the next

**Proposition**

To test \(H_0\) against \(H_1\), the chi-squared test using \(I_1, I_2, I_3, I_4, I_5\) is less powerful as the one using two cells: \(J_1\) and \(J_2\).

\[ \text{a) Let } (\nu_1, \nu_2, \ldots, \nu_5)^T \text{ be the observed frequencies of the sample in the intervals } I_1, \ldots, I_5 \text{ respectively. Then} \]

\[ \mathbf{P}\{ X_1 \in I_i \mid H_0 \} = p_i^{(0)}, \quad i = 1, 2, \ldots, 5 \]

(with \(p_1^{(0)} = p_5^{(0)} = 0.0155; p_2^{(0)} = p_4^{(0)} = 0.207; p_3^{(0)} = 0.555\)). In our case we shall use the standard statistic of Pearson:

\[ X_n^2 = \sum_{i=1}^{5} \frac{(\nu_i - np_i^{(0)})^2}{np_i^{(0)}}. \]

Under \(H_0\), \(X_n^2\) is asymptotically \(\chi_4^2\) and we shall reject \(H_0\) if \(X_n^2 \geq c_{4,\alpha}\). Since

\[ \mathbf{P}\{ X_n^2 \geq c_{4,\alpha} \mid H_0 \} \approx \mathbf{P}\{ \chi_4^2 \geq c_{4,\alpha} \} = \alpha. \]

The power of this test is \(P_5 = \mathbf{P}\{ X_n^2 \geq c_{4,\alpha} \mid H_1 \} \). Or,

\[ \mathbf{P}\{ X_1 \in I_i \mid H_1 \} = p_i^{(1)}, \quad i = 1, 2, \ldots, 5, \]

\(p_1^{(1)} = p_5^{(1)} = 0.011; p_2^{(1)} = p_4^{(1)} = 0.234; p_3^{(1)} = 0.51\).
Under $H_1$, $X^2_n$ has asymptotically a non-central chi-square distribution with 4 degrees of freedom and the parameter of noncentrality $\lambda$:

$$\lambda = n \sum_{i=1}^{5} \frac{(p_i^{(0)} - p_i^{(1)})^2}{p_i^{(0)}}, \quad (\lambda \sim 0.0133n).$$

Using the approximation of Patnaik we can compute

$$P\{X^2_{k-1}(\lambda) \geq c_\alpha\} \sim P\{X^2_{k-1} \geq c_\alpha \left(1 - \frac{\lambda}{k-1}\right)\}. \tag{63}$$

For example, for $\alpha = 0.05$ we have $c_{4,0.05} = 9.49$, from which it follows that

in order to have $P_5 > 0.5 \ 0.74 \ 0.9 \ 0.99$,

it is necessary to take $n > 195 \ 240 \ 269 \ 292$

respectively.

b) Let $\nu$ be the observed frequency on $J_1$. We have

$$P\{X_1 \in J_1 \mid H_0\} = \omega^{(0)} \text{ with } \omega^{(0)} = 0.586$$

and

$$P\{X_1 \in J_1 \mid H_1\} = \omega^{(1)} \text{ with } \omega^{(1)} = 0.532. \tag{64}$$

To test $H_0$ against $H_1$ we shall use again the standard statistic of Pearson

$$X^2_n = \frac{(\nu - n\omega^{(0)})^2}{n\omega^{(0)}(1 - \omega^{(0)})}.$$ 

Under $H_0$ the statistic $X^2_n$ is distributed ($n \to \infty$) asymptotically as $\chi^2_1$. We shall reject $H_0$ if $X^2_n \geq c_{1,\alpha}$, since $P\{\chi^2_1 \geq c_{1,\alpha}\} = \alpha$ and the power of this test is

$$P_2 = P\{X^2_n \geq c_{1,\alpha} \mid H_1\}.$$ 

Under $H_1$, $X^2_n$ has asymptotically ($n \to \infty$) a non-central chi-square distribution with one degree of freedom and the parameter of noncentrality

$$\lambda' = \frac{[\omega^{(0)} - \omega^{(1)}]^2}{\omega^{(0)}(1 - \omega^{(0)})}, \quad (\lambda' \sim 0.012n).$$

Using the same approximation (63):

$$P\{\chi^2_1(\lambda') \geq c_\alpha\} \sim P\{\chi^2_1 \geq c_{1,\alpha}(1 - \lambda')\},$$

331
we have $c_{1, a} = 3.84$ for $a = 0.05$, from which it follows that

\[
\text{in order to have } P_2 > 0.5 \ 0.75 \ 0.9 \ 0.99,
\]

it is necessary to take $n > 74 \ 82 \ 83 \ 84$

respectively. One can see that in the case b) the chi-square test based on $J_1$ and $J_2$ is more powerful, then in the case a). The same approach may be used for composite hypotheses (see, for example, Nikulin & Voinov, 1989).

