N E W S
OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 5, Number 297 (2014), 42 – 48

UDK 519.224

TESTING A DIFFERENTIAL CONDITION AND LOCAL NORMALITY OF DENSITIES

K. T. Mynbaev[1], A. S. Aipenova[2]

[1]Kazakh British Technical University, Almaty, Kazakhstan,
RSE «Institute of Mathematics and Mathematical Modeling», Almaty, Kazakhstan;
[2]Kazakh National University after al-Farabi, Almaty,
RSE «Institute of Mathematics and Mathematical Modeling», Almaty, Kazakhstan.
E-mail: kairat_mynbayev@yahoo.com, a.aipenova@mail.ru

Key words: testing, local normality test, alternative hypothesis, null hypothesis, asymptotic normality.

Abstract. In this paper, we consider testing if a density satisfies a differential equation. This result can be applied to see if a density belongs to a particular family of distributions. For example, the standard normal density $f(t)=(2\pi)^{-1/2}\exp(-t^2/2)$ satisfies the differential equation $f'(t)+tf(t)=0$. If a density satisfies this equation at that point *t*, then it is called *locally standard normal* at that point. Thus, there is a practical need to test whether a density satisfies a certain differential equation. We consider the expression $F(x)=\sum_{l=0}^{L} g_l(x) f^{(l)}(x)$. We can test the null hypothesis H_0: *f* satisfies the equation $F(x) = 0$ against the alternative hypothesis H_a: $F(x) \neq 0$. The testing procedure is accompanied by an asymptotic normality statement.

1. Introduction. In statistics, normality tests are used to determine if a data set is well-modeled by a normal distribution and to compute how likely it is for a random variable underlying the data set to be normally distributed. Nevertheless normality tests are useful in many areas of forecasting and econometric inference as complements to other diagnostic tests. The tests are a form of model selection, and can be interpreted several ways, depending on one's interpretations of probability, for example, in frequentist statistics, statistical hypothesis testing, data are tested against the null hypothesis that it is normally distributed.

In this paper, we propose a local normality test. Consider, for example, the standard normal density $f(t)=(2\pi)^{-1/2}\exp(-t^2/2)$. It satisfies a differential equation $f'(t)+tf(t)=0$. The general solution of this equation is $f(t)=c\exp(-t^2/2)$, and if it is to be a density, one has to put $c=(2\pi)^{-1/2}$. We say that a density *f* is *locally standard normal* at point *t* if it satisfies the above differential equation at that point. Thus, there is a practical need to test whether a density satisfies a certain differential equation. The testing procedure is accompanied by an asymptotic normality statement.

There are a variety of statistical tests designed specifically to test the normality of data distribution. Different tests of normality often produce different results. The latest references concerning normality testing include [1–7]. According to [1], the Shapiro-Wilk test has the best power for a given level of significance, followed closely by the Anderson-Darling test, Lilliefors test and Kolmogorov-Smirnov test. [2] showed that the Jarque-Bera test is superior in power to its competitors for symmetric distributions with medium up to long tails and for slightly skewed distributions with long tails. The power of the Jarque-Bera test is poor for distributions with short tails, especially if the shape is bimodal, sometimes the test is even biased. According to [3], for testing other distributions, the statistics based on generalized sample spacings and the modified Anderson–Darling statistic provide the most powerful tests. [6] compared the power of six formal tests of normality and showed that D'Agostino-Pearson test achieves the highest power under all conditions for large sample size. Most existing tests are based on some global properties of normal distributions. Our test is local. Both global and local approaches have their advantages and deficiencies. The main difference between the global and local approaches consists in the amount of calculation: rejecting normality locally is enough to reject it globally.

2. Main results. Let $X_1, X_2, \dots X_n$ be independent identically distributed observation from distribution having unknown density f. The Rosenblat-Parzen estimator for the density f evaluated at $x \in R$ is defined by

$$\hat{f}_R(x) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{h}K\left(\frac{x-X_j}{h}\right),$$

where $h > 0$ is a bandwidth and K is a kernel on R satisfying $\int_R K(t)dt = 1$ Denote $\alpha_j(K) = \int_R x^j K(x)dx$ the j-th moment of K and let K be a kernel of order q, that is $\alpha_j(K) = 0$, $j = 1,\dots,q-1$, $\alpha_q(K) \neq 0$

If f and K are l times continuously differentiable, then differentiation of $\hat{f}_R(x)$ leads to the estimator of $f^{(l)}(x)$

$$\hat{f}^{(l)}(x) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{h^{l+1}}K^{(l)}\left(\frac{x-X_j}{h}\right). \tag{1}$$

In asymptotic statements the sample size n tends to infinity and the bandwidth h depends on n but this dependence usually is not reflected in the notation.

