Wisconsin Clinical Cancer Center

BIOSTATISTICS

Technical Report # 34

October 1985

AN EVALUATION OF GOODNESS-OF-FIT TESTS
FOR THE PROPORTIONAL HAZARDS MODEL

Mari Palta

UNIVERSITY OF WISCONSIN—MADISON
AN EVALUATION OF GOODNESS-OF-FIT TESTS
FOR THE PROPORTIONAL HAZARDS MODEL

Mari Palta
Biostatistics Center
University of Wisconsin-Madison

Technical Report No. 34
October 1985

Typist: Mary T. Metcalf
EVALUATION OF GOODNESS-OF-FIT TESTS

AN EVALUATION OF GOODNESS-OF-FIT TESTS
FOR THE PROPORTIONAL HAZARDS MODEL

Mari Palta
Biostatistics Center
Room 6729 Medical Sciences Center
420 North Charter Street
Madison, Wisconsin 53706

Summary:

Recently, several goodness-of-fit tests have been proposed for the Cox proportional hazards model. The power of these tests is evaluated and set in relation to asymptotic relative efficiency of the Cox model versus alternative procedures. Formulas are derived which can be used for obtaining approximate power of two of the goodness-of-fit tests in given situations. Additionally, Monte Carlo results are presented for a variety of distributions.

Key Words: Cox model; goodness-of-fit; relative efficiency; survival studies

The work was supported by Grant R01-CA18332-10 from the National Cancer Institute to the Wisconsin Clinical Cancer Center.
1. **Introduction:**

Introduced by Cox (1972), the proportional hazards model is now a standard technique in survival analysis. Associated with the model are score tests based on the partial likelihood (see, e.g., Kalbfleisch and Prentice, 1980). The unweighted score test for comparing two or more treatment groups, known as the log rank test, is fully efficient in the proportional hazards (ph) situation. For other types of hazard differences, weighted tests may be more efficient (Tarone and Ware, 1977).

The fit of the Cox model has traditionally been assessed by various graphical methods and by time dependent covariates (Cox, 1972). In recent years, several omnibus test statistics have been proposed for investigating the fit of the model (Schoenfeld, 1980; Andersen, 1982; and Nagelkerke, Oosting and Hart, 1984). The test statistic of Wei (1984) is designed specifically for one binary covariate.

The advantage of goodness-of-fit test statistics over graphical techniques is objectivity and quantitation of fit. It seems important, however, to ensure that the test statistics have reasonable power in practical situations. Few examples of application have been published to date. In addition, it is necessary to know if rejection of the Cox model by the test statistics, is directly associated with situations of unacceptable robustness. Ideally, the tests should be most sensitive in situations where the relative efficiency vis-à-vis alternative procedures is the worst. Problems in this respect have been noted for certain graphical approaches (Crowley and Storer, 1983).

The purpose of the present study, therefore, is to relate the power of the first three goodness-of-fit tests directly to relative efficiency. Asymptotic relative efficiency calculations follow the method of Schoenfeld (1981)
generalized to include strata or covariates (Palta, 1983). Additionally, certain questions of how to apply the tests will be addressed.

2. **Test Statistics Investigated**

The test statistics are evaluated in the setting of two treatment groups, i.e., a Cox model with one binary covariate. We assume equal allocation of subjects. Censoring, when present, is assumed uniform over a given time interval and equal in the two groups.

Formulas for the three test statistics are given here, only for the above situation. The reader is referred to the original publications for more general formulas and greater detail. The Schoenfeld (1980) and Andersen (1982) test statistics both require division of the time axis into $J$ intervals. It is recommended that, for small sample sizes, two intervals be used which divide the death times approximately in half. A wide range of Monte Carlo studies performed (as in Section 5 below) failed to produce any situation where $J > 2$ intervals presented a noticeable advantage. Obviously, with the 0-1 covariate, no division of covariate axes is necessary. The Schoenfeld test statistic is then given by:

$$
\chi^2(J-1) = \sum_{j=1}^{J} \frac{(d_j - e_j)^2}{V_{jj}}
$$

(1)

where $d_j$ = observed number of deaths in treatment 1 interval j

$$
e_j = \sum_{i \in D_j} \frac{\hat{p}_i}{\sum_{i \in D_j} \frac{n_{0i} \exp(\hat{\beta})}{n_{0i} + n_{1i} \exp(\hat{\beta})}}
$$

expected number of deaths in treatment 1 interval j
with \( D_j \) = index set of death times falling in interval \( j \).

