Survival Analysis Assignment 1

Master of Statistics (M.Stat.) Year II, Session 2020-21

Name: Subhrajyoty Roy
Roll: MB1911

February 27, 2021

Problem 1. Compare performances of the exact confidence interval for the parameter θ for Type 2 censoring scheme and different asymptotic confidence intervals for θ for Type 1 censoring scheme when the lifetime distribution is an exponential distribution with mean θ.

Solution. Let, r denotes the total number of failures and V denotes the total time on tests. Under both censoring scheme, the maximum likelihood estimate of θ is given by $\widehat{\theta} = V/r$.
In Type 2 censoring case, the exact $100(1-\alpha)\%$ confidence interval for θ is given by

$$CI_5 = \left(\frac{2r\widehat{\theta}}{\chi^2_{2r,1-\alpha/2}}, \frac{2r\widehat{\theta}}{\chi^2_{2r,\alpha/2}} \right),$$

where $\chi^2_{a,\alpha}$ denotes the α-th quantile of chi-squared distribution with a degrees of freedom.
On the other hand, for Type 1 censoring scheme, 4 types of asymptotic $100(1-\alpha)\%$ confidence interval exist.

1. Based on the expected Fisher's information, one has

$$CI_1 = \left(\widehat{\theta} - \frac{z_{1-\alpha/2}}{\sqrt{n}} \frac{\widehat{\theta}}{\sqrt{1 - \exp\left[-T_0/\widehat{\theta}\right]}}, \widehat{\theta} + \frac{z_{1-\alpha/2}}{\sqrt{n}} \frac{\widehat{\theta}}{\sqrt{1 - \exp\left[-T_0/\widehat{\theta}\right]}} \right),$$

   where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-th quantile of the standard normal distribution.

2. Similar confidence interval based on the observed Fisher's information is given by

$$CI_2 = \left(\widehat{\theta} - \frac{z_{1-\alpha/2}}{\sqrt{r}}\widehat{\theta}, \widehat{\theta} + \frac{z_{1-\alpha/2}}{\sqrt{r}}\widehat{\theta} \right).$$

3. Cox (1953) proposed another confidence interval as

$$CI_3 = \left(\frac{2r\widehat{\theta}}{\chi^2_{2r+1,1-\alpha/2}}, \frac{2r\widehat{\theta}}{\chi^2_{2r+1,\alpha/2}} \right)$$

4. Based on likelihood ratio test for $H_0 : \theta = \theta_0$, one may obtain that

$$\mathbb{P}\left(2r\log(\theta) + 2\frac{\sum_{i=1}^n x_i}{\theta} < \chi^2_{1,1-\alpha} + 2r\log(\widehat{\theta}) + 2r \right) \approx (1-\alpha),$$

   where x_i is the i-th censored lifetime. Hence, one can numerically invert this to obtain bounds for θ such that $\mathbb{P}(L(\widehat{\theta}) < \theta < U(\widehat{\theta})) \approx (1-\alpha)$, which gives an approximate confidence interval as $CI_4 = (L(\widehat{\theta}), U(\widehat{\theta}))$.

To compare performances between these two different setups, we choose r in Type 2 censoring as $[n(1-\exp(-T_0/\theta))]$ which is the expected number of failures under Type 1 censoring. To compare between different confidence interval schemes, we obtain $B = 1000$ resamples of sample size n (to be varied) and compute the confidence intervals based on each of these schemes, and finally obtain the approximate coverage probability (proportion of B resamples for which the confidence interval contains the true value of θ) and average length of the intervals (averaged over all B resamples). We shall choose T_0 to be such that a pre-defined proportion of samples in expectation is censored, $e^{-T_0/\theta} = p$, i.e. $T_0 = \theta \log(1/p)$, where p is a pre-defined constant ($5\%, 10\%$ and 20%). As true parameter θ, we choose $\theta = 1$ and to see the effect of the sample size, we choose $n = 5, 10, 25, 50, 100$. Table 1 and Table 2 contains the detailed description of the approximate coverage probabilities and the average length of the confidence intervals under different censoring proportion p and sample size n.

