Hybrid Censoring; An Introduction [2]

Debasis Kundu

Department of Mathematics & Statistics
Indian Institute of Technology Kanpur

23-rd November, 2010

[2] This is a joint work with N. Balakrishnan

Outline

Outline

Type-I Hybrid Censoring Scheme

- Type-I and Type-II censoring schemes are the two most popular censoring schemes.
- In Type-I censoring scheme, the experimental time is fixed, but the number of failures is a random variable.
- In Type-II censoring scheme, number of failures is fixed, but the experimental time is a random variable.
- The mixture of Type-I and Type-II censoring schemes is known as the hybrid censoring scheme.

Type-I HCS

A total of n units is placed on a life testing experiment. The lifetimes of the sample units are independent and identically $(i.i.d.)$ random variables . Let the ordered lifetimes of these items be denoted by $X_{1:n}, \cdots, X_{n:n}$ respectively. The test is terminated when a pre-chosen number, $r < n$, out of n items are failed, or when a pre-determined time, T, on test has been reached, $i.e.$ the test is terminated at a random time $T_* = \min\{X_{r:n}, T\}$. It is also usually assumed that the failed items are not replaced.

Available Data

When the data are Type-I hybrid censored, we have one of the following two types of observations;

Case I: $\{x_{1:n} < \cdots < x_{r:n}\}$ if $x_{r:n} \leq T$

Case II: $\{x_{1:n} < \cdots < x_{d:n}\}$ if $x_{r:n} > T$,

here d denotes the number of observed failures that occur before time point T.

Associated Problems

- Under parametric assumption of the underlying distribution, estimation of the parameter(s)
- Finding the distribution of the parameter(s)
- Finding the confidence interval(s)
- Finding the optimum censoring scheme (Choosing r and T).

Exponential Distribution

- It is assumed that the lifetime distribution is a one parameter exponential distribution with mean θ, *i.e.*, the PDF is

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}; \quad x > 0, \quad \theta > 0$$

- Based on the observed sample, it can be easily seen that the MLE of θ does not exist if $d = 0$, and if $d > 0$, it is given by

$$\widehat{\theta} = \begin{cases} \frac{1}{d}\sum_{i=1}^{d} x_{i:n} + (n-d)T & \text{if} \quad x_{r:n} \leq T \\ \frac{1}{r}\sum_{i=1}^{d} x_{i:n} + (n-r)x_{r:n} & \text{if} \quad x_{r:n} > T. \end{cases}$$

Conditional Distribution of the MLE

Now using the moment generating functions of gamma and shifted gamma distribution, one can obtain the PDF of $\widehat{\theta}$ for $0 < x < nT$ as

$$f_{\widehat{\theta}}(x) = (1-q^n)^{-1}\left[\sum_{d=1}^{r-1}\sum_{k=0}^{d} C_{k,d}\, g\left(x - T_{k,d}; \frac{d}{\theta}, d\right) + g\left(x; \frac{r}{\theta}, r\right) + r\binom{n}{r}\sum_{k=1}^{r}\frac{(-1)^k q^{n-r+k}}{n-r+k}\binom{r-1}{k-1} g\left(x - T_{k,r}; \frac{r}{\theta}, r\right)\right],$$

where $q = e^{-T/\theta}$, $C_{k,d} = (-1)^k\binom{n}{d}\binom{d}{k}q^{n-d+k}$, $T_{k,d} = (n-d+k)T/d$, and $g(.)$ is a gamma PDF.

Conditional Distribution of the MLE

It is clear that the PDF of $\widehat{\theta}$, is a generalized mixture of gamma and shifted gamma PDFs, when the mixing coefficients may be negative also. The cumulative distribution function (CDF) and the survival function (SF) of $\widehat{\theta}$ can be obtained in terms of incomplete gamma function.

Observations

- Because of the explicit expression of the PDF of the MLE, different moments can be easily obtained. It is observed that the MLE of θ is a biased estimate of θ.

- It can also be observed that as $T \to \infty$,

$$f_{\widehat{\theta}}(x) = g\left(x; \frac{r}{\theta}, r\right).$$

  It is the well known result that $\dfrac{2r\widehat{\theta}}{\theta}$ has a chi-square distribution with $2r$ degrees of freedom.

- Substitution $r = n$, we get the PDF of the MLE of θ for the usual Type-I censored case.

