The Annals of Statistics
1986, Vol. 14, No. 2, 759–765

CHI-SQUARE GOODNESS-OF-FIT TESTS FOR RANDOMLY CENSORED DATA[1]

BY M. G. HABIB[2] AND D. R. THOMAS

Oregon State University

Two Pearson-type goodness-of-fit test statistics for parametric families are considered for randomly right-censored data. Asymptotic distribution theory for the test statistics is based on the result that the product-limit process with MLE for nuisance parameters converges weakly to a Gaussian process. The Chernoff–Lehmann (1954) result extends to a generalized Pearson statistic. A modified Pearson statistic is shown to have a limiting chi-square null distribution.

1. Introduction. In this paper we consider the problem of testing the goodness of fit of a parametric family $\{F(t;\theta);\ \theta \in \Theta\}$ of survival distributions from arbitrary right-censored data. Pearson-type chi-squared statistics which compare the Kaplan–Meier (1958) estimate $\hat{F}_N(t)$ to the parametric MLE $F(t;\hat{\theta}_N)$ are studied. The random functions $N^{1/2}[\hat{F}_N(t) - F(t;\hat{\theta}_N)]$ are shown to have a limiting Gaussian process, which generalizes the result of Breslow and Crowley (1974) for $N^{1/2}[\hat{F}_N(t) - F(t;\theta)]$ where θ is the true value. From this result limiting distributions of the Pearson-type statistics are obtained. The limiting process result may be of more general use than for the Pearson-type statistics considered here.

We use the random censorship model. There are N pairs of independent nonnegative random variables $(X_1, U_1), (X_2, U_2), \ldots, (X_N, U_N)$, where the X's denote failure times and the U's the random censoring times. The observed data consist only of $Y_i = \min(X_i; U_i)$ and the indicator functions $\delta_i = I_{[X_i \le U_i]}$ for $i = 1, \ldots, N$. Let $H(u) = P(U > u)$ denote the unknown absolutely continuous survival function for the censoring variable and assume that the distribution of X belongs to a family of absolutely continuous survival functions $\{F(x;\theta):\ \theta \in \Theta\}$ where Θ is an open set in k-dimensional Euclidean space R^k. We consider MLE $\hat{\theta}_N$ of the parameter θ based on a random sample from the joint distribution of Y and δ with density function

$$g(t, \delta; \theta) = [f(t;\theta)H(t)]^{\delta}[F(t;\theta)h(t)]^{1-\delta} \tag{1}$$

with respect to the product of Lebesgue measure on $(0, \infty)$ and counting measure on $\{0, 1\}$, where $f(t;\theta)$ and $h(t)$ are the density functions corresponding to $F(t;\theta)$ and $H(t)$, respectively.

Received September 1981; revised September 1985.

[1]This research was supported in part by USPHS grant CA27532 from the National Cancer Institute, DHHS.

[2]Currently at Kuwait University.

AMS 1980 *subject classifications.* Primary 62G10; secondary 62E20.

Key words and phrases. Goodness-of-fit, chi-square tests, censored data, product-limit process.

We consider generalizations of Pearson type statistics to randomly censored data. Let $0 = t_0 < t_1 < \cdots < t_r < \infty$ denote boundaries for $r + 1$ cells. The cell boundaries could be random, e.g., for specified survival probabilities P_i we may select $\hat{t}_{Ni}$ satisfying $F(\hat{t}_{Ni}; \hat{\theta}_N) = P_i$ as our boundaries. The test statistics are quadratic forms in the random vector

$$\hat{Z}_N = N^{1/2}(\hat{F}_N - F_{\hat{\theta}}), \tag{2}$$

where $\hat{F}_N = (\hat{F}_N(t_1), \ldots, \hat{F}_N(t_r))'$ and $F_{\hat{\theta}} = (F(t_1; \hat{\theta}_N), \ldots, F(t_r; \hat{\theta}_N))'$ are respectively the product-limit estimator and the MLE for the survival function.