Proportion of Censoring (**p**)	Sample size (**n**)	$\mathbf{CI_1}$	$\mathbf{CI_2}$	$\mathbf{CI_3}$	$\mathbf{CI_4}$	$\mathbf{CI_5}$
	5	0.848	0.847	0.936	0.943	0.949
	10	0.894	0.894	0.936	0.942	0.939
5%	25	0.938	0.94	0.945	0.949	0.955
	50	0.951	0.953	0.956	0.959	0.954
	100	0.949	0.949	0.953	0.952	0.952
	5	0.849	0.847	0.936	0.942	0.952
	10	0.891	0.894	0.939	0.945	0.95
10%	25	0.936	0.935	0.949	0.95	0.956
	50	0.949	0.948	0.961	0.959	0.955
	100	0.95	0.949	0.954	0.955	0.951
	5	0.856	0.85	0.938	0.941	0.952
	10	0.897	0.9	0.943	0.946	0.952
20%	25	0.939	0.94	0.957	0.959	0.961
	50	0.947	0.95	0.966	0.967	0.954
	100	0.947	0.947	0.945	0.944	0.943

Table 1: Approximate coverage probabilities of different confidence intervals

From the metrics shown in Table 1 and Table 2, we may conclude the following.

1. As the sample size n increases, all of these confidence intervals perform equivalently, since all of these confidence intervals are asymptotically equivalent.

2. For small sample sizes, Cox's confidence interval and the interval obtained through LRT approximately attains the 95% coverage probability. The exact confidence interval for Type 2 censoring almost always attains the 95% coverage probability irrespective of the sample size and proportion of censoring.

3. Proportion of censoring do not have a considerable effect on the coverage probabilities. However, the length of the confidence intervals increases as proportion of censored datapoint increases.

4. In terms of length of the confidence intervals, the exact confidence interval for Type 2 censoring (CI_5) has highest length. Asymptotic CI's based on observed and expected information matrix have similar lengths.

Proportion of Censoring (**p**)	Sample size (**n**)	$\mathbf{CI_1}$	$\mathbf{CI_2}$	$\mathbf{CI_3}$	$\mathbf{CI_4}$	$\mathbf{CI_5}$
5%	5	1.839	1.828	2.31	2.403	3.857
	10	1.282	1.279	1.438	1.489	1.756
	25	0.808	0.807	0.846	0.856	0.918
	50	0.572	0.571	0.585	0.589	0.608
	100	0.404	0.403	0.408	0.409	0.417
10%	5	1.984	1.976	2.57	2.32	5.568
	10	1.338	1.336	1.515	1.573	1.928
	25	0.836	0.835	0.878	0.89	0.973
	50	0.59	0.589	0.604	0.608	0.631
	100	0.416	0.415	0.42	0.422	0.43
20%	5	2.255	2.241	3.112	1.955	5.568
	10	1.452	1.446	1.672	1.749	2.166
	25	0.896	0.895	0.947	0.962	1.038
	50	0.629	0.629	0.647	0.651	0.675
	100	0.443	0.443	0.449	0.451	0.458

Table 2: Average length of the different confidence intervals

5. As the sample size increases, the length of the confidence interval shrinks for each of the method.

In summary, for small samples, CI_1 and CI_2 based on observed and expected Fisher's information matrix performs poorly, whereas, Cox's confidence interval and the CI obtained from inverting LRT performs better. The exact CI based on Type 2 censoring achieves the correct coverage probability, but at the cost of a longer interval.

Problem 2. Compare finite sample performances of the different asymptotic tests for H_0 : $\lambda = 1$ vs $H_1 : \lambda \neq 1$ under random censoring scheme where the lifetime distribution is $\exp(\lambda)$ (λ is the rate parameter).

Solution. Let d denotes the total number of failures and V denotes the total time on tests. Then, under random censoring scheme where the lifetime distribution is exponential with rate parameter λ, the maximum likelihood estimator of λ is $\widehat{\lambda} = d/V$. Based on this, there are three types of asymptotic tests for testing the null hypothesis $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$.

