DOI 10.1007/s10958-023-06539-2
Journal of Mathematical Sciences, Vol. 273, No. 5, July, 2023

CHI-SQUARED TEST FOR HYPOTHESIS TESTING OF HOMOGENEITY

M. S. Ermakov* UDC 519.2

We provide necessary and sufficient conditions for uniform consistency of nonparametric sets of alternatives of chi-squared test for testing the hypothesis of homogeneity. The number of cells of the chi-squared test increases with sample size growth. The nonparametric sets of alternatives can be defined in terms of distribution functions or densities. Bibliography: 18 *titles.*

1. Introduction

For goodness-of-fit testing, the asymptotic behavior of chi-squared tests was rather comprehensively explored both in the case of a fixed number of cells and of a number of cells increasing with growth of sample size [4–7, 13, 14, 16–18]. In the present paper, we explore the last setup for hypothesis testing on homogeneity.

Now, for a setup of goodness-of-fit testing, we remind the results [6, 7] that are akin to those obtained in the present paper.

Let $X_1, \dots, X_n$ be a sample of i.i.d. random variables with values on the interval $[0, 1]$ and having a c.d.f. F_n. Denote by $\widehat{F}_n$ an empirical c.d.f. of the sample. Denote by $\Im$ a set of all distribution functions. Denote by F_0 c.d.f. of the uniform distribution on the interval $[0, 1]$. In goodness-of-fit testing, we explore the problem of testing the hypothesis $\mathbb{H}_0 : F_n = F_0$ versus the alternatives $\mathbb{H}_n : F_n \in \Psi_n \subset \Im$, where Ψ_n is a nonparametric set of alternatives.

Denote by $T_n(\widehat{F}_n)$ test statistics of chi-squared tests and by $T_n(F)$, $F \in \Im$, functionals generating test statistics $T_n(\widehat{F}_n)$.

For goodness-of-fit testing, we have showed in [6] that the sequences of chi-squared tests having an increasing number of cells with growth of sample size are uniformly consistent on the sets of alternatives $\Im(b_n) = \{F : T_n(F) > b_n, F \in \Im\}$. Here $b_n > 0$, $n = 1, 2, 3, \dots$ depend on the number of cells and on the sample size n. Thus, a sequence of sets of the alternatives $\Omega_n \subset \Im$ is uniformly consistent if and only if $\Omega_n \subset \Im(b_n)$. In [7], we explored type II error probabilities of chi-squared tests having cells of equal length and a number of cells growing with an increasing sample size if alternatives are defined in terms of densities. We described all uniformly consistent sequences of simple alternatives for this setup.

The aim of the present paper is to provide similar results for hypothesis testing of homogeneity and to describe all uniformly consistent sequences of alternatives defined in terms of the distribution functions or densities. The problem is more difficult than for goodness-of-fit testing [6, 7]. For hypothesis testing of homogeneity, the answer depends on the distribution functions of two samples. Note that the problem of hypothesis testing of homogeneity has been intensively studied in recent papers [2, 8–11, 19].

Let the interval $[0, 1]$ be divided into $m = m_n$ subintervals

$$?I_{nj} = [e_{nj}, e_{n,j+1}), \quad p_{nj} = e_{n,j+1} - e_{nj} > 0, \quad e_{n0} = 0, \quad e_{nm} = 1,$$

*Institute for Problems in Mechanical Engineering, RAS; St.Petersburg State University, St.Petersburg, Russia, e-mail: `erm2512@gmail.com`.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 501, 2021, pp. 160–180. Original article submitted November 25, 2022.

$1 \le j \le m = m_n$, where $m_n \to \infty$ as $n \to \infty$. The functional T_n generating chi-squared test statistics for goodness-of-fit testing equals

$$T_n(F - F_0) = n \sum_{j=1}^{m} \frac{(r_{nj} - p_{nj})^2}{p_{nj}},$$

where $r_{nj} = F(e_{nj}) - F(e_{n,j-1})$ for all $1 \le j \le m_n$ and $F_0(x) = x$, $x \in [0,1]$.

Then $T_n(\widehat{F}_n - F_0)$ is chi-squared test statistics.

For test K_n, we denote $\alpha(K_n)$ its type I error probability and $\beta(K_n, F_n)$ its type II error probability for alternative F_n.

Let S_n be sequence of test statistics. We say that the sequence of sets of alternatives $\Psi_n \subset \Im$ is uniformly consistent for test statistics S_n if for tests K_n, $\alpha(K_n) = \alpha(1 + o(1))$, $0 < \alpha < 1$, generated test statistics S_n, we have

$$\limsup_{n\to\infty} \sup_{F\in\Psi_n} \beta(K_n, F) < 1 - \alpha.$$

We use similar notation and terminology for the problem of testing of the hypothesis of homogeneity as well.

Recall that for goodness-of-fit testing the chi-squared tests are uniformly consistent for sets of alternatives $\Im(b_n)$ with special order of asymptotics of b_n. Moreover [6], for any sequence of simple alternatives $F_n \in \Im$, for the type II error probabilities $\beta(K_n, F_n)$ of tests K_n, $\alpha(K_n) = \alpha(1 + o(1))$, $0 < \alpha < 1$, generated by test statistics $T_n(\widehat{F}_n - F_0)$, we have

$$\beta(K_n, F_n) = \Phi(x_\alpha - 2^{-1/2} m_n^{-1/2} T_n(F_n - F_0)) + o(1) \tag{1.1}$$

as $n \to \infty$. Here x_α is defined by the equation $\alpha = 1 - \Phi(x_\alpha)$, where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-2\, t^2/2\}\, d\, t$$

is a standard normal distribution function, $x \in \mathbb{R}^1$.

Such asymptotics of the type II error probabilities and the asymptotic minimaxity of chi-squared tests [6] substantiates the consideration of these tests as one of the implementations of the distance method.

In the present paper, we establish similar results for testing of hypothesis of homogeneity with sets of alternatives generated by the differences of distribution functions F_n and G_{l_n} of two independent samples. We assume additionally that $\mathbb{L}_2$-norms of densities of distribution functions F_n or G_n are uniformly bounded by a constant. It turns out that uniform consistency of sets of alternatives is given by the value of functionals T_n defined by differences of distribution functions $F_n - G_{l_n}$. This allows to extend the results of [7] on this setup and to establish necessary and sufficient conditions of uniform consistency of sets of alternatives defined in terms of densities. We implement the technique of [2, 7, 14] for the interpretation of the results in terms of smoothness of densities.

We use letters c and C as a generic notation for positive constants. Denote $[a]$ the whole part of a real number a. For any two sequences of the positive real numbers a_n and b_n, $a_n \asymp b_n$ implies $c < a_n/b_n < C$ for all n and $a_n = o(b_n)$ implies $a_n/b_n \to 0$ as $n \to \infty$. For any complex number z, denote by $\bar{z}$ a complex conjugate number.