2. Composite hypotheses. Let now $X = (X_1, \ldots, X_n)^T$ be a sample, $E(X_1) = \mu$ and $\text{Var}(X_1) = \sigma^2$, $\theta = (\mu, \sigma^2)^T$, $\theta$ is unknown, and let us test $H_0$ according to which $X_1$ follows a logistic distribution (39):

\[
P\{X_1 \leq x \mid H_0\} = G(x, \theta) = G\left(\frac{x - \mu}{\sigma}\right)
\]

against the hypothesis of normality $H_1$ according to which

\[
P\{X_1 \leq x \mid H_1\} = \Phi\left(\frac{x - \mu}{\sigma}\right).
\]

Let $\hat{\theta}_n$ be an estimator which satisfies to (21). According to the precedent study and to §5, we shall take the two cells with random boundaries (41):

\[
J_1(\hat{\theta}_n) = \left[-\infty, -\hat{\sigma}_n x_5 + \hat{\mu}_n\right] \cup \left[-\hat{\sigma}_n x_3 + \hat{\mu}_n, \hat{\sigma}_n x_3 + \hat{\mu}_n\right] \cup \left[\hat{\sigma}_n x_5 + \hat{\mu}_n, +\infty\right],
\]

\[
J_2(\hat{\theta}_n) = \mathbb{R}^1 \setminus J_1(\hat{\theta}_n).
\]

Let $\nu = (\nu, n - \nu)^T$ be the vector of frequencies obtained in the result of the groupement of the sample $X = (X_1, \ldots, X_n)^T$ into the intervals $J_1(\hat{\theta}_n)$ and $J_2(\hat{\theta}_n)$. According to definitions (43) to (47) we calculate

\[
a_1' = -\sigma \frac{\partial}{\partial \mu} P\{X_i \in J_1(\theta) \mid H_0\} = a_1 + a_3 + a_5 = 0 = a_2',
\]

\[
b_1' = -\sigma \frac{\partial}{\partial \sigma} P\{X_i \in J_1(\theta) \mid H_0\} = b_1 + b_3 + b_5 = 2[x_3 g(x_3) - x_5 g(x_5)] = -b_2' \approx 0.296,
\]

\[\lambda_1 = I_{11} - \frac{a_1'^2}{\omega(0)} - \frac{a_2'^2}{1 - \omega(0)} = I_{11} = \frac{\pi^2}{9},\]

332
\[
\lambda_2 = I_{22} - \frac{b'^2_1}{\omega(0)} - \frac{b'^2_2}{1 - \omega(0)} = \frac{\pi^2 + 3}{9} - \frac{b'^2_1}{\omega(0)(1 - \omega(0))},
\]
\[
\lambda_3 = I_{12} - 0 = 0,
\]
\[
\alpha(\nu) = \frac{a'_1 \nu}{\omega(0)} + \frac{a'_2 (n - \nu)}{1 - \omega(0)} = 0,
\]
\[
\beta(\nu) = \frac{b'_1 \nu}{\omega(0)} + \frac{b'_2 (n - \nu)}{1 - \omega(0)} = b'_1 \frac{(\nu - n\omega(0))}{\omega(0)(1 - \omega(0))},
\]
\[
X_n^2 = \frac{(\nu - n\omega(0))^2}{n\omega(0)(1 - \omega(0))},
\]
and finally we obtain the statistic

(67) \[ Y_n^2 = X_n^2 + \frac{\beta(\nu)}{n\lambda_2} = X_n^2 + \frac{\frac{b'^2_1}{\omega(0)(1 - \omega(0))}}{\frac{\pi^2 + 3}{9} - \frac{b'^2_1}{\omega(0)(1 - \omega(0))}} X_n^2 \sim 1.34 X_n^2,
\]

for testing the composite hypothesis \(H_0\), given by (39), against the hypothesis of normality \(H_1\), and

\[ P\{Y_n^2 \geq x \mid H_0\} \rightarrow P\{X_1^2 \geq x\}, \quad (n \rightarrow \infty). \]

**ACKNOWLEDGEMENT**

We would like to thank Professors S. Kotz, A. Rukhin, C. Huber, N. Balakrishnan and J. Antoch and referees for helpful discussion, advice and encouragement during the writing of the paper.
REFERENCES