Exact Confidence Intervals

In an interesting paper, Fairbanks, Madsan and Dykstra in their JASA paper, proposed a set of two sided 100(1-α)% confidence intervals for θ in the Type-I HCS, as follows;

$$\left[\frac{2S}{\chi^2_{2,\alpha/2}}, \infty\right] \qquad \text{if} \qquad d = 0$$

$$\left[\frac{2S}{\chi^2_{2d+2,\alpha/2}}, \frac{2S}{\chi^2_{2d,1-\alpha/2}}\right] \quad \text{if} \quad 1 \leq d \leq r-1,$$

$$\left[\frac{2S}{\chi^2_{2r,\alpha/2}}, \frac{2S}{\chi^2_{2r,1-\alpha/2}}\right] \qquad \text{if} \qquad d = r,$$

here S is the total time on test, *i.e.*

$$S = \begin{cases} \sum_{i=1}^{d} x_{i:n} + (n-d)T & \text{if} \quad x_{r:n} \leq T \\ \sum_{i=1}^{d} x_{i:n} + (n-r)x_{r:n} & \text{if} \quad x_{r:n} > T, \end{cases}$$

Comments

The authors provided a formal proof in the 'with replacement' case, and mentioned that the proof can be extended for the without replacement' case also. It is not very clear that how the proof will be in the 'without replacement' case, as one of the major assumption in their proof ('with replacement' case) is that for $0 \leq j \leq r-1$,

$$P\{\text{j items fail at the decision time}\} = \frac{e^{-nT/\theta}\,(nt/\theta)^j}{j!}.$$

Clearly, the above assumption is no longer valid for the 'without replacement' case and their proof very much depend on that assumption.

Exact Confidence Interval

Based on the exact distribution of $\widehat{\theta}$, and based on the fact that $P_\theta(\widehat{\theta} > b)$ is an increasing function of θ for fixed b, the exact two sided 100(1-α)% symmetric confidence interval of θ, then θ_L and θ_U can be obtained by solving the following two non-linear equations;

$$\frac{\alpha}{2} = F_{\theta_L}(\widehat{\theta}), \quad 1 - \frac{\alpha}{2} = F_{\theta_U}(\widehat{\theta}).$$

Here

$$F_\theta(x) = P(\widehat{\theta} \leq x).$$

θ_L and θ_U need to be computed numerically by solving the above two non-linear equations.

Bayes Estimates

Bayesian solution seems to be a very natural choice in this case, and fortunately it turns out to be quite simple.

Prior:
Draper and Guttman assumed that θ has a inverted gamma prior with the following PDF;

$$\pi(\theta) = \frac{\lambda^{\beta}}{\Gamma(\beta)}\theta^{-(\beta+1)}e^{-\lambda/\theta}; \quad \theta > 0,$$

here $\beta > 0$, and $\lambda > 0$ are the hyper parameters.

Based on the above prior, the Bayes estimate with respect to squared error loss function is

$$\widehat{\theta}_{Bayes} = \frac{(S_d + \lambda)}{(d + \beta - 1)},$$

if $(d + \beta) > 1$, where

$$S_d = \begin{cases} nT & \text{if} \quad d = 0 \\ \sum_{i=1}^{d} x_{i:n} + (n-d)T & \text{if} \quad 1 \leq d \leq r-1 \\ \sum_{i=1}^{r} x_{i:n} + (n-r)x_{r:n} & \text{if} \quad d = r. \end{cases}$$

Under the noninformative prior $\beta = \lambda = 0$, the Bayes estimate matches with the MLE.

Credible Interval

Interestingly, a 100(1-α)% credible interval of θ can be easily obtained as

$$\left(\frac{2(S_d + \lambda)}{\chi^2_{2(d+\beta),\alpha/2}}, \frac{2(S_d + \lambda)}{\chi^2_{2(d+\beta),1-\alpha/2}} \right),$$

if $d + \beta > 0$, and here $\chi^2_{m,\alpha}$ denotes the point of the central χ^2 distribution with m degrees of freedom that leaves an area α in the upper tail.

Two-Parameter Exponential Distribution

Recently some work has been done on two-parameter exponential distribution, *i.e.* when both the location and scale parameters are present, *i.e.* the individual item has the PDF

$$f(x;\theta,\mu) = \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\mu)}; \quad \theta > 0, \quad x > \mu.$$

It is observed that for $n \geq 2$, the MLEs of both θ and μ exist, and they can be seen easily seen as

$$\widehat{\mu} = x_{1:n},$$

and the MLE of θ is same as before and it can be obtained by replacing $x_{i:n}$ with $x_{i:n} - x_{1:n}$.

