

W. KLONECKI (Wrocław)

A NOTE ON OPTIMAL $C(\alpha)$ -TESTS FOR HOMOGENEITY OF THE POISSON DISTRIBUTION

1. Introduction. Neyman and Scott have shown [2] that the test for homogeneity of the Poisson distribution, based on the classical dispersion coefficient (for short, test Z), is an optimal $C(\alpha)$ -test for a vast family of sequences of alternatives converging to a Poisson distribution, each sequence specifying a different character of possible inhomogeneity.

In this paper*, necessary and sufficient conditions for an optimal $C(\alpha)$ -test for homogeneity of the Poisson distribution to be the same as the test Z are given. These conditions show that the test Z is an optimal $C(\alpha)$ -test for a much larger family of sequences of alternatives than it has been pointed out by Neyman and Scott. Also, two examples of optimal $C(\alpha)$ -tests for homogeneity of the Poisson distribution, which are not equivalent to the test Z , are given. The last section contains some remarks concerning the performance of the derived optimal $C(\alpha)$ -tests for testing the homogeneity of the Poisson distribution.

2. General expression for optimal $C(\alpha)$ -tests for homogeneity. Consider N experimental units U_1, \dots, U_N and suppose that, for every unit U_i , $i = 1, \dots, N$, a Poisson random variable X_i is observed. It is assumed that the expected value, say λ , can vary from one experimental unit to the next and is subject to a probability distribution, say $F(\lambda)$. In that case the unconditional probability distribution of X_i is

$$p(x) = \frac{1}{x!} \int_0^{\infty} e^{-\lambda} \lambda^x dF(\lambda).$$

The random variables X_1, \dots, X_N are assumed to be independent. The problem is to test the hypothesis that all the experimental units

* This investigation, conducted at the Statistical Laboratory, University of California, Berkeley, was partially supported by PHS Grant No. GM-10525-09, National Institute of Health, Public Health Service, U.S.A.

are characterized by the same, though unknown, value, say λ_0 , or, equivalently, that $A = \lambda_0$ with probability one.

In order to bring the problem into the general framework in which the theory of $C(\alpha)$ -tests is applicable, one has, as indicated by Neyman and Scott [2], to postulate a family of distributions $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, such that $\{F_\xi(\lambda)\}$ converges as $\xi \rightarrow 0$ to a distribution degenerated at $A = \lambda_0$, say $\varepsilon(\lambda - \lambda_0)$, so that the corresponding sequence of unconditional probability distributions of X_i , say $\{p(x|\xi)\}$, converges to the Poisson distribution with expected value λ_0 . Clearly, there are infinitely many such families of distributions. Let $p(x, 0) = e^{-\lambda_0} \lambda_0^x / x!$.

In this setup the problem of testing the homogeneity of the Poisson distribution reduces to the problem of testing the hypothesis $H_0: \xi = 0$ against the alternative $H_1: \xi > 0$, while λ_0 is a nuisance parameter. Clearly, with appropriate restrictions of regularity of the functions involved, this problem can be tested by the optimal $C(\alpha)$ -test.

In this paper considerations are restricted to families of distributions $\{F_\xi(\lambda)\}$ satisfying the following conditions (conditions A, for short):

- (a) $F_\xi(\lambda) = 0$ for $\lambda < 0$ and $\xi \in (0, a]$;
- (b) $F_\xi(\lambda) \rightarrow \varepsilon(\lambda - \lambda_0)$ as $\xi \rightarrow 0$;
- (c) for all $k \geq 1$, the limits

$$(1) \quad c_k = \lim_{\xi \rightarrow 0} \frac{\mathbb{E}(A_\xi - \lambda_0)^k}{\xi}$$

exist and are finite;

(d) there exists a value ξ_0 , $0 < \xi_0 < a$, such that (5) can be integrated termwise for all ξ , $0 \leq \xi \leq \xi_0$, and that the order of lim and \sum can be subsequently reversed.