In Section 2 the product-limit process with estimated parameters $\hat{Z}_N(t) = N^{1/2}[\hat{F}_N(t) - F(t; \hat{\theta}_N)]$ is shown to converge weakly to a Gaussian process under the null hypothesis H_0: $F(t) \in \{F(t, \theta);\ \theta \in \Theta\}$. This generalizes the result of Breslow and Crowley (1974) for a completely specified survival function $F(t)$. In Section 3 a modified Pearson statistic $\hat{Q}_N(\hat{\theta}_N)$ and a generalized Pearson statistic $Q_N(\hat{\theta}_N)$ are considered [see (7) and (8)]. The modified Pearson statistic $\hat{Q}_N(\hat{\theta}_N)$ is shown in Theorem 2 to have a limiting χ^2_r distribution. For uncensored data the statistic $\hat{Q}_N(\hat{\theta}_N)$ reduces to that proposed by Rao and Robson (1974) and Nikulin (1973), with further development by Moore and Spruill (1975) and Moore (1977). The limiting distribution of the generalized Pearson statistic $Q_N(\hat{\theta}_N)$ is shown in Theorem 3 to be bounded by χ^2_{r-k} and χ^2_r distributions, which is a generalization of the Chernoff and Lehmann (1954) result. These asymptotic results hold for random cell boundaries as well as for fixed cell boundaries.

Chen (1975) and Turnbull and Weiss (1978) proposed goodness-of-fit tests for composite null hypotheses with randomly censored data. Chen considered a generalized Pearson statistic $Q(\tilde{\theta})$, of the same form as (8), based on a modified minimum χ^3 estimator $\tilde{\theta}$. The statistic $Q(\tilde{\theta})$ was shown to have a limiting χ^2_{r-k} distribution under composite null hypotheses. Turnbull and Weiss considered a likelihood ratio test based on the more restrictive model where both the failure distribution and the censoring distribution are assumed to be discrete with finite support. Several tests have been suggested for the case of a simple null hypothesis with randomly censored data; see Koziol and Green (1976), Hollander and Proschan (1979), Fleming et al. (1980), Fleming and Harrington (1981), and Nair (1981). For the case of Type II censoring (when censoring occurs at specified ordered failures), Mihalko and Moore (1980) used sample percentiles as cell boundaries to obtain Pearson-type tests of fit that have limiting chi-square distributions for composite null hypotheses.

2. Weak convergence of the process $\hat{Z}_N(t)$. Let the random function $Z_N(t) = N^{1/2}[\hat{F}_N(t) - F(t; \theta)]$ be defined on an interval $[0, T]$ where $H(T)F(T; \theta) > 0$. Breslow and Crowley (1974) proved that $Z_N(t)$ converges weakly to a mean 0 Gaussian process $Z(t)$ with

$$\operatorname{Cov}(Z(t), Z(s)) = F(t; \theta)F(s; \theta)\int_0^s \frac{f(z; \theta)}{H(z)F^2(z; \theta)}\, dz \tag{3}$$

for $0 \le s \le t \le T$.

To show weak convergence of $\hat{Z}_N(t)$ we make the following assumptions:

(A.1) $F(t; \theta)$ and $f(t; \theta)$ are twice differentiable in θ with continuous derivatives.

(A.2) The information matrix $J = J[\theta, H]$ satisfies

$$J_{ij} = -\int \frac{\partial^2 \ln f(t, \theta)}{\partial\theta_i\, \partial\theta_j} H(t) f(t; \theta)\, dt - \int \frac{\partial^2 \ln F(t, \theta)}{\partial\theta_i\, \partial\theta_j} F(t; \theta) h(t)\, dt$$

for $i, j = 1, \ldots, k$, is positive definite, and is continuous in θ.

