ISSN: 0128-7680
Pertanika J. Sci. & Technol. **5(1)**: 77-84 (1997)

Confidence Intervals for Parallel Systems with Covariates

Ayman Baklizi, Isa Daud and Noor Akma Ibrahim
Department of Mathematics
Faculty of Science and Environmental Studies
Universiti Pertanian Malaysia
43400 UPM, Serdang, Selangor, Malaysia

Received: 10 April 1996

ABSTRAK

Lazimnya bagi model regresi dengan data tertapis selang keyakinannya yang tepat tidak mudah disingkap kembali dan dengan demikian hampiran kepada selang ini diperlukan. Hampiran selang keyakinan ini lazimnya dapat dibina dengan menggunakan kaedah yang berdasarkan taburan normal asimptot dari penganggar kebolehjadian maksimum. Selang-selang seperti ini mudah dikira dan sering dimanfaatkan oleh kebanyakan pakej statistik berkomputer. Walau bagaimanapun selang-selang ini mempunyai beberapa kecacatan. Bagi saiz sampel kecil atau sederhana kecil, selang-selang ini bersifat antikonservatif serta mempunyai kebarangkalian hujung atas dan hujung bawah yang tidak simetri. Oleh yang demikian kaedah alternatif disarankan. Berdasarkan asimptot penganggar kebolehjadian maksimum, kaedah alternatif ini menyarankan selang anggaran dibina dari ujian nisbah kebolehjadian songsang. Kemampuan selang ini diselidik bagi model regresi yang merangkumi data rawak tertapis di sebelah kanan dengan kovariat. Pendekatan model adalah berlandaskan sistem selari. Daripada keputusan, selang-selang yang dibina berdasarkan ujian nisbah kebolehjadian songsang adalah lebih mantap. Selang-selang ini mempunyai kebarangkalian liputan yang hampir kepada nominalnya dengan kebarangkalian di hujung atas dan di hujung bawah hampir simetri.

ABSTRACT

Exact confidence intervals for regression models with censored data are often not tractable, and hence approximate intervals are derived. The most common method of obtaining these approximate intervals is based on the asymptotic normal distribution of the maximum likelihood estimator. These intervals are easy to compute and they are used in most computer statistical packages. However, these intervals have some limitations. When the sample size is small or even moderate they tend to be anticonservative and have asymmetric upper and lower tail probabilities. An alternative method based on the asymptotics of the maximum likelihood estimator is to construct intervals from the inverted likelihood ratio tests. The performance of these intervals is investigated for the regression models based on parallel systems with covariates, and with randomly right censored data for finite samples. The simulation results show that the intervals based on the inverted likelihood ratio test have better performance. They have coverage probability that is close to the nominal one, and have nearly symmetric upper and lower tail probabilities.

Keywords: inverted likelihood tests, random censorship, simulation, asymptotic normality

INTRODUCTION

Recent research on confidence interval estimation based on the asymptotics of the maximum likelihood estimation shows that intervals derived from inverted likelihood tests have desirable properties, and generally perform better than those based on the asymptotic normality of the maximum likelihood estimator (Meeker 1987; Doganaksoy and Schmee 1991; Doganaksoy 1991; Vander Wiel and Meeker 1990).

We shall consider these intervals for the parameters of the regression model based on parallel systems with covariates. A parallel system is a multi-component system that fails only when all of its components fail. These systems are used in industry to increase the reliability of certain products (Bain 1978). In a medical setting, lungs and kidneys are examples of such a parallel system (Elandt-Johnson and Johnson 1979). Parallel systems can be considered as a special kind of nomination sample which consists of independently distributed maxima from subsamples with the same underlying distribution (Kvam and Samaniego 1993).

We shall assume that the life time of each component in the system has an exponential distribution, and the systems have equal numbers of components. The effect of the covariates will be incorporated in the model by expressing the parameters of the distribution as a function of these covariates (Elandt-Johnson and Johnson 1979). In this model, exact confidence intervals based on the maximum likelihood estimator are not tractable, and therefore, large sample approximations are needed.

We shall consider two types of these approximations:

1. Confidence intervals based on the asymptotic normality of the maximum likelihood estimator.
2. Confidence intervals based on inverted likelihood ratio tests.

Under the random censorship mechanism, the behaviour of these intervals will be investigated in terms of their attainment of the nominal level, conservativeness, and the symmetry of their tail probabilities.