\( n_{ki} \) = number in risk set in treatment \( k \) at death time \( i \).

\( \hat{\beta} \) = regression parameter as obtained from the Cox proportional hazards model.

\[
V_{jj} = \sum_{i \in D_j} P_i (1-P_i)
\]

The Andersen test statistic is given by:

\[
\chi^2(J-1) = \sum_{j=1}^{J} \{d_{0j} (\log \hat{\lambda}_{0j} - \hat{\xi}_j)^2 + d_{1j} (\log \hat{\lambda}_{1j} - \hat{\xi}_j - \hat{\alpha})^2\}
\]  
(2)

where \( \hat{\lambda}_{kj} \) = maximum likelihood estimator for the exponential parameter in interval \( j \) treatment \( k \).

\( \hat{\alpha} \) = number of deaths in treatment \( k \) interval \( j \).

\[
\hat{\alpha} = \frac{\sum_{j=1}^{J} (\log \hat{\lambda}_{1j} - \log \hat{\lambda}_{0j})/(d_{1j}^{-1} + d_{0j}^{-1})}{\sum_{j=1}^{J} (d_{1j}^{-1} + d_{0j}^{-1})^{-1}}
\]

estimate of the log hazard ratio under proportional hazards
\[ \xi_j = \frac{d_{0j} (\log(\hat{\lambda}_{0j}) + d_{1j} (\log(\hat{\lambda}_{1j}) - \hat{\alpha}))}{(d_{0j} + d_{1j})} \]

estimate of \( \log(\lambda_{0j}) \) under \( H_0 \)

The Nagelkerke et al. (1984) test statistic is given by

\[ Z = \{ \sum_i U_i(\hat{\beta}) U_i(\hat{\beta}) - \mu \}/\sigma \stackrel{d}{=} N(0,1) \quad (3) \]

where \( U_i(\hat{\beta}) \) is the score from Cox partial likelihood at death time \( i \) given by \(-P_i\) for deaths in treatment 0, and \((1-P_i)\) for deaths in treatment 1. Also

\[ \mu = -\frac{1}{N} \sum_i \{U_i(\hat{\beta})\^2 \}
\]

\[ \sigma^2 = \left[ (N^2 - N + 1) \left\{ \sum_i U_i(\hat{\beta})^2 \right\}^2 - (N^2 + N) \sum_i U_i(\hat{\beta})^4 \right]/(N^3 - N^2) \]

where \( N \) is the total number of deaths, and \( \mu \) and \( \sigma^2 \) are permutation mean and variance respectively.

3. Noncentrality Parameters for Test Statistics

For contiguous alternatives (Schoenfeld, 1980), the noncentrality parameter for (1) can be obtained by replacing \( d_j \), \( e_j \), and \( V_{jj} \) by their expected values under the alternative. A similar approach can be taken for the Andersen statistic (2).

A. Schoenfeld Statistic Noncentrality Parameter Formula

Assume that the time axis is divided into two intervals by the cut
point $T$ ($J = 2$). For simplicity, also assume no censoring (although the same formulas would apply if, where appropriate, survival functions and densities were replaced by the corresponding functions, describing the joint action of death and censoring).

The expected value of $d_1$ in (1) is

$$\overline{f}_1 = n \int_0^T f_1(t) dt$$

where $f_1$ is the density of death in group 1 and $n$ is the sample size per group. Asymptotically, the unconditionally expected value of $e_1$ is

$$\overline{e}_1 = \sum_{D_1} p_i = n \int_0^T [f_0(t) + f_1(t)] \frac{S_1(t) \exp(\beta)}{S_0(t) + S_1(t) \exp(\beta)} dt$$

$$= \int_0^T [f_0(t) + f_1(t)] p(t) dt$$

where $f_0$ is the density of deaths in group 0; $S_0$, $S_1$ are survival functions and $\exp(\beta)$ is the hazard ratio. Where proportional hazards do not hold, $\beta$ is the expected value of the estimate obtained by "erroneously" fitting the proportional hazards model. For practical purposes, we replace $\beta$ by

$$\beta = \frac{1}{2} \int_0^\infty (f_0(t) + f_1(t)) \ln [\Delta(t)] dt$$

(4)

where $\Delta(t)$ is the true hazard ratio at time $t$. Similarly, for the denominators...
\[ V_{ij} = n \int_{\tau_1}^{\tau_2} \left[ f_0(t) + f_1(t) \right] \frac{S_0(t)S_i(t)\exp(\beta)}{\{S_0(t) + S_i(t)\exp(\beta)\}^2} \, dt \]

where for \( j = 1 \): \( \tau_1 = 0 \) and \( \tau_2 = T \), and for \( j = 2 \): \( \tau_1 = T \) and \( \tau_2 = \infty \).

