Exact Non-parametric Confidence Intervals for Quantiles with Progressive Type-II Censoring

Olivier Guilbaud
Clinical Science
AstraZeneca
S-15185 Södertälje
Sweden
olivier.guilbaud@astrazeneca.com

Abstract

It will be shown in this talk how certain well-known exact non-parametric inferences based on order statistics can be generalized to situations with progressive type-II censoring. These inferences include confidence intervals for a quantile, prediction intervals for a future order statistic, tolerance intervals, and outer confidence intervals for a quantile interval.

1. Exact inferences based on order statistics when there is no censoring

Let $X_1 < \cdots < X_n$ be the order statistics of a random sample from an *X*-distribution with continuous distribution function $F(t) = \Pr[X \le t]$. As is well known (David 1981), intervals of the form $[X_r, X_s]$ can be used for various exact non-parametric inferences, including:

(a) confidence intervals for a *p*-quantile ξ_p of the *X*-distribution, because

$$\Pr[X_r \le \xi_p \le X_s] \text{ is known;}$$

(b) prediction intervals for an order statistic X'_t in a future sample from the *X*-distribution, because

$$\Pr[X_r \le X'_t \le X_s] \text{ is known;}$$

(c) tolerance intervals that contains at least a given proportion γ of the *X*-distribution, because

$$\Pr[F(X_s) - F(X_r) \ge \gamma] \text{ is known;}$$

(d) outer confidence intervals for a quantile interval $[\xi_p, \xi_q]$ of the *X*-distribution, because

$$\Pr[X_r \le \xi_p \le \xi_q \le X_s] \text{ is known.}$$

These inferences are exact and nonparametric in that the relevant probabilities are known exactly without any assumption about the shape of the distribution function *F* – the only requirement is that *F* is continuous.

It will be shown in this talk how these and related inferences can be generalized to situations with progressive type-II censoring.

2. Progressive type-II censoring

A typical situation is as follows. A group of *n* experimental units is put on test at time zero in a life-testing experiment. Here *X* represents the time to failure (or death) of experimental units, and $X_1 < \cdots < X_n$ correspond to the successive times to failure of the *n* units. In the present context these order statistics are typically not all observable because of the following progressive type-II censoring.

Immediately after observing the first failure at time Y_1 $(\equiv X_1)$, a prespecified number $R_1 \geq 0$ of the $n-1$ remaining units are selected at random and removed, so that only $n-1-R_1$ units then remain under observation in the experiment. Immediately after observing the second failure at time Y_2, a prespecified number $R_2 \geq 0$ of the $n-1-R_1-1$ remaining units are selected at random and removed, so that only $n-1-R_1-1-R_2$ units then remain under observation in the experiment, and so on using $m \leq n$ prespecified numbers $R_i \geq 0$. Removed units thus become right censored at the time of failure of other units. The *m* numbers R_i specify the censoring scheme. This progressive censoring leads to *m* observable order statistics $Y_1 < \cdots < Y_m$ that are available for inferences.

The situation with no censoring corresponds to the special case with $m = n$ and $R_1 = \cdots = R_m = 0$, whereas the situation with ordinary type-II right censoring at a given order statistic corresponds to the special case with $m < n$, $R_1 = \cdots = R_{m-1} = 0$ and $R_m = n - m$.

Aggarwalla (2001) gave a nice review of the background and developments in this area. Inferences based on $Y_1 < \cdots < Y_m$ have typically been developed assuming a certain parametric model for the *X*-distribution. Models used in this context include exponential, Weibull and log-normal models. Certain exact conditional inferences under such parametric assumptions are available (Viveros and Balakrishnan 1994), including confidence intervals for a *p*-quantile ξ_p, and prediction intervals for a given order statistic X'_t in a future sample.

3. Exact inferences based on order statistics with progressive type-II censoring

It was recently shown how certain exact nonparametric inferences based on intervals of the form $[Y_r, Y_s]$, as well as related inferences, can be made. More precisely, it was shown that probabilities of the form,

$$\Pr[Y_r \leq \xi_p \leq Y_s],$$

$$\Pr[Y_r \leq Y'_t \leq Y_s],$$

can easily be evaluated using certain mixture representations (Guilbaud 2001, theorems 1 & 2) of Y_is in terms of underlying X_js. Thus, exact non-parametric confidence intervals for a *p*-quantile, and prediction intervals for a given order statistic in a future sample (possibly also subject to progressive type-II censoring), are now available. These non-parametric inferences are valid for any continuous distribution function *F*. Of course, corresponding inferences based on the assumption that a certain parametric model holds may be more efficient under that model, but the crucial question then is whether the model assumption is correct.

These results will be considered in this talk.

Although not dealt with in Guilbaud (2001), tolerance intervals and outer confidence intervals for quantile intervals, i.e. the two last kinds of inferences mentioned in section 1, can also be generalized to situations with progressive type-II censoring. That is, probabilities of the form

$$\Pr\left[F(Y_s) - F(Y_r) \geq \gamma\right],$$

$$\Pr\left[Y_r \leq \xi_p \leq \xi_q \leq Y_s\right],$$

can also be evaluated rather easily. A slightly extended variant of the mixture representations in Guilbaud (2001) has however to be used.

These additional results will also be considered in this talk.

References

Aggarwala, R. (2001). Progressive Censoring: A Review. In N. Balakrishnan and C. R. Rao (Eds.), *Handbook of Statistics, Vol. 20*, pp. 373-429. New York: Elsevier.

David, H. A. (1981). *Order Statistics*, 2nd edn. New York: Wiley.

Guilbaud, O. (2001). Exact non-parametric confidence intervals for quantiles with progressive type-II censoring. *Scandinavian Journal of Statistics 28*, 699-713.

Viveros, R. & Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics 36*, 84-91.