(A.3) The MLE $\hat{\theta}_N$ exists and is efficient with $N^{1/2}(\hat{\theta}_N - \theta) = J^{-1}W_N + o_P(1)$, where W_N is the normalized score vector

$$W_N = N^{-1/2} \sum_{i=1}^{N} \frac{\partial \ln(g(Y_i, \delta_i; \theta))}{\partial\theta}.$$

THEOREM 1. *Let $T < \infty$ satisfy $H(T)F(T; \theta) > 0$ for $\theta \in \Theta$. Then, under the Assumptions* A, *the random function $\hat{Z}_N(t)$, for $0 < t < T$, converges weakly to a mean* 0 *Gaussian process $\hat{Z}(t)$ with*

$$\operatorname{Cov}[\hat{Z}(s), \hat{Z}(t)] = \operatorname{Cov}[Z(s), Z(t)] - \frac{\partial F(s; \theta)'}{\partial\theta} J^{-1} \frac{\partial F(t; \theta)}{\partial\theta} \quad \text{for } 0 < s \le t < T.$$

PROOF. Expand $\hat{Z}_N(t)$ around $\hat{\theta}_N = \theta$ to give

$$\hat{Z}_N(t) = Z_N(t) + Z_N^*(t) + R_N(t), \tag{4}$$

where

$$Z_N^*(t) = \frac{\partial F(t; \theta)}{\partial\theta} N^{1/2}(\hat{\theta}_N - \theta)$$

and $R_N(t) \to 0$ in probability uniformly in t.

First we show convergence of finite-dimensional distributions of $Z_N(t) + Z_N^*(t)$. For an arbitrary partition $0 < t_1 < \cdots < t_r < T$, let $Z_N = (Z_N(t_1), \ldots, Z_N(t_r))'$ and $Z_N^* = (Z_N^*(t_1), \ldots, Z_N^*(t_r))'$, so that $Z_n^* = BN^{1/2}(\hat{\theta}_N - \theta)$, where the elements of B are $B_{ij} = \partial F(t_i; \theta)/\partial\theta_j$. The components of Z_N and of $N^{1/2}(\hat{\theta}_N - \theta)$ can each be written, to order $o_P(1)$, as a normalized sum of continuous functions of $(Y_1, \delta_1), \ldots, (Y_N, \delta_N)$. This follows from Breslow and Crowley's (1974) results (7.9), (7.12), and Assumption A.3, respectively. Hence, from the Central Limit Theorem, we have

$$\begin{bmatrix} Z_N \\ N^{1/2}(\hat{\theta}_N - \theta) \end{bmatrix} \to_L \begin{bmatrix} Z \\ \eta \end{bmatrix} \sim N\left(0, \begin{bmatrix} V & C \\ C' & J^{-1} \end{bmatrix}\right), \tag{5}$$

where the elements of V are given by (3) with $V_{ij} = \operatorname{Cov}[Z(t_i), Z(t_j)]$ and J is the information matrix given in A.2. Hence $Z_N = Z_N^* \to \hat{Z} = Z + B\eta \sim N(0, \Sigma)$.

Further, Σ^* can be evaluated without direct computation of C in (5) as follows. Under A.1–A.3 it follows from the result of Pierce (1982) that $\hat{Z}$ and η are independent, and thus that $V = \Sigma^* + BJ^{-1}B'$. This gives the main result here that

$$\text{Var}(\hat{Z}) = \Sigma^* = V - BJ^{-1}B'. \tag{6}$$

Having shown this, the weak convergence of $Z_N(t) + Z_N^*(t)$ will then follow from marginal weak convergence of $Z_N(t)$ and $Z_N^*(t)$ to continuous limits; see, for example, the argument used by Breslow and Crowley (1974, Theorem 4). The convergence of $Z_N(t)$ is a standard result and that of $Z_N^*(t)$ is clear, since it is a nonrandom vector function of t multiplied by a random vector (free of t) with a limiting distribution. □

3. The test statistics. Let $\hat{V}$ and $\hat{\Sigma}^*$ denote respectively the estimators obtained from the covariance matrices V and Σ^* by replacing θ by the MLE $\hat{\theta}_N$ and the censoring distribution H by the product-limit estimator $\hat{H}_N$.