1. **Score Test** at level-α rejects the null hypothesis if

$$T_s = \frac{(d - V\lambda_0)^2}{d} \frac{\widehat{\lambda}^2}{\lambda_0^2} > \chi^2_{1,1-\alpha}$$

   where $\chi^2_{1,1-\alpha}$ is the $(1-\alpha)$-th quantile of χ^2 distribution with 1 degrees of freedom.

2. **Wald's Test** at level-α rejects the null hypothesis if

$$T_w = \frac{(d - V\lambda_0)^2}{d} > \chi^2_{1,1-\alpha}$$

3. **Likelihood Ratio Test (LRT)** at level-α rejects the null hypothesis if

$$T_{LR} = 2(\lambda_0 - \widehat{\lambda})V - 2d\log\left(\frac{\lambda_0}{\widehat{\lambda}}\right) > \chi^2_{1,1-\alpha}$$

To compare performances of these tests, without loss of generality we fix $\alpha = 0.05$ and $\lambda_0 = 1$. Then, we perform B resamples, where in each of the resamples, we generate data according to the random censoring scheme for a range of values of λ, and then note where the aforementioned tests rejects or fails to reject the null hypothesis. Then, we can approximate the power of a test at $\lambda = \lambda_i$ as the proportion of resamples where the test rejects the null hypothesis for datasets generated from $\lambda = \lambda_i$. While this gives us an approximate power curve for each of these tests, we consider two metrices for ease of comparison.

1. The approximate size of the test i.e. the proportion of resamples for which a test rejects the null hypothesis when the data is generated from the parameter setting $\lambda = 1$.

2. The approximate area under the power curve, for which we shall use trapezoidal rule, i.e.

$$\int_1^{\lambda_{\max}} \text{Power}(\lambda)d\lambda \approx \frac{1}{2(\Delta\lambda)}\left(\text{Power}(\lambda_1) + 2\text{Power}(\lambda_2) + \cdots + 2\text{Power}(\lambda_{n-1}) + \text{Power}(\lambda_n)\right)$$

   Since this area depends on the choice of $\lambda_{\max}$, the maximum value of λ under the alternative, we consider $\frac{1}{\lambda_{\max}}\int_1^{\lambda_{\max}} \text{Power}(\lambda)d\lambda$ instead which will reside between 0 and 1.

Now, finally for sample sizes, we shall use $n = 5, 10$ and 25. Let T and C denote the random variables denoting lifetime and censoring. Then, we consider three different types of distribution for censoring.

1. If $C \sim \exp(\tau)$ be the censoring distribution, then the expected proportion of censoring is

$$\mathbb{P}(T > C) = \int_0^\infty e^{-\lambda c} \tau e^{-\tau c} dc = \frac{\tau}{\tau + \lambda}$$

   Therefore, once we fix p as the expected proportion of censoring, then we choose $\tau = p\lambda/(1-p)$.

2. If $C \sim \text{Unif}(0, \theta)$, then

$$\mathbb{P}(T > C) = \int_0^\theta e^{-\lambda c} \theta^{-1} dc = \frac{(1 - e^{-\lambda\theta})}{\theta} \approx \frac{\lambda\theta - \lambda^2\theta^2/2}{\theta} = \lambda - \frac{\lambda^2}{2}\theta$$

   Therefore, fixing the expected proportion of censoring as p, we may choose $\theta = 2(\lambda - p)/\lambda^2$.

3. If $C \sim \text{Weibull}(\tau, 2)$, then

$$\mathbb{P}(T > C) = 1 - \mathbb{P}(T \leq C) = 1 - \int_0^\infty \lambda e^{-(\tau t)^2 - \lambda t} dt = 1 - \frac{\sqrt{\pi}}{\tau} e^{\lambda^2/2\tau^2} \left(1 - \Phi(\lambda/2\tau)\right)$$