The following theorem gives a general expression for test functions of optimal $C(\alpha)$ -tests with respect to families of probability distribution functions which satisfy conditions A:

THEOREM 1. *Let $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, be a family of distributions satisfying conditions A. Then the corresponding optimal $C(\alpha)$ -test criterion for testing $H_0: \xi = 0$ against $H_1: \xi > 0$ has the form*

$$(2) \quad Z_N = \frac{\sum_{i=1}^N g(X_i, \hat{\lambda}_0)}{\sqrt{N\sigma^2(\hat{\lambda}_0)}},$$

where

$$g(x, \lambda) = \frac{1}{e^{-\lambda} \lambda^x} \sum_{k=2}^{\infty} \frac{c_k}{k!} \frac{d^k(e^{-\lambda} \lambda^x)}{d\lambda^k} \quad \text{and} \quad \sigma^2(\lambda) = \text{Var } g(X_i, \lambda),$$

while $\hat{\lambda}_0$ stands for a locally root N consistent estimator for λ_0 .

Proof. As it is well known (cf. [1] and [2]), the test function $g = g(x, \lambda)$ of the optimal $C(\alpha)$ -criterion has the form

$$g = \varphi_\xi - a\varphi_{\lambda_0},$$

where

$$\varphi_\xi = \left. \frac{\partial \log p(x|\xi)}{\partial \xi} \right|_{\xi=0}, \quad \varphi_{\lambda_0} = \left. \frac{\partial \log p(x|0)}{\partial \lambda} \right|_{\lambda=\lambda_0},$$

and

$$a = \frac{\text{cov}(\varphi_\xi, \varphi_{\lambda_0})}{\text{Var} \varphi_{\lambda_0}}.$$

Because

$$(3) \quad p(x|0) = \frac{1}{x!} e^{-\lambda_0} \lambda_0^x,$$

it is easily seen that

$$(4) \quad \varphi_\xi = \frac{1}{x!} \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_0^\infty (e^{-\lambda} \lambda^x - e^{-\lambda_0} \lambda_0^x) dF_\xi(\lambda).$$

Replacing $e^{-\lambda} \lambda^x$ by its expansion around $\lambda = \lambda_0$,

$$e^{-\lambda} \lambda^x = e^{-\lambda_0} \lambda_0^x + \sum_{k=1}^{\infty} \left. \frac{d^k (e^{-\lambda} \lambda^x)}{d\lambda^k} \right|_{\lambda=\lambda_0} \frac{(\lambda - \lambda_0)^k}{k!}$$

reduces (4) to

$$(5) \quad \varphi_\xi = \frac{1}{x!} \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_0^\infty \sum_{k=1}^{\infty} \frac{1}{k!} \left. \frac{d^k (e^{-\lambda} \lambda^x)}{d\lambda^k} \right|_{\lambda=\lambda_0} (\lambda - \lambda_0)^k dF_\xi(\lambda).$$

Finally, in view of assumptions A,

$$(6) \quad \varphi_\xi = \frac{1}{x!} \sum_{k=1}^{\infty} \frac{c_k}{k!} \left. \frac{d^k (e^{-\lambda} \lambda^x)}{d\lambda^k} \right|_{\lambda=\lambda_0}.$$

Using (3), it is easily seen that

$$(7) \quad \varphi_{\lambda_0} = \frac{x}{\lambda_0} - 1.$$

To find $\text{cov}(\varphi_\xi, \varphi_{\lambda_0})$, first note that

$$\mathbb{E} \frac{X_i}{\lambda^{X_i}} \frac{d^k(e^{-\lambda} \lambda^{X_i})}{d\lambda^k} = e^{-\lambda} \frac{d^k \lambda}{d\lambda^k} \quad \text{for } k \geq 1,$$

and, therefore,

$$\mathbb{E} \frac{X_i}{\lambda^{X_i}} \frac{d^k(e^{-\lambda} \lambda^{X_i})}{d\lambda^k} = \begin{cases} e^{-\lambda} & \text{if } k = 1, \\ 0 & \text{if } k > 1. \end{cases}$$