Further simplification leads to the noncentrality parameter

\[ NC = \left\{ \int_0^T f_0(t)p(t) \left[ \frac{\Delta(t)}{\exp(\beta)} - 1 \right] dt \right\}^2 / V_{11} + \left\{ \int_T^\infty f_0(t)p(t) \left[ \frac{\Delta(t)}{\exp(\beta)} - 1 \right] dt \right\}^2 / V_{22} \]  

(5)

Although equation (5) represents a relatively simple expression for the noncentrality of the test statistic (1), interdependence of the parameters makes general interpretation difficult.

Examples of formula (5) evaluated by numerical integration and compared to the results of Monte Carlo studies are presented in Table 1.

B. Evaluation of the Schoenfeld Statistic Noncentrality Parameter for Step Function Hazards

Although formula (5) can be evaluated for specified situations by numerical integration, it is useful to obtain an approximation which can be easily computed as a guideline to the sensitivity of the test statistic. To obtain such an approximation, we evaluate (5) for a step function hazard with breakpoint at time \( T \). The hazard function is defined by

\[ \lambda_k(t) = \begin{cases} \lambda_{k1} & t < T \\ \lambda_{k2} & t > T \end{cases} \quad k = 0,1 \]  

(6)

We denote the hazard ratio \( \lambda_j / \lambda_0 \) in interval \( j \) by \( \Delta_j \) and the overall survival to time \( T \) by \( \overline{S}(T) \). Evaluation of (4) gives
\[ \exp(\beta) = \Delta_1 \left( 1 - \tilde{S}(T) \right), \quad \Delta_2 \tilde{S}(T) = \Delta_1 \left( \frac{\Delta_2}{\Delta_1} \right) \]

which inserted in the first part of (5) gives

\[
\begin{align*}
\int_0^T & \exp(-\lambda_{11} t) \frac{\exp(\beta)}{\exp(-\lambda_{01} t) + \exp(-\lambda_{11} t) \exp(\beta)} \left( \frac{\Delta_1}{\Delta_2} \right)^2 dt \\
& \int_0^T \frac{\exp(-\lambda_{01} t) \exp(-\lambda_{11} t) \exp(\beta)}{\left( \exp(-\lambda_{01} t) + \exp(-\lambda_{11} t) \exp(\beta) \right)^2} \\
& \int_0^T \left[ \lambda_{01} \exp(-\lambda_{01} t) + \lambda_{11} \exp(-\lambda_{11} t) \right] \frac{\exp(-\lambda_{01} t) \exp(-\lambda_{11} t) \exp(\beta)}{\left( \exp(-\lambda_{01} t) + \exp(-\lambda_{11} t) \exp(\beta) \right)^2} \\
& \int_0^T \frac{\exp(-\lambda_{01} t) \exp(-\lambda_{11} t) \exp(\beta)}{\left( \exp(-\lambda_{01} t) + \exp(-\lambda_{11} t) \exp(\beta) \right)^2}
\end{align*}
\]

Although the two integrals can still not be explicitly evaluated, replacing \( \lambda_{01} \) and \( \lambda_{11} \) in the integrand denominators by

\[ \bar{\lambda}_1 = \frac{1}{2} (\lambda_{01} + \lambda_{11}) \]

provides a close approximation for small treatment differences. (Note that our interest in the fit of proportional hazards is also greatest for small treatment differences.)