Now we turn to the testing for local normality. More generally, consider the expression

$$F(x) = \sum_{l=0}^{L} g_l(x) f^{(l)}(x), \tag{2}$$

where $\{g_l(x)\}$ are given functions and the senior coefficient g_L is different from zero at the given point x. We can test

a) the null hypothesis H_0: f satisfies the equation $F(x) = 0$ against

b) the alternative hypothesis H_a: $F(x) \neq 0$.

It is convenient to use the differential operator D defined by $(Df)(x) = F(x)$. Since the derivative $\hat{f}^{(l)}(x)$ is estimated by (1), it is natural to estimate $F(x)$ by

$$\hat{F}_h(x) = \sum_{l=0}^{L} g_l(x)\hat{f}_h^{(l)}(x,K) = \frac{1}{n}\sum_{j=1}^{n}\sum_{l=0}^{L}\frac{g_l(x)}{h^{l+1}}K^{(l)}\left(\frac{x-X_j}{h}\right). \tag{3}$$

As one can see from part (a) of the next theorem, under the null hypothesis it also makes sense to consider the random variable $\hat{G}_h(x) = \hat{F}_h(x)/h$. Provided that

$$f(x)g_L(x) \neq 0, \tag{4}$$

let Ψ denote a normal variable distributed as $N\left(0, f(x)\left[\int g_L(x)K^{(L)}(t)\right]^2 dt\right)$.

Assumption 1.

a) f is infinitely differentiable.

b) K has l continuous derivatives and $\max_{l=0,\dots,L} \| K^{(l)} \|_{C(R)} < \infty$.

Theorem 1. Suppose that Assumption 1 holds. Then the following statements are true:

a) The bias of (3) is given by

$$E\hat{F}_h(x) - F(x) = g_L(x)\frac{f^{(q+L)}(x)}{q!}(-h)^q\alpha_q(K) + O(h). \tag{5}$$

Consequently, under H_0

$$E\hat{F}_h(x) = O(h^q) \tag{6}$$

If, however, $E\hat{F}_h(x) \to const \neq 0$, as $h \to 0$, then $F(x) \neq 0$ and H_0 can be rejected.

b) If $nh^{2L+1} \to \infty$ and (4) holds, then under the null $p\lim \hat{F}_h(x) = 0$ (this equation is preferable to (6) because in practice $E\hat{F}_h(x)$ is unknown).

c) If $nh \to \infty$ and (4) holds, then

d) $$\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x) - E\hat{F}_h(x)\right] \xrightarrow{d} \Psi. \quad (7)$$

If, in addition, $nh^{2L+3} \to 0$, then

$$\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x) - F(x)\right] \xrightarrow{d} \Psi$$

e) If $nh \to \infty$, and $nh^{2L+3} \to 0$ and (4) holds, then under the null $\left(nh^{2L+1}\right)^{1/2} \hat{G}_h(x) \xrightarrow{d} \Psi$.

Proof: a) Denoting

$$\omega_j = \sum_{l=0}^{L} \frac{g_l(x)}{h^{l+1}} K^{(l)}\left(\frac{x - X_j}{h}\right), j = 1, \ldots, n$$

we rewrite (3) as

$$\hat{F}_h(x) = \frac{1}{n}\sum_{j=1}^{n} \omega_j, \quad (8)$$

$$E\hat{F}_h(x) = \frac{1}{n}\sum_{j=1}^{n}\left(\sum_{l=0}^{L} E\frac{g_l(x)}{h^{l+1}} K^{(l)}\left(\frac{x - X_j}{h}\right)\right) = \frac{1}{n}\sum_{j=1}^{n} E\omega_j = E\omega_1. \quad (9)$$

Assumption 2.