An application of this formula gives

$$\begin{aligned} (8) \quad \text{cov}(\varphi_\xi, \varphi_{\lambda_0}) &= \mathbb{E} \frac{1}{e^{-\lambda_0} \lambda_0^{X_i}} \sum_{k=1}^{\infty} \frac{c_k}{k!} \frac{d^k(e^{-\lambda} \lambda^{X_i})}{d\lambda^k} \Bigg|_{\lambda=\lambda_0} \left(\frac{X_i}{\lambda_0} - 1 \right) \\ &= \frac{c_1}{e^{-\lambda_0} \lambda_0} \mathbb{E} \frac{X_i}{\lambda_0^{X_i}} \frac{d(e^{-\lambda} \lambda^{X_i})}{d\lambda} \Bigg|_{\lambda=\lambda_0} = \frac{c_1}{\lambda_0}. \end{aligned}$$

Using (6), (7) and (8), it is easily seen that

$$g = \frac{1}{e^{-\lambda_0} \lambda_0^x} \sum_{k=2}^{\infty} \frac{c_k}{k!} \frac{d^k(e^{-\lambda} \lambda^x)}{d\lambda^k} \Bigg|_{\lambda=\lambda_0}.$$

This completes the proof of theorem 1.

COROLLARY. Let $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, be a family of probability distributions satisfying conditions A. Then the optimal $C(a)$ -test criterion of the hypothesis $H_0: \xi = 0$ against the alternative $H_1: \xi > 0$ has the form

$$(9) \quad Z_N = \frac{\sum_{i=1}^N [(X_i - \hat{\lambda}_0)^2 - X_i]}{\sqrt{2N} \hat{\lambda}_0}$$

if and only if $c_2 > 0$, while $c_k = 0$ for $k \geq 3$. Here $\hat{\lambda}_0$ stands for any root N consistent estimator of the nuisance parameter λ_0 .

Proof. The formula for Z_N follows immediately from theorem 1. In fact, simple arithmetic calculations show that

$$\text{Varg}(X_i, \lambda_0) = \frac{c_2^2}{4\lambda_0^4} \mathbb{E}[(X_i - \lambda_0)^2 - X_i]^2 = \frac{c_2^2}{2\lambda_0^2},$$

which combined with (2) leads to (9). On the other hand, if Z_N is of the form as given by (9), it can be easily shown that $c_k = 0$ for $k \geq 3$.

3. A theorem on the class of sequences $\{c_k\}$. In view of theorem 1, the optimal $C(a)$ -test with respect to the family $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, is uniquely determined by the sequence $\{c_k\}$ of limits defined by (1). This

being the case, the question arises how to characterize the class of those sequences of finite numbers $\{c_k\}$ for which there exist families of distributions $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, satisfying conditions A. A partial answer gives the following theorem:

THEOREM 2. *Let $\{F_\xi(\lambda)\}$, $\xi \in (0, a]$, be a family of distributions satisfying conditions A. If $c_2 > 0$ and if there exists an even integer $k_1 > 2$ such that $c_{k_1} = 0$, then $c_k = 0$ for all $k > 2$. If $c_2 = 0$, then $c_k = 0$ for all $k \geq 2$.*

The proof of theorem 2 will be divided in lemmas 1 and 2. It is assumed that assumptions A hold.

LEMMA 1. *If k_1 and k_2 are even, $k_1 < k_2$ and $c_{k_2} = 0$, while $c_{k_1} > 0$, then $c_k = 0$ for $k_1 < k \leq k_2$.*

Proof. Without loss of generality, it can be assumed that $EA_\xi = 0$. Let $\varepsilon > 0$ and let $k_1 < k \leq k_2$. Now

$$EA_\xi^k = \int_{|\lambda| \leq \varepsilon} \lambda^k dF_\xi(\lambda) + \int_{|\lambda| > \varepsilon} \lambda^k dF_\xi(\lambda).$$

Because $\lambda^k < \varepsilon^{k-k_2} \lambda^{k_2}$ for $|\lambda| > \varepsilon$, it is easily seen that

$$\int_{|\lambda| > \varepsilon} \lambda^k dF_\xi(\lambda) \leq \varepsilon^{k-k_2} \int_{|\lambda| > \varepsilon} \lambda^{k_2} dF_\xi(\lambda) \leq \varepsilon^{k-k_2} \int \lambda^{k_2} dF_\xi(\lambda).$$