2. Main results

2.1. Setup. In comparison with goodness-of-fit-testing, the problem is more difficult. We have two samples $X_1, \dots, X_n$ and $Y_1, \dots, Y_{l_n}$ of i.i.d. random variables taking values on the interval $[0,1]$ and having distribution functions F_n and G_{l_n}, respectively. Thus, the criterion of uniform consistency we search in terms of differences $G_n - F_n$ and nuisance parameter F_n or G_n.

Denote by $\Im \times \Im$ a set of all pairs of distribution functions (F, G).

On the set $\Im \times \Im$, we define a functional

$$T_{1n}(F-G) = nm\sum_{j=1}^{m}(r_{nj}-s_{nj})^2, \qquad (F,G)\in \Im\times\Im,$$

where $s_{nj} = G(e_{nj}) - G(e_{n,j-1})$ for all $1 \le j \le m_n$.

Denote by $\widehat{G}_{l_n}(x)$ an empirical distribution function of sample $Y_1, \dots, Y_{l_n}$.

Denote $a_n = \frac{n}{l_n}$ and assume that $0 < c < a_n < C < \infty$.

Chi-squared test statistics has the following form:

$$T_{1n}(\widehat{F}_n - \widehat{G}_{l_n}) = nm\sum_{j=1}^{m}(\widehat{r}_{nj}-\widehat{s}_{nj})^2,$$

where $\widehat{s}_{nj} = \widehat{G}_{l_n}(e_{nj}) - \widehat{G}_{l_n}(e_{n,j-1})$ for all $1 \le j \le m_n$.

We explore a more general setup as well. For test statistics

$$T_{2n}(\widehat{F}_n - \widehat{G}_{l_n}) = n\sum_{j=1}^{m} g_{nj}\frac{(\widehat{r}_{nj}-\widehat{s}_{nj})^2}{p_{nj}},$$

generated by the functional

$$T_{2n}(F-G) = n\sum_{j=1}^{m} g_{nj}\frac{(r_{nj}-s_{nj})^2}{p_{nj}}, \quad 0 < c < g_{jn} < C < \infty,$$

similar results are established.

Proof is provided for test statistics

$$T_{n}(\widehat{F}_n - \widehat{G}_{l_n}) = n\sum_{j=1}^{m} \frac{(\widehat{r}_{nj}-\widehat{s}_{nj})^2}{p_{nj}},$$

generated by the functional

$$T_{n}(F-G) = n\sum_{j=1}^{m} \frac{(r_{nj}-s_{nj})^2}{p_{nj}}.$$

For test statistics $T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})$, the reasoning is almost the same and, therefore, the differences are not indicated.

We assume that the nuisance parameter F_n has a density $f_n(x) = \frac{dF_n(x)}{dx}$, $x \in [0,1]$, and a priori information is provided that there is a positive constant C such that we have

$$F_n \in \Xi(C) = \left\{ F \,:\, \|f\|^2 < C,\ f(x) = \frac{dF(x)}{dx},\ F \in \Im \right\},$$

where $\|f\|^2 = \int\limits_0^1 f^2(x)\, d\,x$.

The assumption on the distribution function F_n could be naturally replaced with the same assumption on the distribution function G_n.

The main term of the asymptotics for variance of the chi-squared test statistics is significantly simplified if we assume additionally

$$F_n \in \Xi_{1n} = \left\{ F \, : \, \sup_{x\in[0,1]} |f(x)| < c_n m_n^{1/2}, \; f(x) = \frac{dF(x)}{dx}, \; F \in \Im \right\},$$

where $c_n \to 0$ as $n \to \infty$.

For a sequence $b_n > 0$, for $i = 1, 2$, define sets of alternatives

$$\Psi_i(b_n) = \big\{(F, G) \, : \, T_{\mathrm{in}}(F - G) \geq b_n, \; (F, G) \in \Im \times \Im\big\}.$$

We establish uniform consistency of the test statistics $T_{in}(\widehat{F}_n - \widehat{G}_{l_n})$, $i = 1, 2$, in the problems of hypothesis testing

$$\mathbb{H}_0 \; : \; F_n(x) = G_{l_n}(x), \qquad x \in [0, 1]$$

versus the alternatives

$$\mathbb{H}_n \; : \; (F_n, G_{l_n}) \in \Psi_i(b_n) \cap \Xi(C)$$

for the sequences b_n, satisfying

$$0 < \liminf_{n\to\infty} m_n^{-1/2} b_n \leq \limsup_{n\to\infty} m_n^{-1/2} b_n < \infty. \tag{2.1}$$

Assume that for all j, $1 \leq j \leq m_n$, we have

$$0 < c < \; m_n \, p_{nj} \; < C_1 < \infty \tag{2.2}$$

for some positive constants c and C_1.

Assume also that $m_n = o(n)$ as $n \to \infty$.

Proof of the theorems is based on the methods proposed in [6] for the study of chi-squared tests for goodness-of-fit testing.

On the set $\Im \times \Im$, we define the functional

$$T_n(F - G) = n \sum_{j=1}^{m} \left(\int_0^1 \phi_{nj}(x) d(F(x) - G(x)) \right)^2 p_{nj}^{-1},$$

where $\phi_{nj}(x) = \mathbf{1}_{\{x\in I_{nj}\}} - p_{nj}$, $x \in [0, 1]$, $1 \leq j \leq m_n$ and $\mathbf{1}_{\{A\}}$ denotes the indicator of event A. After that we explore test statistics as the test statistics generated by this functional.

This approach allows one to prove easily the results following the numerous results in [7, 8, 15, 19], established for nonparametric hypothesis testing on a density based on the expansions of series of orthogonal functions. However, in this case, functions ϕ_{nj} are not orthogonal.

In this notation, the test statistics $T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})$ have the following form:

$$T_{2n}(\widehat{F}_n - \widehat{G}_{l_n}) = n \sum_{j=1}^{n} g_{nj} \left(\int_0^1 \phi_{nj}(x) d(\widehat{F}_n(x) - \widehat{G}_{l_n}(x)) \right)^2 p_{nj}^{-1}.$$

Note that if the hypothesis holds, $\mathbf{E}[T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})]$ depends on the unknown distribution function $F_n = G_{l_n}$ and, in the case of an alternative, $\mathbf{E}[T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})]$ depends on both unknown distribution functions F_n and G_{l_n}. This is caused by the term

$$W_n = n \sum_{j=1}^{n} g_{nj} \left(\int_0^1 \phi_{nj}^2(x) d\widehat{F}_n(x) + \int_0^1 \phi_{nj}^2(x) d\,\widehat{G}_{l_n}(x) \right) p_{nj}^{-1}$$

included in test statistics.