To justify integration by parts below, for any $h > 0$ and $j = 0, \ldots, l-1, l \geq 1$

$$\lim_{s \to -\infty} K^{(j)}(-s) f^{(l-1-j)}(x + sh) = \lim_{s \to +\infty} K^{(j)}(-s) f^{(l-1-j)}(x + sh) = 0$$

Using (9),

$$E\omega_1 = E\hat{F}_h(x) = \frac{1}{n}\sum_{j=1}^{n}\left(\sum_{l=0}^{L} E\frac{g_l(x)}{h^{l+1}} K^{(l)}\left(\frac{x - X_j}{h}\right)\right) = \sum_{l=0}^{L}\frac{g_l(x)}{h^{l+1}}\int_R K^{(l)}\left(\frac{x - s}{h}\right) f(s)ds =$$

(changing variables)

$$= \sum_{l=0}^{L}\frac{g_l(x)}{h^l}\int_R K^{(l)}(-t) f(x + ht)dt = \sum_{l=0}^{L} g_l(x)\left(\frac{1}{h^l}\int_R K^{(l)}(-t) f(x + ht)dt\right) =$$

(then Assumption 2 allows us to integrate l times by parts)

$$= \sum_{l=0}^{L} g_l(x)\left(-\frac{1}{h^l} K^{(l-1)}(-t) f(x + ht)\right)\Bigg|_{-\infty}^{+\infty} + \sum_{l=0}^{L} g_l(x)\left(\frac{1}{h^{l-1}}\int_R K^{(l-1)}(-t) f'(x + ht)dt\right)$$

$$= \sum_{l=0}^{L} g_l(x)\left(-\frac{1}{h} K(-t) f^{(l-1)}(x + ht)\right)\Bigg|_{-\infty}^{+\infty} + \sum_{l=0}^{L} g_l(x)\left(\int_R K(-t) f^{(l)}(x + ht)dt\right) = \ldots =$$

$$= \sum_{l=0}^{L} g_l(x)\left(\int_R K(-t) f^{(l)}(x + ht)dt\right)$$

This integral typically is not analytically solvable, so we approximate it using Taylor expansion of $f^{(l)}(x + ht)$ in the argument ht, which is valid as $h \to 0$. For a q-th order kernel we take the expansion out to the q-th term

$$f^{(l)}(x + ht) = \sum_{i=0}^{\infty}\frac{f^{(i+l)}(x)}{i!}(ht)^i = f^{(l)}(x) + \frac{f^{(q+l)}(x)}{q!}(ht)^q + o(h^{q+1})$$

Using $\int_{-\infty}^{+\infty} K(t)dt = 1$ and $\alpha_q(K) = \int_R x^q K(x)dx$, we get

$$E\omega_1 = \sum_{l=0}^{L} g_l(x) f^{(l)}(x) + \sum_{l=0}^{L} g_l(x) \frac{f^{(q+l)}(x)}{q!}(-h)^q \left(\int_R K(-t)(-t)^q dt \right) + o(h^{q+1}) =$$

$$= (Df)(x) + g_L(x) \frac{f^{(q+L)}(x)}{q!}(-h)^q \alpha_q(K) + o(h^{q+1}). \quad (10)$$

Equation (5) follows from (3) and (10)

$$E\hat{F}_h(x) - F(x) = g_L(x) \frac{f^{(q+L)}(x)}{q!}(-h)^q \alpha_q(K) + o(h^{q+1}).$$

The rest of part a) is an obvious consequence of (5).

b) We need an asymptotic expression for the variance of $\hat{F}_h(x)$. By the i.i.d. assumption

$$\operatorname{var}\left(\hat{F}_h(x)\right) = \frac{1}{n^2}\sum_{j=1}^{n} \operatorname{var}\left(\omega_j\right) = \frac{1}{n}\left[E\omega_1^2 - \left(E\omega_1\right)^2\right].$$

We need to evaluate $E\omega_1^2$ and $\left(E\omega_1\right)^2$.

$$E\omega_1^2 = \sum_{l,m=0}^{L} \frac{g_l(x) g_m(x)}{h^{l+m+2}} EK^{(l)}\left(\frac{x - X_1}{h}\right) K^{(m)}\left(\frac{x - X_1}{h}\right) =$$

$$= \sum_{l,m=0}^{L} \frac{g_l(x) g_m(x)}{h^{l+m+2}} \int_R K^{(l)}\left(\frac{x-s}{h}\right) K^{(m)}\left(\frac{x-s}{h}\right) f(s)ds =$$

$$= \sum_{l,m=0}^{L} \frac{g_l(x) g_m(x)}{h^{l+m+1}} \int_R K^{(l)}(-t) K^{(m)}(-t) f(x+ht)dt =$$