THE MODEL

Suppose that the life times of the components in the system, denoted by $X_1, X_2, \ldots, X_m$ are independently and identically distributed as exponential with parameter λ. The life time of the system is equivalent to the life time of the longest-lasting component, and is therefore given by $T = \max(X_1, \ldots, X_m)$. The density function of T is given by

$$f(t,\lambda) = \begin{cases} \frac{m}{\lambda}\, exp\left(-\frac{t}{\lambda}\right)\left(1 - exp\left(-\frac{t}{\lambda}\right)\right)^{(m-1)} & t, \lambda > 0. \\ 0 & \text{otherwise} \end{cases} \qquad [1]$$

Putting $\lambda = \exp(\beta'\mathbf{Z})$, where $\mathbf{Z} = (1, z_1, \ldots, z_p)'$ is a vector of covariates and $\beta = (\beta_0, \beta_1 \ldots, \beta_p)'$ is a vector of regression parameters. Transforming $U = \log(T)$ we get.

$$g\left(u,\mathbf{Z}\right) = m\, exp\left(u - \beta'\mathbf{Z}\right) exp\left(-exp\left(u - \beta'\mathbf{Z}\right)\right)\left(1 - exp\left(-exp\left(u - \beta'\mathbf{Z}\right)\right)\right)^{(m-1)}$$
$$-\infty < u < \infty, \qquad [2]$$

which defines a regression model given by $u = \beta'\mathbf{Z} + \varepsilon$ where ε is an error term with density function given by

$$h(\varepsilon) = m\, exp(\varepsilon)\, exp\left(-exp(\varepsilon)\right)\left(1 - exp\left(-exp(\varepsilon)\right)\right)^{(m-1)} \quad -\infty < \varepsilon < \infty \qquad [3]$$

With m = 1 this model reduces to the exponential distribution with error term distributed as an extreme value random variable (Kalbfleisch and Prentice 1980).

The observed data is of the form $(y_i, \delta_i), i = 1, \ldots, n$, where y_i is equal to (u_i, c_i), $c_1, c_2, \ldots, c_n$ being the censor values, and δ_i is an indicator variable given by

$$\delta_i = \begin{cases} 1 \;;\; u_i \le c_i \\ 0 \;;\; u_i > c_i \end{cases} \qquad [4]$$

CONFIDENCE INTERVALS

This section describes two methods of obtaining approximate intervals for the regression parameters. These approximations are based on the asymptotics of the maximum likelihood estimator (Kalbfleisch and Prentice 1980; Lawless 1982).

APPROXIMATE INTERVALS BASED ON THE ASYMPTOTIC NORMALITY

Intervals based on the asymptotic normality are widely used. Most of the common statistical packages use this kind of approximation. These intervals are easy to calculate and they are reasonably accurate when the sample is large. A two-sided confidence interval for β_i is given by

$$\left[\hat{\beta}_i - Z_{(1-\alpha/2)}\, s\left(\hat{\beta}_i\right),\ \hat{\beta}_i + Z_{(1-\alpha/2)}\, s\left(\hat{\beta}_i\right)\right] \qquad [5]$$

where i = 0, 1, ..., p respectively, and where s(.) denotes the sample standard deviation for the given estimator, and α denotes the nominal level of the confidence interval.

CONFIDENCE INTERVALS BASED ON INVERTED LIKELIHOOD RATIO TESTS

These confidence intervals are constructed using the fact that, asymptotically, $-2 \log (R(\beta_i))$; $(i = 0, 1, \ldots, p)$, has a chi-squared distribution with one degree of freedom (Vander Wiel and Meeker 1990), where for $p = 1$ with parameters β_0, β_1 we have

$$R(\beta_0) = \max_{\beta_1}\left(L(\beta_0, \beta_1) / L(\hat{\beta}_0, \hat{\beta}_1)\right) \quad [6]$$

$$R(\beta_1) = \max_{\beta_0}\left(L(\beta_0, \beta_1) / L(\hat{\beta}_0, \hat{\beta}_1)\right) \quad [7]$$

where L denotes the likelihood function, and $\hat{\beta}_0, \hat{\beta}_1$ are the maximum likelihood estimators of β_0 and β_1. The bounds of these intervals are given as the solutions of

$$-2 \log \left(R(\beta_i)\right) - X^2\left(_{1-\alpha, 1}\right) = 0. \quad [8]$$

These intervals are somewhat complicated to compute; however, Venzon and Moolgavkar (1988) provide an efficient algorithm to compute the bounds.