To delete this dependence, we subtract this term from the test statistics in one of the setups. Note that we do not have such an influence of W_n on the test statistics in the case of the test statistics $T_{1n}(\widehat{F}_n - \widehat{G}_{l_n})$.

Without loss of generality, we can assume that the distribution functions F_n and G_{l_n} have densities

$$f_n(x) = 1 + \sum_{j=1}^{m} \theta_{nj}\, \phi_{nj}(x), \quad x \in [0,1]$$

and

$$g_{l_n}(x) = 1 + \sum_{j=1}^{m} \tau_{nj}\, \phi_{nj}(x), \quad x \in [0,1],$$

respectively, and

$$\sum_{j=1}^{m} \theta_{nj} p_{nj} = 0, \qquad \sum_{j=1}^{m} \tau_{nj} p_{nj} = 0.$$

Denote $\eta_{nj} = \theta_{nj} - \tau_{nj}$.

2.2. Test statistics T_{1n}. Denote $M_{1n}(\eta) = nm \sum\limits_{j=1}^{m} p_{nj}^2 \eta_{nj}^2$ and denote

$$\sigma_{1n}^2 = 2m^2 \sum_{j=1}^{m} p_{nj}^2 (1 + \theta_{nj} + a_n + a_n \tau_{nj})^2.$$

Lemma 2.1. *We have*

$$\mathbf{E}[T_{1n}(\widehat{F}_n - \widehat{G}_{l_n})] - (m-1)(1 + a_n) = nm \sum_{j=1}^{m} p_{nj}^2 \eta_{nj}^2 (1 + o(1)), \tag{2.3}$$

and

$$\begin{aligned} \mathbf{Var}[T_{1n}(\widehat{F}_n - \widehat{G}_{l_n})] &= \sigma_{1n}^2 (1 + o(1)) \\ &+ nm^2 \sum_{j=1}^{m} p_{nj}^3 (1 + \theta_{nj} + a_n + a_n \tau_{nj}) \eta_{nj}^2 (1 + o(1)) \doteq \sigma_{11n}^2 (1 + o(1)) \end{aligned} \tag{2.4}$$

as $n \to \infty$.

Note that the second term in the right-hand side of (2.4) equals zero if the hypothesis holds. Thus, we have an interesting situation. The sets of alternatives are so rich that the asymptotic variance for the alternatives approaching the hypothesis is significantly different from the asymptotic variance for the hypothesis.

By (3.25), $\sigma_{11n}^2 - \sigma_{1n}^2 > 0$. If $F_n \in \Xi_{2n}$, then $\sigma_{11n}^2 - \sigma_{1n}^2 = o(\sigma_{1n}^2)$ as $n \to \infty$.

Note that we can substitute the estimators

$$\widehat{\theta}_{nj} = \frac{\widehat{r}_{nj}}{p_{nj}} - 1, \quad \text{and} \quad \widehat{\tau}_{nj} = \frac{\widehat{s}_{nj}}{p_{nj}} - 1$$

of parameters θ_{nj} and τ_{nj} into (2.4). After that, as we show, we get a consistent estimator

$$\widehat{\sigma}_{1n}^2 = 2m^2 \sum_{j=1}^{m} (\widehat{r}_{nj} + a_n \widehat{s}_{nj})^2$$

of variance σ_{1n}^2.

Other methods of the estimation of variance are considered in [1, 8, 9].

Define tests

$$K_{1n} = \mathbf{1}_{\{\widehat{\sigma}_{1n}^{-1}(T_{1n}(\widehat{F}_n - \widehat{G}_{l_n}) - m(1+a_n)) > x_\alpha\}},$$

where x_α is defined by the equation $1 - \alpha = \Phi(x_\alpha)$, $0 < \alpha < 1$.

Theorem 2.1. *Assume* (2.1), (2.2) *and let* $m_n = o(n)$ *as* $n \to \infty$. *Then the sequence of sets of alternatives* $\Psi_{1n}(b_n) \cap \Xi(C)$ *is uniformly consistent for a sequence of tests* K_{1n}, *generated by tests statistics* $T_{1n}(\widehat{F}_n - \widehat{G}_{l_n})$.

We have $\alpha(K_{1n}) = \alpha(1 + o(1))$ *and*

$$\beta(K_{1n}, \Psi_{1n}(b_n)) = \Phi(\sigma_{11n}^{-1}(\sigma_{1n} x_\alpha - M_{1n}(\eta))) + o(1) \tag{2.5}$$

as $n \to \infty$.

2.3. Test statistics T_{2n} and T_{3n}. Denote $M_{2n}(\eta) = n \sum\limits_{j=1}^{m} g_{nj} p_{nj} \eta_{nj}^2$ and denote

$$\sigma_{2n}^2 = 2 \sum_{j=1}^{m} g_{nj}^2 (1 + \theta_{nj} + a_n + a_n \tau_{nj})^2.$$

We show that

$$\widehat{\sigma}_{2n}^2 = 2 \sum_{j=1}^{m} g_{nj}^2 p_{nj}^{-2} (\widehat{r}_{nj} + a_n \widehat{s}_{nj})^2$$

is the consistent estimator of σ_{2n}^2.

Tests for the test statistics $T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})$ are based on the following asymptotics.

Lemma 2.2. *We have*

$$\mathbf{E}[T_{2n}(\widehat{F}_n - \widehat{G}_{l_n})] = M_{2n}(\eta)(1 + o(1)) + \mathbf{E}[W_n], \tag{2.6}$$

$$\begin{aligned} \mathbf{E}[W_n] = \sum_{j=1}^{m} g_{nj} ((1 - p_{nj} + \theta_{nj}(1 - p_{nj}) - p_{nj}\theta_{nj}^2) \\ + a_n(1 - p_{nj} + \tau_{nj}(1 - p_{nj}) - p_{nj}\tau_{nj}^2)) \doteq e_n, \end{aligned} \tag{2.7}$$

$$\begin{aligned} \mathbf{Var}[T_{2n}(\widehat{F}_n - \widehat{G}_n)] &= \sigma_{2n}^2(1 + o(1)) \\ &+ n \sum_{j=1}^{m} g_{nj}^2 p_{nj} (1 + \theta_{nj} + a_n + a_n \tau_{nj}) \eta_{nj}^2 (1 + o(1)) \doteq \sigma_{21n}^2 (1 + o(1)) \end{aligned} \tag{2.8}$$

as $n \to \infty$.

As we show, if $m_n = o(n^{2/3})$, then we have

$$e_n = \sum_{j=1}^{m} g_{nj}(1 + a_n + \theta_{nj} + \tau_{nj}) + O(1). \tag{2.9}$$

Note that we can substitute the estimators $\widehat{\theta}_{nj}$ and $\widehat{\tau}_{nj}$ of parameters θ_{nj} and τ_{nj} into (2.7) and obtain a consistent estimator $\widehat{e}_n$ for e_n.