$$= \frac{g_L^2(x)}{h^{2L+1}}\left\{ f(x)\int_R \left(K^{(L)}(t)\right)^2 dt \right\} + \frac{g_L^2(x)}{h^{2L+1}} \sum_{\substack{0 \le l,m \le L \\ l+m \le 2L-1}} \frac{g_l(x) g_m(x)}{g_L^2(x)} h^{2L-l-m} \left[f(x) \int_R K^{(l)}(-t) K^{(m)}(-t) dt \right] =$$

$$= \frac{g_L^2(x)}{h^{2L+1}}\left\{ f(x)\int_R \left(K^{(L)}(t)\right)^2 dt + O(h) \right\}. \quad (11)$$

Combining (10) and (11), we obtain the expression for the variance

$$\operatorname{var}\left(\hat{F}_h(x)\right) = \frac{g_L^2(x)}{nh^{2L+1}}\left\{ f(x)\alpha_0\left[K^{(L)}\right]^2 + O(h) \right\} - \frac{1}{n}\left[F(x) + O(h)\right]^2 =$$

$$= \frac{1}{nh^{2L+1}}\left\{ f(x)\alpha_0\left[g_L(x)K^{(L)}\right]^2 + O(h) \right\}. \quad (12)$$

If H_0 holds then by (5) $E\hat{F}_h(x) \to 0$, $h \to 0$ and by the Chebyshov inequality (12) implies

$$P\left(\left|\hat{F}_h(x) - E\hat{F}_h(x)\right| \ge \varepsilon\right) \le \frac{1}{\varepsilon^2}\operatorname{var}\left(\hat{F}_h(x)\right) \to 0, h \to 0, \text{ for any } \varepsilon > 0.$$

Hence, $\hat{F}_h(x) = \left[\hat{F}_h(x) - E\hat{F}_h(x)\right] + E\hat{F}_h(x) \xrightarrow{P} 0$.

c) Let us prove convergence in distribution of the standardized version $S_n = \left[\hat{F}_h(x) - E\hat{F}_h(x)\right] / \left[\operatorname{var}(\hat{F}_h(x))\right]^{1/2}$ of $\hat{F}_h(x)$. Using (8), we have

$$S_n=\frac{1}{n}\sum_{j=1}^{n}\frac{\omega_j-E\omega_j}{\left[\operatorname{var}(\hat{F}_h(x))\right]^{1/2}}=\sum_{j=1}^{n}X_{nj}$$

where $X_{nj}=\omega_j-E\omega_j/\left[\operatorname{var}(\hat{F}_h(x))\right]^{1/2}$.

It is easy to see that by the i.i.d property

$$EX_{nj}=0,\quad \operatorname{var}(X_{nj})=\frac{\operatorname{var}(\omega_j)}{n^2\operatorname{var}(\hat{F}_h(x))}=\frac{1}{n},\quad \operatorname{var}(S_n)=1.$$

Alternatively, using the notation in the Lindeberg-Feller theorem [7], we can rewrite

$$\mu_{nj}=0,\quad \sigma_{nj}=\frac{1}{n},\quad \sigma_n=1.$$

Let F_{nj} be the distribution function of X_{nj}. Since X_{nj} are i.i.d., all F_{nj} coincide with F_{n1} and the Lindeberg function takes the form

$$\lambda\equiv\frac{1}{\sigma_n^2}\sum_{j=1}^{n}\int_{|x|>\varepsilon}x^2dF_{nj}=n\int_{|x|>\varepsilon}x^2dF_{n1}(x)\le\frac{n}{\varepsilon^{\delta}}\int|x|^{2+\delta}dF_{n1}(x)=$$
$$=\frac{n}{\varepsilon^{\delta}}E|X_{n1}|^{2+\delta}=\frac{E|\omega_1-E\omega_1|^{2+\delta}}{\varepsilon^{\delta}n^{1+\delta}\left[\operatorname{var}(\hat{F}_h(x))\right]^{1+\delta/2}}. \tag{13}$$

By Holder's inequality

$$\left(E|\omega_1-E\omega_1|^{2+\delta}\right)^{\frac{1}{2+\delta}}\le 2\left(E|\omega_1|^{2+\delta}\right)^{\frac{1}{2+\delta}}\le$$