Difficult Part

Although, the MLEs of θ and μ can be obtained in explicit form, it is not easy to obtain the joint PDF of $\widehat{\mu}$ and $\widehat{\theta}$ when they exist. The joint moment generating function of $\widehat{\theta}$ to $\widehat{\mu}$ is possible to obtain, but from their the inversion has not yet been possible.

The PDFs of $\widehat{\theta}$ and $\widehat{\mu}$ can be obtained along the same line. From the marginal PDFs, the corresponding confidence intervals can be obtained.

Bayesian Inference

Prior:
In this case it is assumed that $\lambda = 1/\theta$ has a Gamma distribution and μ has a non-informative prior.

Based on these priors, the posterior distribution of λ given μ is also a gamma distribution. The posterior distribution of μ is

$$l(\mu|Data) \propto \frac{1}{(A_0 - \mu)^{n+d}}; \quad \mu < x_{1:n}.$$

Therefore, it is possible to generate samples from the joint posterior distribution function of λ and μ, and that can be used to compute Bayes estimate and also to construct the credible intervals.

Weibull Distribution

Estimation and the associated inferential procedures are quite different when the lifetime distribution of the individual item is a Weibull random variables. Let us assume that the shape and scale parameters as α and λ respectively, *i.e.* it has the following PDF;

$$f(x;\alpha,\lambda) = \frac{\alpha}{\lambda}\left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-(x/\lambda)^{\alpha}}, \quad x > 0,$$

where $\alpha > 0$ and $\lambda > 0$ are the natural parameter space.

Available Data

Remember we have one of the following two types of observations;

Case I: $\{x_{1:n} < \cdots < x_{r:n}\}$ if $x_{r:n} \leq T$

Case II: $\{x_{1:n} < \cdots < x_{d:n}\}$ if $x_{r:n} > T$,

here d denotes the number of observed failures that occur before time point T.

Likelihood

Based on the available sample as described in the beginning of this section, the likelihood function for Case I and Case II are

$$l(\alpha,\lambda) = \left(\frac{\alpha}{\lambda}\right)^r \prod_{i=1}^{r} \left(\frac{x_{i:n}}{\lambda}\right)^{\alpha-1} e^{-\left\{\sum_{i=1}^{r}(x_{i:n}/\lambda)^\alpha + (n-r)(x_{r:n}/\lambda)^\alpha\right\}}$$

and

$$l(\alpha,\lambda) = \begin{cases} \left(\frac{\alpha}{\lambda}\right)^d \prod_{i=1}^{d} \left(\frac{x_{i:n}}{\lambda}\right)^{\alpha-1} e^{-\left\{\sum_{i=1}^{d}(x_{i:n}/\lambda)^\alpha + (n-r)(T/\lambda)^\alpha\right\}} & \text{if} \quad d > 0 \\ e^{-n(T/\lambda)^\alpha} & \text{if} \quad d = 0 \end{cases}$$

respectively. The MLEs can be obtained by maximizing the log-likelihood function with respect to the unknown parameters α and λ.

MLE of the Unknown Parameters

As expected the MLEs of α and λ cannot be obtained in explicit form. It has been observed that the MLE of α can be obtained by finding the unique solution of a fixed point type equation

$$h(\alpha) = \alpha,$$

where for Case I

$$h(\alpha) = -r\left(\sum_{i=1}^{r} \ln x_{i:n} + \frac{1}{u(\alpha)}\left[\sum_{i=1}^{r} x_{i:n}^{\alpha} \ln x_{i:n} + (n-r)x_{r:n}^{\alpha} \ln x_{r:n}\right]\right)^{-1}$$

and for Case II

$$h(\alpha) = -d\left(\sum_{i=1}^{d} \ln x_{i:n} + \frac{1}{u(\alpha)}\left[\sum_{i=1}^{d} x_{i:n}^{\alpha} \ln x_{i:n} + (n-d)\,T^{\alpha} \ln T\right]\right)^{-1}$$

MLE: Contd.

$$u(\alpha) = \begin{cases} r\left(\sum_{i=1}^{r} x_{i:n}^{\alpha} + (n-r)x_{r:n}^{\alpha}\right)^{-1} & \text{for} \quad \text{Case I} \\ d\left(\sum_{i=1}^{d} x_{i:n}^{\alpha} + (n-d)T^{\alpha}\right)^{-1} & \text{for} \quad \text{Case II} \end{cases}$$

Simple iterative scheme can be used to solve for $\widehat{\alpha}$, the MLE of α, from the above equation. Start with some initial guess of α, say $\alpha^{(0)}$, obtain $\alpha^{(1)} = h(\alpha^{(0)})$, and proceeding in this manner we can obtain $\alpha^{(k+1)} = h(\alpha^{(k)})$. Stop the iteration procedure, when $|\alpha^{(k+1)} - \alpha^{(k)}| < \epsilon$, some pre-assigned tolerance level. Once $\widehat{\alpha}$ is obtained then $\widehat{\lambda}$, the MLE of λ, can be obtained as $\widehat{\lambda} = u(\widehat{\alpha})^{1/\widehat{\alpha}}$.