Thus, for any $0 < \xi \leq a$,

$$E \frac{A_\xi^k}{\xi} \leq \int_{|\lambda| \leq \varepsilon} \frac{\lambda^k}{\xi} dF_\xi(\lambda) + \varepsilon^{k-k_2} \int \frac{\lambda^{k_2}}{\xi} dF_\xi(\lambda)$$

and, therefore,

$$E \left| \frac{A_\xi^k}{\xi} \right| \leq \varepsilon^{k-k_1} \int_{|\lambda| \leq \varepsilon} \frac{\lambda^{k_1}}{\xi} dF_\xi(\lambda) + \varepsilon^{k-k_2} \int \frac{\lambda^{k_2}}{\xi} dF_\xi(\lambda)$$

since $|\lambda|^k \leq \varepsilon^{k-k_1} \lambda^{k_1}$ for $|\lambda| \leq \varepsilon$. Hence, in view of the assumptions,

$$|c_k| \leq \varepsilon^{k-k_1} \lim_{\xi \rightarrow 0} \int_{|\lambda| \leq \varepsilon} \frac{\lambda^{k_1}}{\xi} dF_\xi(\lambda) \leq \varepsilon^{k-k_1} \lim_{\xi \rightarrow 0} \int \frac{\lambda^{k_1}}{\xi} dF_\xi(\lambda)$$

or

$$|c_k| \leq \varepsilon^{k-k_1} c_{k_1}.$$

This implies that $c_k = 0$, since $c_{k_1} \geq 0$ and ε can be chosen arbitrarily small.

LEMMA 2. *If k_1 and k_2 are even, $k_2 > k_1$, and $c_{k_1} = 0$, then $c_{k_2} = 0$.*

Proof. Similarly as in the proof of lemma 1, it is assumed that

$E A_\xi = 0$. Suppose that $c_{k_2} > 0$. For any $\varepsilon > 0$

$$E A_\xi^{k_1} \geq \int_{|\lambda| \leq \varepsilon} \lambda^{k_1} dF_\xi(\lambda).$$

Since $\lambda^{k_1} > \varepsilon^{k_1-k_2} \lambda^{k_2}$ for $|\lambda| \leq \varepsilon$, it follows that

$$E A_\xi^{k_1} \geq \varepsilon^{k_1-k_2} \int_{|\lambda| \leq \varepsilon} \lambda^{k_2} dF_\xi(\lambda).$$

Hence

$$c_{k_1} \geq \varepsilon^{k_1-k_2} \overline{\lim}_{\xi \rightarrow 0} \int_{|\lambda| \leq \varepsilon} \frac{\lambda^{k_2}}{\xi} dF_\xi(\lambda).$$

To complete the proof of lemma 2, it is sufficient to note that by assumptions A there exists $\varepsilon_0 > 0$ such that

$$c_{k_2}(\varepsilon_0) = \overline{\lim}_{\xi \rightarrow 0} \int_{|\lambda| \leq \varepsilon} \frac{\lambda^{k_2}}{\xi} dF_\xi(\lambda) > 0,$$

since $c_{k_2}(\varepsilon_0) > 0$ implies that $c_{k_1} > 0$.

The proof of theorem 2 can be now completed as follows:

Proof of theorem 2. In view of lemma 2, the relation $c_{k_1} = 0$ implies that $c_{k_2} = 0$ for all even $k_2 > k_1$. From lemma 1 it then follows that if $k_2 > 2$ is even and if $c_{k_2} = 0$, while $c_2 > 0$, then $c_k = 0$ for $2 < k \leq k_2$. Thus the first assertion of theorem 2 is proved. To prove the second assertion, note that, in view of lemma 2, the assumption $c_2 = 0$ implies that $c_k = 0$ for all even k . However, this implies that $c_k = 0$ for all $k \geq 2$.