The expression can then be written in closed form as

\[ \Delta_1 \left( \frac{2}{1 + \Delta_1} \right)^2 \left\{ \frac{1 - \left( \frac{\Delta_2}{\Delta_1} \right)}{\Delta_1} \right\} \left( 1 - \frac{\tilde{S}(T)}{\tilde{S}(T)} \right)^2 \]

\[ \left( 1 - \frac{\tilde{S}(T)}{\tilde{S}(T)} \right)^2 \]

\[ \left( 1 - \frac{\tilde{S}(T)}{\tilde{S}(T)} \right)^2 \]

The numerator of the second part of the noncentrality parameter (5) becomes
\[
\begin{array}{c}
\{ \lambda_{02} \exp[-\lambda_{01} T - \lambda_{02} (t-T)] \} \frac{\exp[-\lambda_{11} T - \lambda_{12} (t-T)] \exp(\beta)}{\exp[-\lambda_{01} T - \lambda_{02} (t-T)] + \exp[-\lambda_{11} T - \lambda_{12} (t-T)] \exp(\beta)}
\end{array}
\]

\[
\Delta_2 \frac{1 - S(T)}{1 - S(T)} \cdot \frac{1}{\Delta_1} \]

with a similar expression for the denominator. Again making the assumption of small treatment differences, we replace \(\lambda_{02}\) and \(\lambda_{12}\) in the integrand denominators by

\[
\overline{\lambda}_2 = .5 (\lambda_{02} + \lambda_{12})
\]

This allows evaluation of the expression as:

\[
\Delta_2 \frac{2}{(1+\Delta_2)^2} \left[ 1 - \frac{\Delta_1}{\Delta_2} \right] \frac{1 - S(T)}{\Delta_1} \exp(\beta) \cdot \frac{1}{S(T)} \cdot \frac{1}{\Delta_2} \frac{1}{1 - S(T)} \cdot \frac{1}{\Delta_1} \frac{1}{1 - S(T)} \cdot \frac{1}{\Delta_2} \frac{1}{1 - S(T)} = (8)
\]

Denoting \((\Delta_2/\Delta_1)\) by \(K\) and observing that the expressions in brackets in (7) and (8) take on the same value for \(K\) as for \(1/K\), we obtain for the NC parameter

\[
\text{NC} = \frac{2\Delta_1}{(1+\Delta_1)^2} \frac{1 - K S(T)}{K S(T)} \left( \frac{1}{1 - S(T)} \right)^2 \frac{1 - \lambda_1 T}{1 - \lambda_1 T} \cdot \frac{2\Delta_2}{(1+\Delta_2)^2} \frac{1 - K 1 - S(T)}{K 1 - S(T)} \left( \frac{1}{1 - S(T)} \right)^2 \frac{1 - \lambda_1 T}{1 - \lambda_1 T} \cdot \frac{e^{-2\lambda_1 T}}{S(T)}
\]

Dividing death times in half implies \(S(T) = .5\) leading to
\[ NC = 4 \frac{(1 - \sqrt{k})^2}{\sqrt{k}} \left\{ \frac{\Delta_1}{(1 + \Delta_1)^2} \left( 1 - e^{-\lambda_1 T} \right)^2 + \frac{\Delta_2}{(1 + \Delta_2)^2} e^{-2\lambda_1 T} \right\} \]

We note that the expression \((1 - \sqrt{k})^2/\sqrt{k}\) increases as \(k\) moves away from 1, and takes on the same value for \(k\) as for \(1/k\). Conversely, increasing treatment differences tend to decrease \(NC\). For small treatment differences, the expression in brackets stays close to \(1/8\), giving

\[ NC = 0.5 \frac{(1 - \sqrt{k})^2}{\sqrt{k}} \quad (9) \]

as a simple expression for the noncentrality parameter.

C. Comparison of Asymptotic and Monte Carlo power for the Schoenfeld Test Statistic

Comparisons were made between asymptotic power of the Schoenfeld test statistic, as obtained by numerical integration of formula (5), asymptotic power obtained by the approximate formula (9) and Monte Carlo power based on 1,000 randomly generated samples. The comparisons were made for sample size 50/group and a variety of alternatives to the proportional hazards situation. The rejection region was based on \(\alpha = .05\). The distributions investigated were:

1) Step hazard functions as described above.

2) Pareto distributions with survival

\[ S_k(t) = (1 + a_k t)^{-b} \quad k = 1, 2, \ b > 0 \]
3) Gompertz distributions with survival

\[ S_k(t) = \exp[-\frac{1}{\gamma_k} (e^{\gamma_k t} - 1)] \quad k = 1, 2 \]

4) Weibull distributions with different shape parameters, given by

\[ S_k(t) = \exp[-\gamma_k t] \quad k = 1, 2 \]

Results of the investigation are given in Table 1.