THE SIMULATION STRUCTURE

The simulation structure adopted here has $m = 3$. It is restricted to the case of two parameters, β_0 and β_1 equal to 1. The sample size is varied as 20, 30, and 40. The level of significance α is taken to be 0.05. The covariate values are taken on an equally spaced lattice, 0, 0.2, ... , 1.8. Four censoring mechanisms are considered: exponential censoring, uniform censoring, singly type 1 censoring, and type 1 censoring, with three censoring proportions in each case, that is, 0.1, 0.3, and 0.5. There were 5000 replications for each simulation run, as suggested by Piegorsch (1987).

The simulation program was written in FORTRAN with double precision. The maximum likelihood estimator was found using the Newton-Raphson method; necessary and sufficient conditions for the existence of the maximum likelihood estimator are given in Hamada and Tse (1988). The maximum likelihood estimator is unique because the log-likelihood function is concave; this property follows from the concavity of log(g) where g is the density function of u. See equation [2/], Burridge (1981). The method of Venzon and Moolgavkar (1988) was used to obtain the bounds of the intervals based on inverted likelihood ratio tests. Intervals based on the asymptotic normality were obtained using the formula given in equation [5] with estimates of the standard deviations obtained from the inverse of the observed information matrix. The simulation results are given in Tables 1, 2, 3, and 4.

TABLE 1

Observed error rates of confidence intervals for regression parameters based on 5000 samples under singly type 1 censoring

			β_0			β_1		
SS	**CP**	**M**	**L**	**U**	**T**	**L**	**U**	**T**
20	0.1	LR	0.0212	0.0272	0.0484	0.0288	0.0306	0.0594
		AN	0.0154	0.0322	0.0476	0.0288	0.0306	0.0594
	0.3	LR	0.0252	0.0280	0.0532	0.0316	0.0316	0.0632
		AN	0.0164	0.0330	0.0494	0.0298	0.0300	0.0598
	0.5	LR	0.0286	0.0262	0.0548	0.0274	0.0286	0.0560
		AN	0.0112	0.0320	0.0432	0.0234	0.0234	0.0474
30	0.1	LR	0.0264	0.0256	0.0520	0.0288	0.0278	0.0566
		AN	0.0218	0.0308	0.0526	0.0284	0.0276	0.0560
	0.3	LR	0.0272	0.0282	0.0554	0.0248	0.0258	0.0506
		AN	0.0194	0.0328	0.0522	0.0240	0.0256	0.0496
	0.5	LR	0.0318	0.0276	0.0594	0.0268	0.0278	0.0546
		AN	0.0188	0.0352	0.0540	0.0246	0.0262	0.0508
40	0.1	LR	0.0276	0.0256	0.0532	0.0266	0.0284	0.0550
		AN	0.0230	0.0302	0.0532	0.0266	0.0286	0.0552
	0.3	LR	0.0268	0.0256	0.0524	0.0260	0.0292	0.0552
		AN	0.0210	0.0302	0.0512	0.0248	0.0282	0.0530
	0.5	LR	0.0308	0.0278	0.0586	0.0294	0.0322	0.0616
		AN	0.0234	0.0326	0.0560	0.0266	0.0306	0.0572

SS = sample size, CP = censoring proportion, M = method, L = lower tail error probability, U = upper tail error probability, T = total error probability, LR = likelihood ratio, AN = asymptotic normality

TABLE 2

Observed error rates of confidence intervals for regression parameters based on 5000 samples under uniform censoring