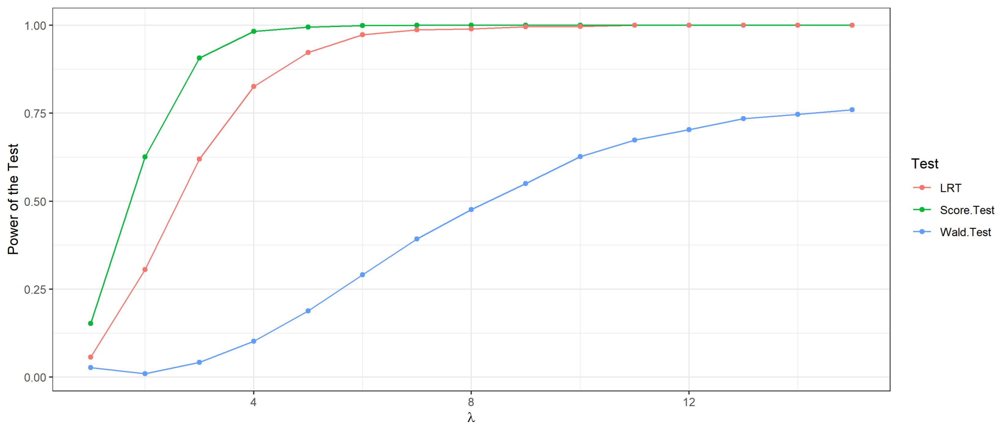
   where Φ is the normal cdf. We can fix this probability as p and numerically solve for τ to obtain the required parameter of Weibull distribution.

We shall consider only 3 proportion of censoring, namely $5\%, 20\%$ and 40%. The performances of these tests when the censoring distribution is exponential, are tabulated in Table 3. Figure 1 shows the approximate power curve of the aforementioned asymptotic tests under different sample sizes, where the censoring distribution is exponential with the parameter adjusted so that the expected censoring proportion is 5%.

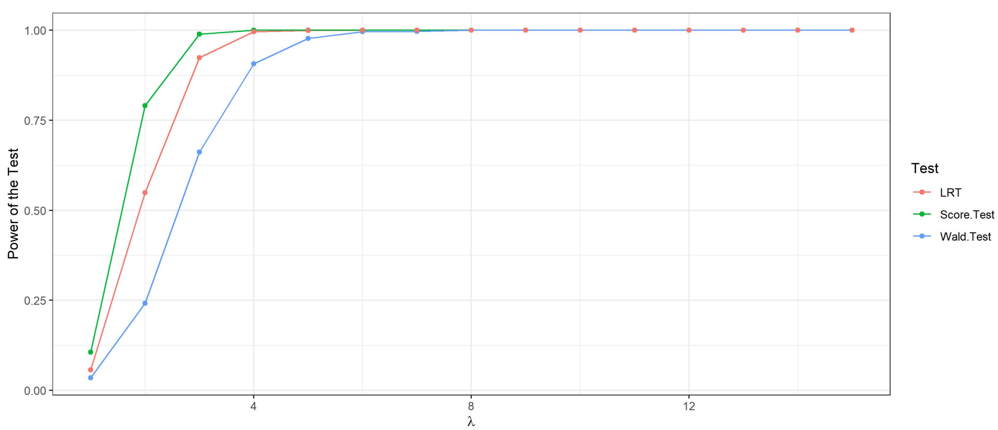
Proportion of Censoring	n	Approximate Size			Area under the power curve		
		Score test	Wald's test	LRT	Score test	Wald's test	LRT
5%	5	0.153	0.027	0.057	0.846	0.394	0.798
	10	0.106	0.035	0.057	0.859	0.811	0.846
	25	0.074	0.04	0.048	0.866	0.861	0.864
20%	5	0.158	0.04	0.053	0.832	0.2	0.748
	10	0.115	0.047	0.051	0.852	0.748	0.831
	25	0.073	0.045	0.047	0.865	0.856	0.862
40%	5	0.148	0.047	0.056	0.802	0.064	0.647
	10	0.116	0.053	0.041	0.84	0.56	0.79
	25	0.076	0.049	0.043	0.861	0.839	0.854

Table 3: Small sample performances of score test, Wald's test and LRT for exponential distribution as the censoring distribution.