The modified Pearson statistic is defined as

$$\hat{Q}_N(\hat{\theta}_N) = \hat{Z}_N' \hat{\Sigma}^{*-1} \hat{Z}_N \tag{7}$$

and the generalized Pearson statistic as

$$Q_N(\hat{\theta}_N) = \hat{Z}_N' \hat{V}^{-1} \hat{Z}_N. \tag{8}$$

The limiting distributions for these test statistics are developed in the following two theorems. The arguments are given for fixed-cell boundaries first, with subsequent extension to random-cell boundaries.

THEOREM 2. *Under composite null hypotheses, Assumptions* A.1–A.3 *and that* Σ^* *is of full rank* r, *the statistic* $\hat{Q}_N(\hat{\theta}_N)$ *has a limiting* χ_r^2 *distribution.*

PROOF. The components V, B, and J are each continuous in θ. It can be shown that V is continuous in H with respect to the supremum metric over that interval $[0, T]$ and J is continuous in H with respect to the supremum metric over the interval $[0, \infty)$. Since $\hat{\theta}_N$ and $\hat{H}_N$ are consistent estimators it then follows that $\hat{V}$ and $\hat{\Sigma}^*$ converges in probability to V and Σ^*, respectively. Theorem 1 can then be used to complete the proof. □

THEOREM 3. *Under composite null hypotheses, Assumptions* A.1–A.3 *and the assumption that the gradient matrix* B *is of full rank* (k) *the statistic* $Q_N(\hat{\theta}_N)$ *has a limiting distributions which is bounded by* χ_{r-k}^2 *and* χ_r^2 *distributions.*

PROOF. From Theorem 1 and the convergence of $\hat{V}$ to V in probability it follows that $Q_N(\hat{\theta}_N) \to_L Z'V^{-1}Z$, where $Z \sim N(0, \Sigma^*)$. Let Λ be a diagonal matrix of eigenvalues of V and P the corresponding orthogonal matrix of eigenvectors. Then let Λ^* be a diagonal matrix of eigenvalues of $\Lambda^{-1/2}P'\Sigma^*P\Lambda^{-1/2}$ and P^* the

corresponding orthogonal matrix of eigenvectors. We can then write

$$Z'V^{-1}Z = \sum_{i=1}^{r} \lambda_i \xi_i^2, \tag{9}$$

where the ξ_i's are independent $N(0,1)$. The eigenvalues λ_i^* satisfy the equation

$$\begin{aligned} 0 &= \left| P\Lambda^{1/2}\left(\Lambda^{-1/2}P'\Sigma P\Lambda^{-1/2} - \lambda_i^* I \right)\Lambda^{1/2}P' \right| \\ &= |\Sigma - \lambda_i V| \\ &= (-1)^k \left| BJ^{-1}B' - (1 - \lambda_i^*)V \right|. \end{aligned}$$

To conclude the proof, note that the nonzero roots of $|BJ^{-1}B' - \mu V| = 0$ and those of $|B'V^{-1}B - \mu J| = 0$ are identical [see Rao (1973, page 68)] and $BJ^{-1}B'$ and $B'V^{-1}B$ are nonnegative definite implies that (9) reduces to