Define tests

$$K_{2n} = \mathbf{1}_{\{\widehat{\sigma}_{2n}^{-1}(T_{2n}(\widehat{F}_n - \widehat{G}_{l_n}) - \widehat{e}_n) > x_\alpha\}},$$

where x_α is defined by the equation $1 - \alpha = \Phi(x_\alpha)$, $0 < \alpha < 1$.

Theorem 2.2. *Assume* (2.1), (2.2) *and let* $m_n = o(n^{2/3})$, *as* $n \to \infty$. *Then the sequence of sets of alternatives* $\Psi_{2n}(b_n) \cap \Xi(C)$ *is uniformly consistent for a sequence of tests* K_{2n}.

Let $m_n = o(n)$ *as* $n \to \infty$ *and let there be a constant* C *such that* $\|g_n\| < C$, $g_n(x) = \frac{dG_n(x)}{dx}$, $x \in [0,1]$. *Then the sequence of sets of alternatives* $\Psi_{2n}(b_n) \cap \Xi(C)$ *is uniformly consistent.*

We have $\alpha(K_{2n}) = \alpha(1 + o(1))$ *and*

$$\beta(K_{2n}, F_n, G_n) = \Phi(\sigma_{21n}^{-1}(\sigma_{2n} x_\alpha - M_{1n}(\eta))) + o(1) \tag{2.10}$$

as $n \to \infty$.

In [1,6,8,9,19], the authors delete the version of term W_n from the version of test statistics T_{2n} for similar setups of nonparametric hypothesis testing and obtain the results for such modified test statistics.

Define test statistics

$$T_{3n}(\widehat{F}_n - \widehat{G}_{l_n}) = T_{2n}(\widehat{F}_n - \widehat{G}_{l_n}) - W_n.$$

Define the corresponding test of hypothesis testing

$$K_{3n} = \mathbf{1}_{\{\widehat{\sigma}_{2n}^{-1} T_{3n}(\widehat{F}_n - \widehat{G}_{l_n}) > x_\alpha\}},$$

where x_α is defined by the equation $1 - \alpha = \Phi(x_\alpha)$, $0 < \alpha < 1$.

Theorem 2.3. *Assume* (2.1), (2.2) *and let* $m_n = o(n)$ *as* $n \to \infty$. *Then the sequence of sets of alternatives* $\Psi_{2n}(b_n) \cap \Xi(C)$ *is uniformly consistent for the sequence of tests* K_{3n}.

We have $\alpha(K_{3n}) = \alpha(1 + o(1))$ *and*

$$\beta(K_{3n}, F_n, G_n) = \Phi(\sigma_{21n}^{-1}(\sigma_{2n} x_\alpha - M_{2n}(\eta))) + o(1) \tag{2.11}$$

as $n \to \infty$.

2.4. Hypothesis testing on homogeneity in terms of densities. The asymptotics of the type II error probabilities in (2.5), (2.10), and (2.11) are exactly the same as the asymptotics [6, 7] of chi-squared tests for goodness-of-fit testing (1.1). By this reason, for the sets of alternatives defined in terms of densities, we can transfer necessary and sufficient conditions [7] of uniform consistency for the problem of goodness-of-fit testing to the case of hypothesis homogeneity.

Assume that the distribution functions F_n and G_{l_n} have densities f_n, g_{l_n}, respectively, and $F_n \in \Xi(C)$, $G_{l_n} \in \Xi(C)$. Denote $h_n = f_n - g_{l_n}$.

We explore the problem of testing the hypothesis

$$\mathbb{H}_0 \,:\, h_n(x) = 0, \qquad x \in [0,1],$$

versus the alternatives

$$\mathbb{H}_n \,:\, h_n \in \Omega_n \subset \Gamma,$$

where $\Gamma = \{h \,:\, h = \frac{d(F-G)(x)}{dx}, \|h\| < \infty, F \in \Xi(C)\}$.

For this setup, all assertions of Theorem 6.1 in [7] hold if we replace the densities $1 + f_n$ with functions h_n. All requirements in condition B that the functions $1 + f_n$ and the functions specially defined by the function $1 + f_n$ should be densities are replaced with the requirement that the functions h_n and the functions similarly specially defined by h_n should be differences of two densities. In particular, this holds if the densities of distribution functions F_n and G_n satisfy B in [7].

This version of Theorem 6.1 in [7] holds only for the sequence of simple alternatives h_n, $\|h_n\| \asymp n^{-r}$, $\frac{1}{4} < r < \frac{1}{2}$, $m_n \asymp n^{2-4r}$. In this setup, following [7], we assume that cells of chi-squared tests have the same length.

3. Proof of Theorems

3.1. Estimate of $\mathbf{E}[T_n]$. The reasoning is provided for the test statistics T_n.

Alternatives satisfy the inequality

$$T_n(F_n - G_n) = n \sum_{j=1}^{m} p_{nj} \eta_{nj}^2 \geq b_n.$$

By $f_n \in \Xi(C)$, we have

$$\sum_{j=1}^{m} p_{nj} \theta_{nj}^2 \leq \|f_n - 1\|^2 < C. \tag{3.1}$$

Lemma 3.1. *For $1 \leq j \leq m$, we have*

$$\mathbf{E}_\theta[\phi_{nj}(X_1)] = \theta_{nj} p_{nj}, \tag{3.2}$$

$$\mathbf{E}_\theta[\phi_{nj}^2(X_1)] = p_{nj}(1 - p_{nj} + \theta_{nj}(1 - 2p_{nj})), \tag{3.3}$$

$$\begin{aligned} \mathbf{E}[\bar{\phi}_{nj_1}^4(X_1)] = p_{nj}(1+\theta_{nj})(1 - 4p_{nj}(1+\theta_{nj}) \\ + 6p_{nj}^2(1+\theta_{nj})^2 - 3p_{nj}^3(1+\theta_{nj})^3) \end{aligned} \tag{3.4}$$

and, for $1 \leq j_1 < j_2 \leq m$, we have

$$\mathbf{E}_\theta[\phi_{nj_1}(X_1)\,\phi_{nj_2}(X_1)] = -p_{nj_1} p_{nj_2}(1 + \theta_{nj}(1 - 2p_{nj}) + \theta_{nj_2}(1 - 2p_{nj_2})), \tag{3.5}$$

$$\begin{aligned} \mathbf{E}_\theta[\bar{\phi}_{nj_1}^2(X_1)\,\bar{\phi}_{nj_2}^2(X_1)] &= p_{nj_1} p_{nj_2}(1+\theta_{nj_1})(1+\theta_{nj_2}) \\ &\quad \times (p_{nj_1}(1+\theta_{nj_1}) + p_{nj_2}(1+\theta_{nj_2}) - 3p_{nj_1}p_{nj_2}(1+\theta_{nj_1})(1+\theta_{nj_2})). \end{aligned} \tag{3.6}$$

Equalities (3.2)–(3.6) are obtained by straightforward calculations and proof is omitted.