			β_0			β_1		
SS	**CP**	**M**	**L**	**U**	**T**	**L**	**U**	**T**
20	0.1	LR	0.0234	0.0264	0.0498	0.0300	0.0268	0.0568
		AN	0.0174	0.0308	0.0482	0.0296	0.0270	0.0566
	0.3	LR	0.0230	0.0260	0.0490	0.0290	0.0276	0.0566
		AN	0.0172	0.0310	0.0482	0.0282	0.0270	0.0552
	0.5	LR	0.0238	0.0254	0.0492	0.0304	0.0298	0.0602
		AN	0.0118	0.0304	0.0422	0.0256	0.0254	0.0510
30	0.1	LR	0.0240	0.0268	0.0508	0.0258	0.0246	0.0504
		AN	0.0210	0.0322	0.0532	0.0254	0.0248	0.0502
	0.3	LR	0.0258	0.0284	0.0542	0.0226	0.0278	0.0504
		AN	0.0188	0.0326	0.0514	0.0220	0.0274	0.0494
	0.5	LR	0.0280	0.0288	0.0568	0.0264	0.0286	0.0550
		AN	0.0166	0.0338	0.0504	0.0240	0.0260	0.0500
40	0.1	LR	0.0270	0.0256	0.0526	0.0260	0.0285	0.0544
		AN	0.0224	0.0296	0.0520	0.0258	0.0280	0.0538
	0.3	LR	0.0288	0.0244	0.0532	0.0268	0.0306	0.0574
		AN	0.0236	0.0288	0.0524	0.0264	0.0300	0.0564
	0.5	LR	0.0298	0.0272	0.0570	0.0278	0.0322	0.0600
		AN	0.0202	0.0314	0.0516	0.0268	0.0300	0.0568

SS = sample size, CP = censoring proportion, M = method, L = lower tail error probability, U = upper tail error probability, T = total error probability, LR = likelihood ratio, AN = asymptotic normality

TABLE 3

Observed error rates of confidence intervals for regression parameters based on 5000 samples under exponential censoring

SS	CP	M	β_0 L	β_0 U	β_0 T	β_1 L	β_1 U	β_1 T
20	0.1	LR	0.0236	0.0272	0.0508	0.0296	0.0266	0.0562
		AN	0.0178	0.0324	0.0502	0.0298	0.0266	0.0564
	0.3	LR	0.0250	0.0248	0.0498	0.0294	0.0282	0.0576
		AN	0.0172	0.0300	0.0472	0.0290	0.0270	0.0560
	0.5	LR	0.0248	0.0234	0.0482	0.0302	0.0278	0.0580
		AN	0.0122	0.0284	0.0406	0.0262	0.0240	0.0502
30	0.1	LR	0.0252	0.0272	0.0524	0.0246	0.0250	0.0496
		AN	0.0208	0.0326	0.0534	0.0246	0.0250	0.0496
	0.3	LR	0.0254	0.0280	0.0534	0.0232	0.0272	0.0504
		AN	0.0202	0.0326	0.0528	0.0230	0.0262	0.0492
	0.5	LR	0.0258	0.0270	0.0528	0.0236	0.0242	0.0478
		AN	0.0178	0.0338	0.0516	0.0222	0.0228	0.0450
40	0.1	LR	0.0278	0.0256	0.0534	0.0268	0.0278	0.0546
		AN	0.0220	0.0294	0.0514	0.0268	0.0280	0.0548
	0.3	LR	0.0300	0.0252	0.0552	0.0260	0.0294	0.0554
		AN	0.0244	0.0284	0.0528	0.0256	0.0290	0.0546
	0.5	LR	0.0270	0.0270	0.0540	0.0276	0.0308	0.0584
		AN	0.0190	0.0324	0.0514	0.0270	0.0290	0.0560

SS = sample size, CP = censoring proportion, M = method, L = lower tail error probability, U = upper tail error probability, T = total error probability, LR = likelihood ratio, AN = asymptotic normality

TABLE 4

Observed error rates of confidence intervals for regression parameters based on 5000 samples under type 1 censoring

SS	CP	M	β_0 L	β_0 U	β_0 T	β_1 L	β_1 U	β_1 T
20	0.1	LR	0.0212	0.0272	0.0483	0.0288	0.0306	0.0594
		AN	0.0154	0.0322	0.0476	0.0288	0.0306	0.0594
	0.3	LR	0.0252	0.0280	0.0532	0.0316	0.0318	0.0634
		AN	0.0164	0.0330	0.0494	0.0298	0.0300	0.0598
	0.5	LR	0.0286	0.0262	0.0548	0.0274	0.0286	0.0560
		AN	0.0112	0.0320	0.0432	0.0234	0.0234	0.0474
30	0.1	LR	0.0264	0.0256	0.0520	0.0288	0.0278	0.0566
		AN	0.0218	0.0308	0.0526	0.0284	0.0276	0.0560
	0.3	LR	0.0274	0.0280	0.0554	0.0248	0.0260	0.0508
		AN	0.0194	0.0328	0.0522	0.0240	0.0256	0.0496
	0.5	LR	0.0318	0.0278	0.0596	0.0270	0.0278	0.0548
		AN	0.0188	0.0356	0.0544	0.0248	0.0262	0.0510
40	0.1	LR	0.0276	0.0256	0.0532	0.0266	0.0284	0.0550
		AN	0.0228	0.0302	0.0530	0.0266	0.0286	0.0552
	0.3	LR	0.0268	0.0256	0.0524	0.0260	0.0292	0.0552
		AN	0.0210	0.0302	0.0512	0.0248	0.0282	0.0530
	0.5	LR	0.0310	0.0278	0.0588	0.0296	0.0321	0.0618
		AN	0.0234	0.0326	0.0560	0.0268	0.0308	0.0576