For non-step function situations, \( K \) was defined as the ratio of

\[ \Lambda_1 = \left\{ \ln S_1(T_{.50}) \right\}/\left\{ \ln S_2(T_{.50}) \right\} \]

\[ \Lambda_2 = \left\{ \ln[S_1(T_{.90})/S_1(T_{.50})] \right\}/\left\{ \ln[S_2(T_{.90})/S_2(T_{.50})] \right\} \]

where \( T_{.50} \) and \( T_{.90} \) are the 50th and 90th percentiles of overall survival, and numerator and denominator were chosen so that \( K > 1 \).

The results in Table 1 indicate generally good agreement between the three methods of obtaining power for the Schoenfeld statistic. Larger discrepancies occur with \( \exp(\beta) \) equal to 2.0 or above and for converging step function. Nonetheless, (9) gives a rough estimate of the power of the goodness-of-fit test. We find that in a sample of 50/group (and no censoring) \( K \) should be at least 3.0 for 80% power of detecting non-proportional hazards. For \( n = 200 \)/group, \( K = 1.75 \) would be detected with 80% power.
D. Andersen Statistic Noncentrality Parameter

Following the above approach, all the component parts of the Andersen statistic are replaced by their expected values. Table 2 shows a comparison between Monte Carlo results and asymptotic power obtained by applying numerical integration to evaluate number of deaths, etc. Results are shown for the Pareto distribution with censoring and for the uncensored Gompertz distribution. It can be seen that asymptotic and Monte Carlo results agree well in the cases investigated. As seen in Section 5 below, however, the Andersen statistic presents problems in the uncensored Pareto situation.

For the step function survival situation, the asymptotic noncentrality parameter becomes

\[
\frac{\ln^2 \frac{\Delta_2}{\Delta_1}}{\exp(-2\lambda_1 T) [1 - 2S(T) + \exp(-2\lambda_1 T)]}
\frac{1}{2S(T) [1 - 2S(T) + \exp(-2\lambda_1 T)] + \exp(-2\lambda_1 T) [2 - 2S(T)]}
\]

which for \( T = T_{.50} \) and small treatment differences simplifies to

\[
NC_A = \frac{1}{8} \ln^2 k \tag{10}
\]

This quantity remains very close to (9), the approximate noncentrality parameter for the Schoenfeld test statistic, for a wide range of \( k \).

However, (10) provides a less good approximation for the power of the Andersen statistic in situations other than step function hazards. As will be seen in the examples, the power of the Andersen statistic tends to be lower than that of the Schoenfeld statistic in many of these situations.
4. **Relationship to Asymptotic Relative Efficiency:**

We now wish to relate power of the goodness-of-fit statistics to asymptotic relative efficiency of the proportional hazards model vis-a-vis alternative procedures. The asymptotic relative efficiency of the log rank statistic versus an optimally weighted test has been derived by Schoenfeld (1981). For the uncensored situation, this relative efficiency is given by:

$$ \text{ARE} = \frac{\int_0^\infty \left( \ln \left( \frac{\lambda_1(t)}{\lambda_0(t)} \right) g(t) dt \right)^2}{\int_0^\infty \left( \ln \left( \frac{\lambda_1(t)}{\lambda_0(t)} \right) \right)^2 g(t) dt} $$  \hspace{1cm} (11)

where $g$ is the overall density function for deaths. For non-proportional hazards, numerical integration is generally required to evaluate this expression. The results have been shown to be accurate in moderate size samples (Palta and Amini, 1982; Palta 1983). Additional Monte Carlo studies were undertaken here, which further verified the accuracy of (11). An extension of formula (11), (Palta, 1983), was used to obtain the ARE of the log rank test versus an appropriate Cox model with two covariate (see examples below).

For the step function hazard situation with $T = T_{.50}$, formula (11) simplifies to:

$$ \text{ARE} = 1 - \left\{ 1 + \left[ \frac{1}{2} \frac{\ln \Delta_2}{\ln K} \right] \right\}^{-1} \quad \text{where} \quad K = \frac{\Delta_2}{\Delta_1} $$

This formula is not a good approximation for non-step function hazards, but shows that at least for step function hazards: 1) the asymptotic relative efficiency inversely parallels the Schoenfeld noncentrality parameter in that ARE decreases
with increasing \( |K-1| \) and increases with increasing \( |\Delta-1| \). 2) On the other hand, the ARE is more sensitive than the noncentrality parameter to \( \Delta_1 \) and \( \Delta_2 \) (for given \( K \)), so that a wide range of ARE is possible for relatively constant noncentrality.