Proof of Lemma 2.2. We begin with proof of (2.6). For $x, y \in [0,1]$, denote

$$\bar{\phi}_{nj}(x) = \phi_{nj}(x) - \mathbf{E}_\theta \phi_{nj}(X_1) = \phi_{nj}(x) - \theta_{nj} p_{nj}$$

and

$$\tilde{\phi}_{nj}(y) = \phi_{nj}(y) - \mathbf{E}_\tau \phi_{nj}(Y_1) = \phi_{nj}(y) - \tau_{nj} p_{nj}.$$

Then we have

$$T_n(\widehat{F}_n - \widehat{G}_n) = I_{1n} + I_{2n} + I_{3n} + W_n, \tag{3.7}$$

with

$$I_{1n} = 2\,I_{11n} + 2\,I_{12n} + 2\,I_{13n},$$

where

$$I_{11n} = \sum_{1 \leq i_1 < i_2 \leq n} U_{1n}(X_{i_1}, X_{i_2}), \qquad I_{12n} = \sum_{1 \leq i_1 < i_2 \leq l_n} U_{2n}(Y_{i_1}, Y_{i_2}),$$

and

$$I_{13n} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{l_n} U_{3n}(X_{i_1}, Y_{i_2}),$$

where

$$U_{1n}(X_{i_1}, X_{i_2}) = \sum_{j=1}^{m} \frac{\bar{\phi}_{nj}(X_{i_1})\bar{\phi}_{nj}(X_{i_2})}{n p_{nj}},$$

$$U_{2n}(Y_{i_1}, Y_{i_2}) = \sum_{j=1}^{m} \frac{\tilde{\phi}_{nj}(Y_{i_1})\tilde{\phi}_{nj}(Y_{i_2})}{n p_{nj}},$$

and

$$U_{3n}(X_{i_1}, Y_{i_2}) = \sum_{j=1}^{m} \frac{\bar{\phi}_{nj}(X_{i_1})\tilde{\phi}_{nj}(Y_{i_2})}{np_{nj}}.$$

We have

$$I_{2n} = \sum_{j=1}^{m} \left(\frac{1}{n}\sum_{i=1}^{n} \bar{\phi}_{nj}(X_i) - \frac{1}{l_n}\sum_{i=1}^{l_n} \tilde{\phi}_{nj}(Y_i) \right) \eta_{nj}, \tag{3.8}$$

$$I_{3n} = M_n(\eta) = n\sum_{j=1}^{n} p_{nj}\eta_{nj}^2 = T_n(F_n - G_{l_n}). \tag{3.9}$$

$$W_n = n^{-1}\sum_{j=1}^{m}\sum_{i=1}^{n} \bar{\phi}_{nj}^2(X_i)p_{nj}^{-1} + nl_n^{-2}\sum_{j=1}^{m}\sum_{i=1}^{l_n} \tilde{\phi}_{nj}^2(Y_i)p_{nj}^{-1}. \tag{3.10}$$

We have

$$\mathbf{E}I_{1n} = 0, \qquad \mathbf{E}I_{2n} = 0, \tag{3.11}$$

$$\begin{aligned}
\mathbf{E}[W_n] &= \sum_{j=1}^{m}(1 - p_{nj} + \theta_{nj}(1 - 2p_{nj}) - p_{nj}\theta_{nj}^2) \\
&\quad + nl_n^{-1}\sum_{j=1}^{m}(1 - p_{nj} + \tau_{nj}(1 - 2p_{nj}) - p_{nj}\tau_{nj}^2) \\
&= (1 + a_n)\sum_{j=1}^{m}(1 - p_{nj} + \theta_{nj}(1 - 2p_{nj}) - p_{nj}\theta_{nj}^2) \\
&\quad + O(n^{-1/2}mM_{1n}^{1/2}(\eta))(1 + n^{-1}M_n(\eta))),
\end{aligned} \tag{3.12}$$

because

$$\begin{aligned}
\sum_{j=1}^{m} |\theta_{nj} - \tau_{nj}| &\leq \max_{1\leq j\leq m} p_{nj}^{-1} \sum_{j=1}^{m} p_{nj}|\eta_{nj}| \\
&\leq Cm\left(\sum_{j=1}^{m} p_{nj}\eta_{nj}^2\right)^{1/2}\left(\sum_{j=1}^{m} p_{nj}\right)^{1/2} \leq Cn^{-1/2}mM_{1n}^{1/2}(\eta)
\end{aligned} \tag{3.13}$$

and

$$\begin{aligned}
\sum_{j=1}^{m} |\theta_{nj}^2 - \tau_{nj}^2| &\leq \max_{1\leq j\leq m} p_{nj}^{-1} \sum_{j=1}^{k} p_{nj}|\eta_{nj}|\,(|\theta_{nj}| + |\tau_{nj}|) \\
&\leq Cm^{-1}\left(\sum_{j=1}^{m} p_{nj}\eta_{nj}^2\right)^{1/2}\left(\sum_{j=1}^{m} p_{nj}(\theta_{nj}^2 + \tau_{nj}^2)\right)^{1/2} \\
&\leq Cn^{-1/2}mM_{1n}^{1/2}(\eta)(N_n(\theta) + N_n(\tau))^{1/2} \leq Cn^{-1/2}mM_{1n}^{1/2}(\eta),
\end{aligned} \tag{3.14}$$

because

$$|N_n^{1/2}(\tau) - N_n^{1/2}(\theta)| \leq n^{-1/2}M_n^{1/2}(\eta).$$

Note that the reminder in the right-hand side of (3.12) is $o(m_n)$ as $n \to \infty$, if $m_n = o(n^{2/3})$. $\square$

3.2. Analysis of $\mathbf{Var}[T_n]$. We have

$$\mathbf{Var}[I_{11n}] = V_{11n} + V_{12n}, \tag{3.15}$$

where

$$\begin{aligned} V_{11n} = 2\sum_{j=1}^{m} p_{nj}^{-2}(\mathbf{Var}[\phi_j(X_1)])^2 &= 2\sum_{j=1}^{m}(1 - p_{nj} + \theta_{nj}(1-2p_{nj}) - p_{nj}^2)^2 \\ &= 2\sum_{j=1}^{m}(1+\theta_{nj})^2(1+o(1)) \end{aligned} \tag{3.16}$$

and

$$\begin{aligned} V_{12n} &= 2\sum_{1\le j_1<j_2\le m} p_{nj_1}^{-1}p_{nj_2}^{-1}(\mathbf{Cov}[\phi_{j_1}(X_1), \phi_{j_2}(X_1)])^2 \\ &= 4\sum_{1\le j_1<j_2\le m} p_{nj_1}p_{nj_2}(1+\theta_{nj_1})^2(1+\theta_{nj_2})^2(1+o(1)) \le (C + N_n^2(\theta_n))(1+o(1)). \end{aligned} \tag{3.17}$$