(plugging ω_1 in and applying Minkowski's inequality)

$$\le 2\sum_{l=0}^{L}\frac{|g_l(x)|}{h^{l+1}}\left[E\left|K^{(l)}\left(\frac{x-X_1}{h}\right)\right|^{2+\delta}\right]^{\frac{1}{2+\delta}}=2\sum_{l=0}^{L}\frac{|g_l(x)|}{h^{l+1}}\left[\int_R\left|K^{(l)}\left(\frac{x-s}{h}\right)\right|^{2+\delta}f(s)ds\right]^{\frac{1}{2+\delta}}=$$
$$=2\sum_{l=0}^{L}\frac{|g_l(x)|}{h^{l+1}}\left[\int_R\left|K^{(l)}(t)\right|^{2+\delta}f(x+ht)dt\right]^{\frac{1}{2+\delta}}\le 3\frac{|g_L(x)|}{h^{L+1}}\left\{\left[f(x)\int_R\left|K^{(L)}(t)\right|^{2+\delta}dt+O(h)\right]^{\frac{1}{2+\delta}}\right\}.$$

Therefore,

$$\left(E|\omega_1-E\omega_1|^{2+\delta}\right)\le 3^{2+\delta}\left(\frac{|g_L(x)|^{2+\delta}}{h^{L+1-1/(2+\delta)}}\left\{f(x)\int_R\left|K^{(L)}(t)\right|^{2+\delta}dt+O(h)\right\}\right). \tag{14}$$

Combing (12) and (14), we get

$$\lambda\le\frac{n^{1+\delta/2}h^{(2L+1)(1+\delta/2)}}{\varepsilon^{\delta}n^{1+\delta}h^{L+1-1/(2+\delta)}}\frac{\left[3|g_L(x)|\right]^{2+\delta}\left\{f(x)\int_R\left|K^{(L)}(t)\right|^{2+\delta}dt+O(h)\right\}}{\left\{f(x)\alpha_0\left[g_L(x)K^{(L)}\right]^2+O(h)\right\}^{1+\delta/2}}=\frac{c(\varepsilon,\delta,x)}{(nh)^{\delta/2}}\to 0,$$

as $nh\to\infty$.

By the Lindeberg-Feller theorem the assumption $nh\to\infty$ implies $S_n\xrightarrow{d}N(0,1)$.

Now we can prove the convergence stated in (7). By (12) the limit in distribution of $\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x)-E\hat{F}_h(x)\right]$ is the same as that of

$$\left\{f(x)\alpha_0\left[g_L(x)K^{(L)}\right]^2+O(h)\right\}^{1/2}\frac{\hat{F}_h(x)-E\hat{F}_h(x)}{\left[\operatorname{var}(\hat{F}_h(x))\right]^{1/2}}=$$

$$=\left\{f(x)\alpha_0\left[g_L(x)K^{(L)}\right]^2+O(h)\right\}^{1/2}S_n\xrightarrow{d}\Psi.$$

In the equation

$$\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x)-F(x)\right]=\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x)-E\hat{F}_h(x)\right]+\left(nh^{2L+1}\right)^{1/2}\left[E\hat{F}_h(x)-F(x)\right]$$

the first term on the right converges in distribution, as we have just proved, and the second is of order $O\left(\left(nh^{2L+1}\right)^{1/2}h\right)=O\left(\left(nh^{2L+3}\right)^{1/2}\right)=o(1)$, according to (5) and the assumption $nh^{2L+3}\to 0$.

(d) Under the null one has the identity $\left(nh^{2L+3}\right)^{1/2}\hat{G}_h(x)=\left(nh^{2L+1}\right)^{1/2}\left[\hat{F}_h(x)-F(x)\right]$, the conclusion follows from part (c).

REFERENCES

[1] Razali N.M., Wah Y.B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. Journal of Statistical Modeling and Analytics. 1983. Vol. 2, N 1. P. 21-33.
[2] Thadewald T., Buning H. Jarque-Bera test and its competitors for testing normality – a power comparison. Journal of Applied Statistics. 2007. Vol. 34. P. 87-105.
[3] Surucu B. A power comparison and simulation study of goodness-of-fit tests. Computers and Mathematics with Applications. 2008. Vol. 56. P. 1617-1625.
[4] Farrell P.J., Rogers-Stewart K. Comprehensive study of tests for normality and symmetry: extending the Spiegel-haltertest. Journal of Statistical Computation and Simulation. 2008. Vol. 76. P. 803-816.
[5] Szekely G.J., Rizzo M.L. A new test for multivariate normality. Journal of Multivariate Analysis. 2005. Vol. 93. P. 87-105.
[6] Saculinggan M., Balase E.A. Empirical power comparison of goodness of fit tests for normality in the presence of outliers. Journal of Physics. 2013. Vol. 435. P. 1-11.
[7] Henze N., Zirkler B. A class of invariant and consistent tests for multivariate normality. Communications in Statistics – Theory and Methods. 1990. Vol. 19. P. 3595–3617.
[8] Davidson J. Stochastic limit theory. New York Oxford University Press, 1994.