4. Examples. We shall present three different test criteria for homogeneity, each of them being an optimal $C(\alpha)$ -test criterion with respect to a particular family of probability distributions characterizing the inhomogeneity of the experimental units.

a. First, consider a family of random variables A_ξ , $\xi \in (0, 1]$, distributed as follows:

For every $\xi \in (0, 1]$, let

$$(10) \quad P(A_\xi = \xi r) = e^{-\lambda_0/\xi} \frac{(\lambda_0/\xi)^r}{r!},$$

where $r = 0, 1, \dots$. Then, whatever $\xi > 0$ is, the random variable X_i has Neyman's A-distribution. It is easily seen that the corresponding family of distributions $\{F_\xi(\lambda)\}$ satisfies conditions A. Moreover, since $E[A_\xi - \lambda_0] = 0$, $E[A_\xi - \lambda_0]^2 = \lambda_0 \xi$ and $E[A_\xi - \lambda_0]^k = o(\xi)$ for $k \geq 3$, it follows that $c_1 = 0$, $c_2 = \lambda_0$ and $c_k = 0$ for $k \geq 3$. Consequently, in view of the corollary, the optimal $C(\alpha)$ -test statistic is given by (9) and represents an optimal $C(\alpha)$ -test

criterion with respect to the family given by (10). Using the maximum likelihood estimate $\hat{\lambda}_0 = \bar{X}$, it reduces to

$$Z_{1N} = \sqrt{\frac{\bar{N}}{2} \left(\frac{S^2}{\bar{X}} - 1 \right)}, \quad \text{where } S^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2.$$

This is the classical dispersion coefficient.

b. Next, consider a family of random variables $A_\xi = \lambda_0 + R_\xi$, $\xi \in (0, 1]$, where $\lambda_0 > 0$ is an unknown constant and R_ξ stands for the Poisson random variable with expectation equal to ξ . This implies that

$$P(A_\xi = \lambda_0 + r) = e^{-\xi} \frac{\xi^r}{r!},$$

where $r = 0, 1, \dots$. The probability generating function corresponding to the distribution of X_i has the following form:

$$G_{X_i}(u|\xi) = \exp\{\xi(e^{1-u} - 1) + \lambda_0(1-u)(1-\xi)\}.$$

Clearly, the family of distributions $\{F_\xi(\lambda)\}$ of the random variables A_ξ , $\xi \in (0, 1]$, satisfies conditions A. However, $c_k = 1$ for $k = 1, 2, \dots$, and the optimal $C(\alpha)$ -test function with respect to the considered family $\{F_\xi(\lambda)\}$ has the following form:

$$Z_{2N} = \frac{\sum_{i=1}^N [e^{-1}((1 + \hat{\lambda}_0)/\hat{\lambda}_0)^{X_i} - X_i/\hat{\lambda}_0]}{\sqrt{N(e^{1/\hat{\lambda}_0} - 1 - 1/\hat{\lambda}_0)}}.$$

With the use of the estimate $\hat{\lambda}_0 = \bar{X}$, this reduces to

$$Z_{2N} = \frac{\sum_{i=1}^N [e^{-1}((1 + \bar{X})/\bar{X})^{X_i} - 1]}{\sqrt{N(e^{1/\bar{X}} - 1 - 1/\bar{X})}}.$$

This is also an optimal $C(\alpha)$ -test function with respect to the family A_ξ , $\xi \in (0, 1]$, where A_ξ assumes only two values, say λ_0 and $\lambda_0 + 1$, with probabilities $1 - \xi$ and ξ , respectively.

c. Finally, consider the family of random variables $\xi \in (0, 1/2]$, where A_ξ assumes the values $\lambda_0 - 1$, λ_0 and $\lambda_0 + 1$ with probabilities $\xi/2$, $1 - \xi$ and $\xi/2$, respectively. Also in this case assumptions A are satisfied, and easy calculations show that $c_{2k} = 1$ and $c_{2k+1} = 0$ for $k = 1, 2, \dots$. An application of (2) and the use of $\hat{\lambda}_0 = \bar{X}$ lead to the following optimal