Therefore, we have

$$\mathbf{Var}[I_{11n}] = 2\sum_{j=1}^{m}(1+\theta_{nj})^2(1+o(1)). \tag{3.18}$$

We have

$$\begin{aligned} \mathbf{Var}[I_{12n}] &= 4a_n\sum_{j=1}^{m} p_{nj}^{-2}\mathbf{Var}[\phi_j(X_1)]\mathbf{Var}[\phi_j(Y_1)] \\ &= 4a_n\sum_{j=1}^{m}(1+\theta_{nj})(1+\tau_{nj})(1+o(1)). \end{aligned} \tag{3.19}$$

Arguing similarly to (3.18), we get

$$\mathbf{Var}[I_{13n}] = 2a_n^2\sum_{j=1}^{m}(1+\tau_{nj})^2(1+o(1)). \tag{3.20}$$

We have

$$\mathbf{Cov}[I_{11n}, I_{12n}] = 0, \quad \mathbf{Cov}[I_{11n}, I_{13n}] = 0, \quad \mathbf{Cov}[I_{12n}, I_{13n}] = 0. \tag{3.21}$$

Thus, by (3.18)– (3.21), we get

$$\mathbf{Var}[I_{1n}] = 2\sum_{j=1} m(1 + a_n + \theta_{nj} + a_n\tau_{nj})^2(1+o(1)). \tag{3.22}$$

We have

$$\mathbf{Var}[I_{2n}] = J_{21n} + J_{22n} + J_{23n} + J_{24n}, \tag{3.23}$$

with

$$\begin{aligned} J_{21n} &= 2n^{-1}(n-1)^2\sum_{1\le j_1<j_2\le m}\mathbf{Cov}[\phi_{j_1}(X_1), \phi_{j_2}(X_1)]\eta_{nj_1}\eta_{nj_2} \\ &= 2n^{-1}(n-1)^2\sum_{1\le j_1<j_2\le m} p_{nj_1}p_{nj_2}(1+\theta_{nj_1})(1+\theta_{nj_2})\eta_{nj_1}\eta_{nj_2}(1+o(1)) \\ &\le C\left(\sum_{j=1}^{m} p_{nj}(1+\theta_{nj_1})^2\right)\left(n\sum_{j=1}^{m} p_{nj}\eta_{nj}^2\right) \le C\, M_{1n}(\eta)(1+N_n(\theta)), \end{aligned} \tag{3.24}$$

and

$$
\begin{aligned}
J_{22n} &= n^{-1}(n-1)^2 \sum_{j=1}^{m} \mathbf{Var}[\phi_{nj}(X_1)]\,\eta_{nj}^2 \\
&= n^{-1}(n-1)^2 \sum_{j=1}^{m} p_{nj}(1 - p_{nj} + \theta_{nj}(1-2p_{nj}) - p_{nj}\theta_{nj}^2)\,\eta_{nj}^2 \\
&= n \sum_{j=1}^{m} p_{nj}(1+\theta_{nj})\,\eta_{nj}^2(1+o(1)) = O(m^{1/2}M_{1n}(\eta)),
\end{aligned}
\tag{3.25}
$$

because

$$
\max_{1\le j\le m} |\theta_{nj}|^2 < C\,m\,N_n(\theta) < C\,m. \tag{3.26}
$$

The terms J_{23n} and J_{24n} are estimated similarly to J_{21n} and J_{22n}, respectively. We omit this reasoning.

We have

$$
\mathbf{Var}[W_n] = A_{1n} + A_{2n} + A_{3n} + A_{4n}, \tag{3.27}
$$

where

$$
A_{1n} = n^{-1} \sum_{1\le j_1 < j_2 \le m} \mathbf{E}[\bar{\phi}_{nj_1}^2(X_1)\,\bar{\phi}_{nj_2}^2(X_1)] p_{nj_1}^{-1} p_{nj_2}^{-1} \tag{3.28}
$$

and

$$
A_{2n} = n^{-1} \sum_{j=1}^{m} \mathbf{E}[\bar{\phi}_{nj_1}^4(X_1)] p_{nj}^{-2}. \tag{3.29}
$$

The terms A_{3n} and A_{4n} are estimated similarly to A_{1n} and A_{2n}, respectively. We omit this reasoning.

Using (3.4) and (3.26), we get

$$
\begin{aligned}
A_{1n} &\le n^{-1} \sum_{1\le j_1<j_2\le m} [p_{nj_1}(1+\theta_{nj_1})^2(1+\theta_{nj_2}) + p_{nj_2}(1+\theta_{nj_1})(1+\theta_{nj_2})^2] \\
&\le Cn^{-1} \sum_{j=1}^{m} p_{nj}(1+|\theta_{nj}|)^2 \left(m + \sum_{j=1}^{m} p_{nj}|\theta_{nj}| \right) \\
&\le Cn^{-1}(C+N_n(\theta))(m + m^{1/2}N^{1/2}(\theta)) \\
&\le Cn^{-1}m + Cn^{-1}mN_n(\theta) + Cn^{-1}m^{1/2}N^{3/2}(\theta).
\end{aligned}
\tag{3.30}
$$

Using (3.6) and (3.26), we get

$$
\begin{aligned}
A_{2n} = n^{-1} \sum_{j=1}^{m} p_{nj}^{-1}(1+\theta_{nj})[1 &- 4p_{nj}(1+\theta_{nj}) \\
&+ 6p_{nj}^2(1+\theta_{nj})^2 - 3p_{nj}^3(1+\theta_{nj})^3].
\end{aligned}
\tag{3.31}
$$

We estimate only two terms in A_{2n}. The other two terms are estimated similarly and have a smaller order.