5. Additional Monte Carlo Results for Selected Situations:

Additional Monte Carlo studies were undertaken to further compare the test statistics. Distribution assumptions were as in Section 1 with uniform censoring imposed in cases indicated. An additional situation investigated was exponential survival with a \( N(0,1) \) covariate related by

\[
\lambda_k(x) = \lambda_k \exp(-bx) \quad \text{where} \quad X \sim N(0,1) \quad \text{and} \quad k = 0,1
\]

(12)

It is of interest to note that this situation is closely related to the Pareto situation, in that (12) describes an exponential distribution with log normal distribution on the parameter and the Pareto distribution can be seen as exponential with gamma distributed parameter (Turnbull, 1974).

Monte Carlo results in Table 3 were chosen to be representative and to illustrate certain features of the statistics. Each entry is based on 1,000 independent generations. Parameter sets below the horizontal line are null situations in each case. It is seen that the Andersen statistic has unsatisfactory behavior in uncensored situations with a long tail in the survival. This is remedied when censoring is imposed. Nonetheless, the Schoenfeld statistic seems most satisfactory overall. Nagelkerke statistic power seems particularly low in the situations investigated. In general, adequate power for \( n < 200 \) group is obtained only for rather extreme situations. The
situation with a normally distributed covariate seems particularly difficult to detect, despite very low ARE versus the full Cox model.

6. **Conclusions:**

The three goodness-of-fit statistics for the Cox proportional hazards model are useful only in large sample sizes. The Schoenfeld statistic seems to exhibit the most satisfactory behavior overall. A rough idea of the power of the Schoenfeld statistic can be obtained from the ratio of the hazard ratios over the first versus the last "half" of survival.

Although there is a correspondence between the power of the test statistics and relative efficiency versus alternative procedures in most practical situations considered, it is possible to construct step function hazard situations with low ARE of the log rank test and low power of rejection by the goodness-of-fit tests.
Table 1

Comparison of Asymptotic and Monte Carlo Power for Schoenfeld Statistic

\[ n = 50/\text{group}, \alpha = .05, 1,000 \text{ simulations/cell} \]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.34</td>
<td>Pareto (1.125; .5; b=2)</td>
<td>.90</td>
<td>1.6</td>
<td>.11</td>
<td>.12</td>
<td>.11</td>
</tr>
<tr>
<td>1.42</td>
<td>Gompertz (1.0; .5)</td>
<td>.63</td>
<td>1.3</td>
<td>.16</td>
<td>.17</td>
<td>.14</td>
</tr>
<tr>
<td>1.50</td>
<td>Pareto (1.5; .5; b=2)</td>
<td>.88</td>
<td>2.2</td>
<td>.13</td>
<td>.17</td>
<td>.18</td>
</tr>
<tr>
<td>1.57</td>
<td>Converging step ( (\Delta_1 = 1.6; \Delta_2 = 1.15) )</td>
<td>.73</td>
<td>1.4</td>
<td>.16</td>
<td>---</td>
<td>.18</td>
</tr>
<tr>
<td>1.60</td>
<td>Diverging step ( (\Delta_1 = 1.6; \Delta_2 = 1.86) )</td>
<td>.73</td>
<td>1.5</td>
<td>.22</td>
<td>---</td>
<td>.22</td>
</tr>
<tr>
<td>1.67</td>
<td>Pareto (2.5; 1.0; b=1.0)</td>
<td>.64</td>
<td>1.3</td>
<td>.25</td>
<td>.26</td>
<td>.25</td>
</tr>
<tr>
<td>1.71</td>
<td>Pareto (2.0; .5; b=2.0)</td>
<td>.87</td>
<td>2.6</td>
<td>.21</td>
<td>.26</td>
<td>.27</td>
</tr>
<tr>
<td>3.3</td>
<td>Weibull (1.0; 2.0)</td>
<td>&lt;.10</td>
<td>1.4</td>
<td>.85</td>
<td>.87</td>
<td>.86</td>
</tr>
<tr>
<td>4.0</td>
<td>Diverging step ( (\Delta_1 = 1; \Delta_2 = 4.0) )</td>
<td>.5</td>
<td>2.0</td>
<td>.71</td>
<td>---</td>
<td>.94</td>
</tr>
<tr>
<td>Converging step ( (\Delta_1 = 4.0; \Delta_2 = 1.0) )</td>
<td>.5</td>
<td>2.0</td>
<td>.88</td>
<td>---</td>
<td>.94</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