$$Z'V^{-1}Z = \sum_{i=1}^{k} \xi_i^2 \lambda_i + \sum_{i=k+1}^{r} \xi_i^2,$$

where $\lambda_i \in (0,1)$ for $i = 1, \ldots, k$. □

The treatment of random change of time on pages 144–145 of Billingsley (1968) can be applied here. He shows that if Φ_N is a random monotone function which converges in probability to a function Φ with Φ_N and Φ having the same finite domain then the random composite function $\hat{Z}_N \circ \Phi_N$ converges weakly to the Gaussian process $\hat{Z} \circ \Phi$. The asymptotic distributions of the test statistics $\hat{Q}_N$ and Q_N given in Theorems 2 and 3 then hold for random partition points which depends on Φ_N. For example, in our chi-square goodness-of-fit application, one can truncate the fitted survival function at $t = T$ and define $F^*(t; \hat{\theta}_N) = F(t; \hat{\theta}_N)$ for $-\infty < t < T$ and 0 otherwise. Then use $\Phi_N(P) = F^{*-1}(P; \hat{\theta}_N)$ for $0 < P < 1$ to produce random cell boundaries based on specified values on the survival scale $1 > P_1 > \cdots > P_r > 0$. For a sample of size N one might need to reduce the number of cells from r^* to $\hat{r}_N$ where $\hat{r}_N = \max\{i\colon F^{*-1}(P_i; \hat{\theta}_N) < T\}$. Then $\hat{r}_N$ converges in probability to r where $r = \max\{i\colon P_i > F(T; \theta)\}$. The asymptotic distribution of the test statistics $\hat{Q}_N$ and Q_N then holds for the random partition points $\hat{t}_{Ni} = F^{*-1}(P_i, \hat{\theta}_N)$.

4. Computation. Let there be c censored observations with censoring times $t_1^* < t_2^* < \cdots < t_c^*$. Define $t_0^* = 0$ and $t_{c+1}^* = \infty$. Let $m(s) = \max\{j;\ t_j^* < s\}$. Then, the uncorrected covariance matrix V and the generalized Pearson statistic $Q_N(\hat{\theta}_N)$ reduce to the following simple forms:

$$\hat{V}_{ij} = F(t_i; \hat{\theta}_N)F(t_j; \hat{\theta}_N)\,\hat{d}_i$$

and

$$Q_N(\hat{\theta}_N) = N \sum_{i=1}^{r} \{\hat{F}_N(t_{i-1})/F(t_{i-1}; \hat{\theta}_N) - \hat{F}_N(t_i)/F(t_i; \hat{\theta}_N)\}/(\hat{d}_i - \hat{d}_{i-1}),$$

where

$$d_i = \int_0^{t_i} \frac{f(y;\hat{\theta}_N)}{\hat{H}_N(y)F^2(y;\hat{\theta}_N)}\,dy$$
$$= \sum_{j=1}^{m(t_i)} \left\{1/F(t_j^*;\hat{\theta}_N) - 1/F(t_{j-1}^*;\hat{\theta}_N)\right\}/\hat{H}_N(t_j^*)$$
$$+\left\{1/F(t_{m(t_i)}^*;\hat{\theta}_N) - 1/F(t_i;\hat{\theta}_N)\right\}/\hat{H}_N(t_{m(t_i)}^*).$$

Unfortunately, the computation of the modified Pearson statistic $\hat{Q}_N(\hat{\theta}_N)$, which uses the corrected covariance matrix $\hat{\Sigma}$, is not as simple as that of $Q_N(\hat{\theta}_N)$. However, one could take advantage of the identity $\hat{\Sigma}^{-1} = \hat{V}^{-1} + \hat{C}$, where $\hat{C} = \hat{V}^{-1}\hat{B}(\hat{B}'\hat{V}^{-1}\hat{B} - \hat{J})^{-1}\hat{B}'\hat{V}^{-1}$ to get $\hat{Q}_N(\hat{\theta}_N) = Q_N(\hat{\theta}_N) + \hat{Z}_N'\hat{C}\hat{Z}_N$. Thus, if $Q_N(\hat{\theta}_N)$ is greater than $\chi^2_{(r,\alpha)}$ then $\hat{Q}_N(\hat{\theta}_N)$ would be also. Further, because $\hat{Z}_N'\hat{C}\hat{Z}_N$ is bounded by χ^2_k, it following that, if $Q_N(\hat{\theta}_N)$ is less than $\chi^2_{r-k}(\alpha)$ then $\hat{Q}_N(\hat{\theta}_N)$ would be less than $\chi^2_r(\alpha)$. Hence, $\hat{Q}_N(\hat{\theta}_N)$ need only be computed when $\chi^2_{r-k}(\alpha) < Q_N(\hat{\theta}_N) < \chi^2_r(\alpha)$. In such a case, the components of the information matrix $\hat{J}$ are required. They are given by