We have

$$
\begin{aligned}
n^{-1}\sum_{j=1}^{m} p_{nj}^{-1}(1+\theta_{nj}) &\le Cn^{-1}m^2\left(1+\sum_{j=1}^{m} p_{nj}|\theta_{nj}|\right) \\
&\le Cn^{-1}m^2\left(1+\left(\sum_{j=1}^{m} p_{nj}\theta_{nj}^2\right)^{1/2}\right) \le Cn^{-1}m^2(1+N_n(\theta)) = o(m)
\end{aligned}
\tag{3.32}
$$

and

$$
n^{-1}\sum_{j=1}^{m} p_{nj}^2(1+\theta_{nj})^4 \le Cn^{-1}m^{-1} + n^{-1}\sum_{j=1}^{m} p_{nj}^2\theta_{nj}^4 \le Cn^{-1}(m^{-1}+N_n^2(\theta)). \tag{3.33}
$$

Therefore, we have

$$
A_{2n} \le Cn^{-1}m^2(1+N_n^{1/2}(\theta)) + n^{-1}N_n^2(\theta). \tag{3.34}
$$

3.3. Consistency of estimators of bias and variance of test statistics T_n. Let us show the consistency of estimators of $\sum_{j=1}^{m} g_{nj}\theta_{nj}$ in (2.7) and (2.8).

We have

$$
\begin{aligned}
\mathbf{Var}\left[\sum_{j=1}^{m} g_{nj}\frac{1}{n}\sum_{i=1}^{n}\frac{\phi_{nj}(X_i)}{p_{nj}}\right] &= \frac{1}{n}\sum_{j=1}^{m} g_{nj}^2\frac{\mathbf{Var}[\phi_{nj}(X_1)]}{p_{nj}^2} \\
&+ \frac{1}{n}\sum_{1\le j_1<j_2\le m} g_{nj_1}g_{nj_2}\frac{\mathbf{Cov}[\phi_{nj_1}(X_1),\phi_{nj_2}(X_1)]}{p_{nj_1}p_{nj_2}} \\
&= \frac{1}{n}\sum_{j=1}^{m} g_{nj}^2\frac{1+\theta_{nj}}{p_{nj}}(1+o(1)) \\
&+ \frac{1}{n}\sum_{1\le j_1<j_2\le m} g_{nj_1}g_{nj_2}(1+\theta_{nj_1}+\theta_{nj_2})(1+o(1)) = o(m),
\end{aligned}
\tag{3.35}
$$

because

$$
n^{-1}\sum_{j=1}^{m}\frac{\theta_{nj}}{p_{nj}} \le Cn^{-1}m^2\sum_{j=1}^{m} p_{nj}\theta_{nj} \le Cn^{-1}m^2N_n^{1/2}(\theta) = o(m) \tag{3.36}
$$

and

$$
n^{-1}m\sum_{j=1}^{m} g_{nj}\theta_{nj} \le Cn^{-1}m\max_{1\le j\le m} p_{nj}^{-1}\sum_{j=1}^{m} p_{nj}\theta_{nj} \le Cn^{-1}m^2N_n^{1/2}(\theta) = o(m). \tag{3.37}
$$

We estimate only one term arising in the estimation of variance. Other terms are estimated similarly.

We have

$$
n^{-4}\mathbf{Var}\left[\sum_{j=1}^{m} g_{nj}^2p_{nj}^{-2}\sum_{1\le i_1<i_2<n}\phi_{nj}(X_{i_1})\phi_{nj}(X_{i_2})\right] \le B_{1n}+B_{2n}, \tag{3.38}
$$

where

$$\begin{aligned} B_{1n} &= Cn^{-2}\sum_{j=1}^{m} p_{nj}^{-4}(\mathbf{Var}[\phi_{nj}(X_1)])^2 \le Cn^{-2}\sum_{j=1}^{m} p_{nj}^{-2}(1+\theta_{nj})^2(1+o(1)) \\ &\le Cn^{-2}\max_{1\le j\le m} p_{nj}^{-3}\sum_{j=1}^{m} p_{nj}(1+\theta_{nj})^2 = O(n^{-2}m^3(1+N_n(\theta)) = o(m) \end{aligned} \tag{3.39}$$

and

$$\begin{aligned} B_{2n} &= Cn^{-2}\sum_{1\le j_1<j_2\le m} \frac{(\mathbf{Cov}[\phi_{nj_1}(X_1),\phi_{nj_2}(X_1)])^2}{p_{nj_1}^2 p_{nj_2}^2} \\ &\le Cn^{-2}\sum_{1\le j_1<j_2\le m}(1+\theta_{nj_1}+\theta_{nj_2})^2 \le cn^{-2}m^2 \\ &+ Cn^{-2}m\max_{1\le j\le m} p_{nj}^{-1}\left(\left|\sum_{j=1}^{m} p_{nj}\theta_{nj}\right| + \sum_{j=1}^{m} p_{nj}\theta_{nj}^2\right) \\ &\le Cn^{-2}m^2(1+N_n(\theta)) = o(1). \end{aligned} \tag{3.40}$$

We have provided the estimates of variance in the case of sample $X_1,\dots,X_n$. In the case of sample $Y_1,\dots,Y_n$, we have the same situation. In this case, in the final estimates, $N_n(\theta_n)$ is replaced with $N_n(\tau_n)$.

We have

$$N_n^{1/2}(\tau_n) \le N_n^{1/2}(\theta_n) + n^{-1/2}M_n^{1/2}(\eta_n). \tag{3.41}$$

Since $N_n^{1/2}(\theta_n) < C < \infty$, it suffices to show that if, in the final estimates, we replace $N_n^{1/2}(\theta_n)$ with $n^{-1}M_n(\eta_n)$, then these estimates have a smaller order than $M_n^2(\eta_n)$.

Note that in (3.12)–(3.40), the largest orders in the final estimates for a distribution function G_{l_n} are $M_{1n}(\eta_n)N_n(\tau_n)$ (the version of (3.24)), $n^{-1}m^2N_n^{1/2}(\tau_n)$ (the version of (3.30)) and $n^{-1}N_n^2(\tau_n)$ (the version of (3.34)).

It suffices to estimate only $n^{-1}m^2N_n^{1/2}(\tau_n)$. We have

$$n^{-3/2}m^2M_n^{1/2}(\eta_n)M_n^{-2}(\eta_n) = O(n^{-3/2}m_n^2m_n^{-3/4}) = o(1), \tag{3.42}$$

if $m_n^{-1/2}M_n(\eta_n)\to\infty$ as $n\to\infty$.

Thus,

$$M_n(\eta_n)\widehat{\sigma}_n \to_P \infty, \tag{3.43}$$

if $m_n^{-1/2}M_n(\eta_n)\to\infty$ as $n\to\infty$.

Therefore, the type II error probabilities of tests K_n tend to zero if $N_n(\tau_n)\to\infty$ as $n\to\infty$.

3.4. Asymptotic normality of test statistics T_n. It suffices to prove the asymptotic normality of statistics I_{1n}. For alternatives, we can assume $(F_n, G_{l_n}) \in \Xi_n(C)\times\Xi_n(C)$ for some $C > 0$. Otherwise, the type II error probabilities tend to zero. The statistics I_{1n} are not U–statistics. However, we can implement the same martingale technique to the proof of asymptotic normality [3, 6, 12, 15] and to get a similar result as in the case of goodness-of-fit tests [6, 15]. Since in [1] similar reasoning for testing of hypothesis of homogeneity is omitted for test statistics based on $\mathbb{L}_2$-norm of kernel estimator of density, we outline this reasoning in the present paper for chi-squared tests.