Approximate MLEs

To avoid the numerical computation, approximate maximum likelihood estimates have been proposed, which have explicit expressions. It has been used in many cases particularly, when the family is a location scale family.

Although, Weibull family is not a location scale family, but log Weibull family is a location scale family..

Suppose a random variable X has Weibull distribution with PDF as given before, then $Y = \ln X$ has the extreme value distribution with PDF

$$f_Y(y;\mu,\sigma) = \frac{1}{\sigma} e^{(y-\mu)/\sigma - e^{(y-\mu)/\sigma}}; \quad -\infty < y < \infty,$$

here $\mu = \ln \lambda$, and $\sigma = 1/\alpha$.

AMLEs

The main idea is to work with $Y = \ln X$, when X has a Weibull distribution. Make a Taylor series approximation to the two normal equations upto first order terms, and obtain the estimates in explicit forms.

Bayesian Inference

This problem has a nice Bayesian solution also. For the Bayesian inference the following Weibull PDF has been considered:

$$f(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}}; \quad x > 0.$$

Prior:
It is assumed that λ has a gamma prior, and α has a PDF which is log-concave, and they are independent.

Posterior:
In this case the posterior distribution of λ given α is gamma, and the posterior distribution α has a log-concave PDF.

Bayes Estimates and Credible Intervals

It is possible to generate samples from the joint posterior density function of α and λ. Using the generated samples, the Bayes estimates and the associated credible intervals can be obtained.

Outline

Type-II HCS

Like conventional Type-I censoring, the main disadvantage of Type-I HCS is that most of the inference results are obtained under the condition that the number of observed failures is at least one, and more over there may be very few failures occurring up to the pre-fixed time T. In that case the efficiency of the estimator(s) may be very low.

Because of this alternative hybrid censoring scheme that would terminate the experiment at the random time $T^* = \max\{X_{r,n}, T\}$ has been proposed. It is called Type-II HCS.

It has the advantage of guaranteeing that at least r failures are observed at the end of the experiment. If the r failure occurs before time T, the experiment continues up to time point T. On the other hand, if the r-th failure does not occur before time T, then the experiment continues until r-th failure takes place.

Comparison

Type-I HCS: In this case the termination time is pre-fixed, which is clearly an advantage. However, if the mean lifetime of the experimental item is not small compared to the pre-fixed termination time T, then with high probability, far fewer than r failures may be observed before the termination T. This will definitely has an adverse effect on the efficiency of the inferential procedure based on Type-I HCS.

Type-II HCS: In this case, the termination time is a random variable, which is clearly a disadvantage. On the other hand for Type-II HCS, more than r failures may be observed at the termination time, and this will result in efficient estimation procedure in this case.

Outline

Generalized Type-I HCS

Suppose n identical items are put on a life test at the time point 0. Fix $r, k \in \{1, 2, \cdots, n\}$ and $T \in (0, \infty)$, such that $k < r < n$.

If the k -th failure occurs before time T, terminate the experiment at $\min\{X_{r:n}, T\}$. If the k-th failure occurs after time T, terminate the experiment at $X_{k:n}$.

It is clear that this HCS modifies the Type-I HCS by allowing the experiment to continue beyond time T if very few failures had been observed up to time point T. Under this censoring scheme, the experimenter would like to observe r failures, but is willing to accept a bare minimum of k failures.

Generalized Type-II HCS

Consider a life-testing experiment in which n items are put on a test. Fix $r \in \{1, 2, \cdots, n\}$, and $T_1, T_2 \in (0, \infty)$, where $T_1 < T_2$.

If the r-th failure occurs before the time point T_1, terminate the experiment at T_1. If the r-th failure occurs between T_1 and T_2, terminate the experiment at $X_{r:n}$. Otherwise, terminate the experiment at T_2.

This hybrid censoring scheme modifies the Type-II HCS by guaranteeing that the experiment will be completed by time T_2. Therefore, T_2 represents the absolute longest that the experimenter allows the experiment to continue.

Thank You