Comparison of Asymptotic and Monte Carlo (MC) Results for Andersen Statistic

\[ n = 50/\mathrm{group}, \ \text{generations} \ 1,000, \ J = 2, \ a = .05 \]

### Compertz (Uncensored)

<table>
<thead>
<tr>
<th>( e^\beta )</th>
<th>( Y_1 )</th>
<th>( Y_2 )</th>
<th>Asympt. Power</th>
<th>MC Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3</td>
<td>1.0</td>
<td>.5</td>
<td>.09</td>
<td>.09</td>
</tr>
</tbody>
</table>

### Pareto (40-50\% Censoring)

<table>
<thead>
<tr>
<th>( e^\beta )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b )</th>
<th>Asympt. Power</th>
<th>MC Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6</td>
<td>1.125</td>
<td>.5</td>
<td>2</td>
<td>.06</td>
<td>.06</td>
</tr>
<tr>
<td>1.4</td>
<td>2.5</td>
<td>1.0</td>
<td>1</td>
<td>.08</td>
<td>.09</td>
</tr>
<tr>
<td>1.3</td>
<td>2.0</td>
<td>6.0</td>
<td>0.5</td>
<td>.11</td>
<td>.12</td>
</tr>
</tbody>
</table>
Table 3

Selected Monte Carlo Results

% Rejected at \( \alpha = .05 \), 1,000 Independent Generations/Cell

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameters</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b )</th>
<th>% ARE vs Opt Wts</th>
<th>% ARE vs full Cox Model</th>
<th>% Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARETO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=50/group</td>
<td>uncensored</td>
<td>1.0</td>
<td>2.5</td>
<td>1.0</td>
<td>78%</td>
<td>--</td>
<td>26%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.0</td>
<td>6.0</td>
<td>0.5</td>
<td>64%</td>
<td>--</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.0</td>
<td>6.0</td>
<td>0.5</td>
<td>--</td>
<td></td>
<td>6.1%</td>
</tr>
<tr>
<td>n=200/group</td>
<td>censoring</td>
<td>1.0</td>
<td>2.5</td>
<td>1.0</td>
<td>93%</td>
<td>--</td>
<td>59%</td>
</tr>
<tr>
<td>~ 50%</td>
<td></td>
<td>2.0</td>
<td>6.0</td>
<td>0.5</td>
<td>86%</td>
<td>--</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.0</td>
<td>6.0</td>
<td>0.5</td>
<td>--</td>
<td></td>
<td>6.0%</td>
</tr>
<tr>
<td>EXPONENTIAL</td>
<td>WITH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COVARIATE</td>
<td>( \lambda_1 )</td>
<td>1.5</td>
<td>4.0</td>
<td>2.0</td>
<td>83%</td>
<td>25%</td>
<td>10%</td>
</tr>
<tr>
<td>n=50/group</td>
<td>uncensored</td>
<td>2.5</td>
<td>2.5</td>
<td>2.0</td>
<td>5.6%</td>
<td>25%</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>5.0</td>
<td>0.0</td>
<td>4.8%</td>
<td>6.9%</td>
<td>7.4%</td>
</tr>
<tr>
<td>n=200/group</td>
<td>uncensored</td>
<td>1.5</td>
<td>4.0</td>
<td>2.0</td>
<td>21%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEIBULL</td>
<td>DISTRIBUTIONS</td>
<td>( \gamma_1 )</td>
<td>( \gamma_2 )</td>
<td></td>
<td></td>
<td>84%</td>
<td>96%</td>
</tr>
<tr>
<td>n=50/group</td>
<td>uncensored</td>
<td>1.0</td>
<td>0.5</td>
<td>&lt; 10%</td>
<td>85%</td>
<td>87%</td>
<td>5.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>2.0</td>
<td></td>
<td>85%</td>
<td>87%</td>
<td>5.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td></td>
<td></td>
<td>5.7%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.0</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td>5.7%</td>
</tr>
<tr>
<td>n=50/group</td>
<td>~ 40%</td>
<td>censoring</td>
<td>1.0</td>
<td>2.0</td>
<td>38%</td>
<td></td>
<td>62%</td>
</tr>
</tbody>
</table>

18
REFERENCES