$$\hat{J}_{ij} = \sum_{l=1}^{c+1} H_N(t_{l-1}^*)\{\hat{K}_{ij}(t_{l-1}^*) - \hat{K}_{ij}(t_l^*)\}$$
$$- \sum_{l=1}^{c+1} \frac{\partial^2 \ln F(t_1^*;\hat{\theta}_N)}{\partial\theta_i\,\partial\theta_j} f(t^*;\hat{\theta}_N)\{\hat{H}_N(t_{l-1}^*) - \hat{H}_N(t_l^*)\},$$

where

$$\hat{K}_{ij}(s) = \int_0^s \frac{\partial^2 \ln f(t;\hat{\theta}_N)}{\partial\theta_i\,\partial\theta_j} f(t;\hat{\theta}_N)\,dt.$$

These components, $\hat{K}_{ij}(s)$, require numerical integration in some applications, for example the two-parameter Weibull and gamma distributions.

Acknowledgment. We are grateful to a referee for suggesting the current version of Theorem 1 which permitted inclusion of random cell boundaries.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

CHEN, J. (1975). Goodness of fit tests under random censorship. Ph.D. thesis, Dept. Statistics, Oregon State Univ.

CHERNOFF, H. and LEHMANN, E. L. (1954). The use of maximum likelihood estimates in χ^2 tests for goodness of fit. *Ann. Math. Statist.* **25** 579–586.

FLEMING, T. and HARRINGTON, D. (1981). A class of hypothesis tests for one and two sample censored survival data. *Comm. Statist. A—Theory Methods* **10** 763–794.

FLEMING, T., O'FALLON, J. and O'BRIEN, P. (1980). Modified Kolmogorov–Smirnov test procedures with application to arbitrarily right-censored data. *Biometrics* **36** 607–625.

HABIB, M. G. (1981). A chi-square goodness-of-fit test for censored data. Ph.D. thesis, Dept. Statistics, Oregon State Univ.

HOLLANDER, M. and PROSCHAN, F. (1979). Testing to determine the underlying distribution using randomly censored data. *Biometrics* **35** 393–401.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–480.

KOZIOL, J. A. and GREEN, S. B. (1976). A Cramér–von Mises statistic for randomly censored data. *Biometrika* **63** 465–464.

MIHALKO, D. P. and MOORE, D. S. (1980). Chi-square tests of fit for Type II censored data. *Ann. Statist.* **8** 625–644.

MOORE, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-square tests of fit. *J. Amer. Statist. Assoc.* **72** 131–137.

MOORE, D. S. and SPRUILL, M. C. (1975). Unified large sample theory of general chi-square statistics for tests of fit. *Ann. Statist.* **3** 599–616.

NAIR, V. N. (1981). Plots and tests for goodness of fit with randomly censored data. *Biometrika* **68** 99–103.

NIKULIN, M. (1973). Chi-square tests for continuous distributions with shift and scale parameters. *Theory Probab. Appl.* **18** 559–568.

PIERCE, D. A. (1982). The asymptotic effect of substituting estimators for parameters in certain types of statistics. *Ann. Statist.* **10** 475–478.

PYKE, R. (1969). Applications of almost surely convergent constructions of weakly convergent processes. *Lecture Notes in Math.* **89** 187–200. Springer, New York.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

RAO, K. C. and ROBSON, D. S. (1974). A chi-square statistic for goodness-of-fit tests within the exponential family. *Comm. Statist.* **3** 1139–1153.

TURNBULL, B. W. and WEISS, L. (1978). A likelihood ratio statistic for testing goodness-of-fit with randomly censored data. *Biometrics* **34** 367–375.

DEPARTMENT OF STATISTICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331