The reasoning is provided for $l_n \le n$. The case $l_n \ge n$ is similar.

Define the martingale W_{ni}, $1 \le i \le n + l_n$, by induction. We put

$$W_{n1} = U_{1n}(X_1, X_1), \quad \text{and} \quad W_{n2} = U_{2n}(Y_1, Y_1) + U_{3n}(X_1, Y_1).$$

If i is odd, we put $j = [i/2]$ and

$$W_{ni} = \sum_{s=1}^{j} U_{1n}(X_j, X_s) + \sum_{s=1}^{j-1} U_{3n}(X_j, Y_s).$$

If i is even, $i \le 2l_n$, we put $j = i/2$ and

$$W_{ni} = \sum_{s=1}^{j} U_{2n}(Y_j, Y_s) + \sum_{s=1}^{j-1} U_{3n}(X_s, Y_j).$$

If $i \ge 2l_n$, we put $j = i - l_n$ and

$$W_{ni} = \sum_{s=1}^{j} U_{1n}(X_j, X_s) + \sum_{s=1}^{l_n} U_{3n}(X_j, Y_s).$$

We can implement the reasoning of [12] to this martingale and obtain a similar result.

Denote

$$V_{1n}(x,y) = \mathbf{E}[U_{1n}(x, X_1)U_{1n}(y, X_1)], \quad V_{2n}(x,y) = \mathbf{E}[U_{1n}(x, Y_1)U_{1n}(y, Y_1)],$$
$$V_{3n}(x,y) = \mathbf{E}[U_{3n}(X_1, x)U_{3n}(X_1, y)], \quad V_{4n}(x,y) = \mathbf{E}[U_{3n}(x, Y_1)U_{3n}(y, Y_1)].$$

Theorem 3.1. *The statistics I_{1n} is asymptotically normal with zero mean and variance σ_1^2 if we have*

$$\begin{aligned}\lim_{n\to\infty} m_n^{-1}[\mathbf{E}[V_{1n}^2(X_1,X_2) + V_{2n}^2(Y_1,Y_2) + V_{3n}^2(X_1,X_2) + V_{4n}^2(Y_1,Y_2)]\\ + n^{-1}\mathbf{E}[U_{1n}^4(X_1,X_2) + U_{2n}^4(Y_1,Y_2) + U_{3n}^4(X_1,Y_1)]] = 0.\end{aligned} \tag{3.44}$$

Proof of the theorem almost repeats the reasoning for the proof of the asymptotic normality in [12] and is omitted.

The verification of (3.44) practically does not differ from the verification of similar conditions in the case of goodness-of-fit testing [6]. Moreover, most of the estimates for proof of (3.44) and the estimates in [6] coincide. Thus, we omit this reasoning.

This research was supported by the RFFI (project 20-01-00273).

Translated by M. S. Ermakov.

REFERENCES

1. N. Anderson, P. Hall, and D. Titterington, "Two-sample test statistics for measuring discrepancies between two multivariate probability density functions using kernel-based density estimates," *J. Multivariate Anal.*, **50**, 41–54 (1994).
2. E. Arias-Castro, B. Pelletier, and V. Saligrama, "Remember the curse of dimensionality: the case of goodness-of-fit testing in arbitrary dimension," *Journal of Nonparametric Statistics*, **30**, 448–471 (2018).
3. B. M. Brown, "Martingale central limit theorems," *Ann. Math. Statist*, **42**, 59–66 (1971).
4. A. R. Barron, "Uniformly powerful goodness of fit tests," *Ann. Statist.*, **17**, 107–124 (1989).
5. D. M. Chibisov, "Asymptotic optimality of the chi-square test with large number of degrees of freedom within the class of symmetric tests," *Math. Methods Statist.*, **1**, 55–82 (1992).
6. M. S. Ermakov, "Asymptotic normality of chi-squared tests," *Theory Probab. Appl. Theory Probab. Appl.*, **42**, No. 4, 589–610 (1998).
7. M. S. Ermakov, "On uniform consistency of nonparametric tests. I.," *Zap. Nauchn. Semin. POMI*, **495**, 147–176 (2020). `https://arxiv.org/abs/2003.05619`.

8. M. Fromont, B. Laurent, M. Lerasle, and P. Reynaud-Bouret, “Kernels based tests with nonasymptotic bootstrap approaches for two-sample problem,” *JMLR: Workshop and Conference Proceedings*, **23**, 23–41 (2012).
9. M. Fromont, B. Laurent, and P. Reynaud-Bouret, “The two-sample problem for Poisson processes: Adaptive tests with a nonasymptotic wild bootstrap approach,” *Ann. Statist.*, **41**, 1431–1461 (2013).
10. A. Gretton, K. Borgwardt, M. Rasch, B. Scheolkopf, and A. Smola, “A kernel two-sample test,” *J. Machine Learning Research*, **13**, 723–773 (2012).
11. A. Gretton, D. Sejdinovic, H. Strathmann, S. Balakrishnan, M. Pontil, K. Fukumizu, and B. K. Sriperumbudur, “Optimal kernel choice for large-scale two-sample tests,” *Advances Neural Information Processing systems*, 1205–1213 (2012).
12. P. Hall, “Central limit theorem for integrated square error of multivariate nonparametric density estimators,” *J. Multivariate Analysis*, **14**, 1–16 (1984).
13. G. I. Ivchenko and Yu. I. Medvedev, “Decomposable statistics and the hypotheses testing for groupped data,” *Theory Probab. Appl.*, **25**, No. 3, 540–551 (1981).
14. Yu. I. Ingster, “On the Comparison of Minimax Properties of Kolmogorov, χ^2 and ω^2-Tests,” *Theory Probab. Appl.*, **32**, No. 2, 346–350 (1987).
15. Yu. I. Ingster and I. A. Suslina, “Nonparametric goodness-of-fit testing under Gaussian models,” *Lect. Notes Statist.*, **169** (2002).
16. H. B. Mann and A. Wald, “On the choice of the number of intervals in the application of chi-squared test,” *Ann. Math. Statist.*, **13**, 306–318 (1942).
17. C. Morris, “Central limit theorems for multinomial sums,” *Ann. Statist.*, **3**, 165–188 (1975).
18. J. Robins, L. Li, E. T. Tchetgen, and Aad van der Vaart, “Asymptotic normality of quadratic estimators,” *Stochastic Processes and their Applications*, **126**, 3733–3759 (2015).
19. T. Li and M. Yuan, “On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives,” p. 42 (2019).