SS = sample size, CP = censoring proportion, M = method, L = lower tail error probability, U = upper tail error probability, T = total error probability, LR = likelihood ratio, AN = asymptotic normality

DISCUSSION AND CONCLUSIONS

To judge the adequacy of a confidence interval in a simulation study, two important observations have to be made: (1) the attainment of the observed error probability to the nominal one, or at least, conservativeness and, (2) the degree of symmetry of the observed lower and upper tail probabilities (Jennings 1987).

Tables 1, 2, 3 and 4 show that both intervals tend to achieve the nominal level. They tend to be symmetric for the slope parameter. However, intervals based on the asymptotic normality of the maximum likelihood estimator for the intercept parameter have an observed upper and lower tail probabilities that are highly asymmetric, especially for small samples. On the other hand, intervals based on inverted likelihood ratio tests have observed upper and lower tail probabilities that are symmetric even for sample size as small as 20, which is an important consideration for one sided confidence limits (Doganaksoy 1991).

As the censoring proportion increases, the intervals tend to get shorter and have less coverage probabilities. The form of the censoring mechanism and the proportion of censored cases do not appear to have a clear effect on the relative performance of the two kinds of confidence intervals.

REFERENCES

BAIN, L.J. 1978. *Statistical Analysis of Reliability and Life Testing Models.* New York: Marcel Dekker.

BURRIDGE, J. 1981. A note on maximum likelihood estimation for regression models using grouped data. *Journal of Royal Statistical Society,* Series B **43:** 41-45.

DOGANAKSOY, N. 1991. Interval estimation from censored and masked system failure data. *IEEE Transactions and Reliability* **40(3):** 280-285.

DOGANAKSOY, N. and J. SCHMEE. 1991. Comparisons of approximate confidence intervals for the smallest extreme value distribution simple linear regression model under time censoring. *Communications in Statistics-Simulation and Computation* **120(4):** 1085-1113.

ELANDT-JOHNSON, R. and L. JOHNSON. 1979. *Survival Models and Data Analysis.* New York: Wiley.

HAMADA, M. and S.K. TSE. 1988. A note on the existence of maximum likelihood estimates in linear regression models using interval censored data. *Journal of Royal Statistical Society,* Series B **50:** 293-296.

JENNINGS, D.E. 1987. How do we judge confidence-interval adequacy? *The American Statistician* **41(4):** 335-337.

KALBFLEISCH, J. and R. PRENTICE. 1980. *Statistical Analysis of Failure Time Data.* New York: Wiley.

KVAM, P. and F. SAMANIEGO. 1993. On estimating distribution functions using nomination samples. *Journal of American Statistical Association* **88(424):** 1317-1322.

LAWLESS, J.F. 1982. *Statistical Models and Methods for Lifetime Data.* New York: Wiley.

MEEKER, W.Q. 1987. Limited failure population life tests: Applications to nonintegrated circuit reliability. *Technometrics* **29(1)**: 51-65.

PIEGORSCH, W. 1987. Performance of likelihood-based interval estimates for two-parameter exponential samples subject to type 1 censoring. *Technometrics* **29(1)**: 41-49.

VANDER WIEL, S.A. and W.Q. MEEER. 1990. Accuracy of approximate confidence regression data from accelerated life tests. *IEEE Transactions and Reliability* **39(3)**: 346-351.

VENZON, D.J. and S.H. MOOLGAVKAR. 1988. A method for computing profile-likelihood-based confidence intervals. *Applied Statistics* **37**: 87-94.