$C(\alpha)$ -criterion:

$$Z_{3N} = \frac{\sum_{i=1}^N [e((\bar{X} - 1)/\bar{X})^{X_i} + e^{-1}((\bar{X} + 1)/\bar{X})^{X_i} - 2]}{\sqrt{2N(e^{1/\bar{X}} + e^{-1/\bar{X}} - 2)}}.$$

5. Comments on the performance of optimal $C(\alpha)$ -tests for homogeneity. The weakness of any asymptotic optimal test is connected with the question how large should the number N of observations be that would insure a reasonable approximation to the limiting properties of the tests. Comparing the three optimal $C(\alpha)$ -tests presented in this paper, there is no doubt that the number N of observations insuring a reasonable approximation for Z_{1N} is substantially smaller than the number of observations needed to insure the same approximation for Z_{2N} and Z_{3N} . Moreover, in view of theorem 1 and theorem 2, one is willing to conclude that the test statistic Z_{1N} converges to the limiting normal distribution faster than any other $C(\alpha)$ -optimal test statistic.

Monte Carlo methods seem to support these conjectures with respect to Z_{1N} , Z_{2N} and Z_{3N} . Under the null hypothesis the empirical distribution of Z_{1N} was always closer to the normal distribution than there were the empirical distributions of Z_{2N} and Z_{3N} . Also, under alternatives, such as considered in examples b and c, the empirical powers of Z_{2N} and Z_{3N} never exceeded the power of Z_{1N} . Only in a few cases, where N was as large as 400, the empirical powers of Z_{2N} and Z_{3N} seemed to coincide with those of Z_{1N} for some particular values of λ_0 and ξ .

Neyman and Scott [2] pointed out that Z_{1N} has a remarkable property of optimality, which they call *robustness of optimality*, namely that Z_{1N} is an optimal $C(\alpha)$ -test criterion with respect to the family of random variables

$$A_\xi = \lambda_0 e^{A\sqrt{\xi}}, \quad \xi \in (0, 1],$$

no matter what the distribution of A is, provided that it has a finite support and $EA = EA^3 = 0$. From the results presented in this paper, it is seen that the classical test of homogeneity for the Poisson distribution has — besides the property mentioned by Neyman and Scott — some further remarkable properties of optimality.

References

- [1] J. Neyman, *Optimal asymptotic tests of composite statistical hypotheses*, p. 213-234, in *The Harald Cramér Volume*, Almqvist & Wiksell, Stockholm, and Wiley, New York 1959.

- [2] — and E. Scott, *On the use of $C(\alpha)$ optimal tests of composite hypotheses*, Bull. Inst. Internat. Statist. 41 (1967), p. 477-496.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WROCLAW

Received on 30. 12. 1972

W. KLONECKI (Wrocław)

**O OPTYMALNYCH TESTACH $C(\alpha)$ DLA WERYFIKACJI HIPOTEZY
O JEDNORODNOŚCI ROZKŁADU POISSONA**

STRESZCZENIE

Pokazuje się, że dla weryfikacji hipotezy o jednorodności rozkładu Poissona test Z , oparty na klasycznym współczynniku zmienności, jest optymalnym testem $C(\alpha)$ dla szerszej niż to wskazali Neyman i Scott [2] klasy rodzin alternatyw nieskończenie bliskich hipotezie o jednorodności. Podane są zarazem rodziny alternatyw (nieskończenie bliskich hipotezie o jednorodności), dla których test Z nie jest optymalnym testem $C(\alpha)$. Głównym wynikiem jest twierdzenie podające warunki konieczne i dostateczne na to, aby optymalny względem danej rodziny alternatyw test $C(\alpha)$ był identyczny z testem Z . Omówione są też krótko wyniki obliczeń, wykonanych na maszynie cyfrowej, w celu porównania mocy testu Z z dwoma innymi optymalnymi testami $C(\alpha)$. Sugerują one, że dla weryfikacji hipotezy o jednorodności rozkładu Poissona test Z jest lepszy niż inne optymalne testy $C(\alpha)$.