ДИФФЕРЕНЦИАЛДЫҚ ШАРТТАРДЫ ТЕСТІЛЕУ ЖӘНЕ ТЫҒЫЗДЫҚТЫҢ ЖЕРГІЛІКТІ НОРМАЛІ

Қ. Т. Мынбаев[1], А. С. Аипенова[2]

[1]Қазақстан-Британ техникалық университеті, Алматы, Қазақстан
РМК «Математика және математикалық модельдеу институты», Алматы, Қазақстан,
[2]әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан,
РМК «Математика және математикалық модельдеу институты», Алматы, Қазақстан

Тірек сөздер: тестілеу, жергілікті нормальды тест, баламалы болжау, нөлдік болжау, асимптотикалық нормальдық.

Аннотация. Жұмыста, тығыздық функциясы дифференциалдық теңдеуді қанағаттандыратын тестілеу әдісі ұсынылады. Бұл нәтиже тығыздық функциясы нақты үлестірімдер жиынына жататынын тексеру үшін қолданылуы мүмкін. Мысалы, стандартты нормальді тығыздық $f(t)=(2\pi)^{-1/2}\exp(-t^2/2)$ жиыны келесі дифференциалдық теңдеуді $f'(t)+tf(t)=0$ қанағаттандырады. Егер тығыздық осы теңдеуді t нүктесінде қанағаттандырса, онда ол осы нүктеде жергілікті стандартты нормальді деп аталады. Сондықтан, тексерудің тәжірибелік қажеттілігі туындайды, яғни тығыздық функциясы қандайда бір дифференциалдық теңдеуді қанағаттандыратынын тексеру керек. Мынадай өрнекті $F(x)=\sum_{l=0}^{L}g_l(x)f^{(l)}(x)$ қарастырамыз. Біз H_0 нөлдік болжауды тексере аламыз, егер f функциясы $F(x)=0$ теңдікті қанағаттандырса баламалы болжауға қарсы H_a: $F(x)\neq 0$.

Тестілеу әдісі асимптотикалық нормальдық ұйғарыммен анықталады.

ТЕСТИРОВАНИЕ ДИФФЕРЕНЦИАЛЬНЫХ УСЛОВИЙ И ЛОКАЛЬНАЯ НОРМАЛЬНОСТЬ ПЛОТНОСТИ

К. Т. Мынбаев[1], А. С. Аипенова[2]

[1]Казахстанско-Британский технический университет, Алматы, Казахстан,
РГП «Институт математики и математического моделирования», Алматы, Казахстан;
[2]Казахский национальный университет им. аль-Фараби, Алматы, Казахстан,
РГП «Институт математики и математического моделирования», Алматы, Казахстан

Ключевые слова: тестирование, тестирование локальной нормальности, альтернативная гипотеза, нулевая гипотеза, асимптотическая нормальность.

Анотация. В этой статье предлагается процедура тестирования ли плотность удовлетворяет ли плотность дифференциальному уравнению. Этот результат может быть применен для проверки, что плотность относится к конкретному семейству распределений. Например, стандартная нормальная плотность $f(t)=(2\pi)^{-1/2}\exp(-t^2/2)$ удовлетворяет дифференциальному уравнению $f'(t)+tf(t)=0$. Если плотность удовлетворяет этому уравнению в этой точке *t*, то она называется локально стандартной нормальной в этой точке. Таким образом, существует практическая необходимость в проверке, удовлетворяет ли плотность некоторому дифференциальному уравнению. Рассмотрим выражение $F(x)=\sum_{l=0}^{L} g_l(x) f^{(l)}(x)$. Мы можем проверить нулевую гипотезу H_0: f удовлетворяет уравнению $F(x) = 0$ против альтернативной гипотезы H_a: $F(x) \neq 0$.

Метод тестирования сопровождается утверждением об асимптотической нормальности.

Поступила 01.10.2014 г.