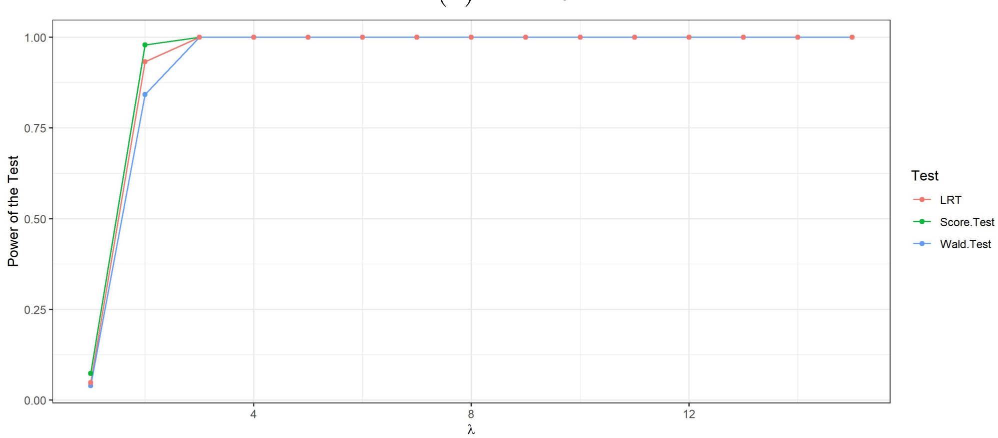
Similarly, Table 4 and Table 5 contains the details of the simulation where the censoring distribution are Uniform and Weibull respectively.
From the results shown in Table 3,Table 4 and Table 5, we conclude that



(a) n = 5



(b) n = 10



(c) n = 25

Figure 1: Approximate power curve of the three asymptotic tests under 5% censored data for different sample sizes

Proportion of Censoring	n	Approximate Size			Area under the power curve		
		Score test	Wald's test	LRT	Score test	Wald's test	LRT
5%	5	0.14	0.074	0.067	0.78	0.068	0.615
	10	0.13	0.061	0.06	0.834	0.519	0.777
	25	0.076	0.052	0.047	0.86	0.833	0.851
20%	5	0.142	0.078	0.065	0.777	0.067	0.609
	10	0.124	0.065	0.058	0.832	0.512	0.774
	25	0.067	0.052	0.044	0.859	0.83	0.851
40%	5	0.152	0.094	0.073	0.774	0.063	0.603
	10	0.125	0.077	0.053	0.83	0.5	0.767
	25	0.082	0.056	0.043	0.859	0.826	0.848

Table 4: Small sample performances of score test, Wald's test and LRT for uniform distribution as the censoring distribution.

1. Generally, Score test has more power than Likelihood Ratio Test (LRT) and LRT has more power than Wald's test.

2. As the sample size n increases, all of the tests perform similarly.

3. While LRT and Wald's test have approximately 5% size for small samples, Score test usually has higher size than intended.

4. As the proportion of censoring increases, the power of all three tests decreases when the censoring distribution is exponential or Weibull distribution. However, when the censoring distribution is uniform, the effect of the proportion of censoring is very limited.

5. Under heavy censoring (40% censored data), Wald's test have very small power (similar to its size) for small samples.

In summary, Wald's test should not be considered if the sample size is small. Likelihood Ratio test should be preferred in such scenario.

Proportion of Censoring	n	Approximate Size			Area under the power curve		
		Score test	Wald's test	LRT	Score test	Wald's test	LRT
5%	5	0.149	0.081	0.068	0.844	0.485	0.8
	10	0.107	0.063	0.053	0.856	0.809	0.842
	25	0.077	0.052	0.042	0.865	0.857	0.862
20%	5	0.155	0.083	0.064	0.845	0.486	0.801
	10	0.112	0.06	0.05	0.856	0.81	0.843
	25	0.072	0.048	0.039	0.865	0.858	0.862
40%	5	0.144	0.06	0.064	0.845	0.486	0.803
	10	0.1	0.051	0.046	0.858	0.811	0.844
	25	0.07	0.042	0.047	0.865	0.859	0.863

Table 5: Small sample performances of score test, Wald's test and LRT for Weibull distribution as the censoring distribution.

Thank You