

Exact Inference for Competing Risks Model with Generalized Type II Progressive Hybrid Censoring

Author and Position

❑ **Subin Cho** : Graduate student, Department of Statistics, Daegu University

❑ **Seonghee Park** : Undergraduate student, Division of Mathematics and Big Data Science, Daegu University

❑ **Kyeongjun Lee** : Assistant professor, Division of Mathematics and Big Data Science, Daegu University

Text

1. Introduction

❖ In a life test that observes an item until failure, the lifespan may not be accurately recorded, and items can be removed before failure to reduce the cost or time associated with the test. In addition, more than one risk factor(RisF) can exist at the same time, and Cox proposed this model as competing risk model(CoRiM).

❖ Lone, S.A. and Rahman, A. (2017) discussed a phased partial life test plan for competitive risk using adaptive type I progressive hybrid screening (AdIPHCS). Experimental If you want to remove live experimental units at a point other than the final end of the experiment, you cannot use the traditional type I and type II screening schemes. Therefore, recently Sobhi, M.M.A. et al. (2016-2020) proposed Pr2CS and AdIPHCS are very popular in life test problems and reliability analysis. However, those assure a pre-assigned number of failures, it has the drawback that it might take a long time to observe a pre-assigned number of failures. Therefore Cho et al. (2015) suggest a generalized type II progressive hybrid censoring scheme (GenIIPrHyCS), a schematic representation of the is presented GenIIPrHyCS.

❖ We describe GenIIPrHyCS, consider a life-testing experiment in which n identical units are put on test. The times T_1, T_2 , and integer m are pre-assigned such that $m \leq n$ and $0 < T_1 < T_2 < \infty$, and also pre-assigned PrCS (R_1, R_2, \dots, R_m) are satisfied $\sum_{i=1}^m (R_i + m) = n$. Let D_1 and D_2 represent the number of failures up to pre-assigned times T_1 and T_2 , respectively. Likewise, let d_1 and d_2 be the observed value of D_1 and D_2 , respectively. When the first failure is observed, the R_1 survival units are removed randomly from the test. Furthermore, when the second failure is observed, the R_2 survival units are removed randomly from the test and so on.

❖ If $X_{m:m:n} < T_1$, terminate the test at pre-assigned time T_1 (Case I). If $T_1 < X_{m:m:n} < T_2$, then instead of terminating the test by removing the all survival units at pre-assigned time T_1 , continue to observe failures, without any removals, up to time m -th failure (Case II). If $T_2 < X_{m:m:n}$, terminate the test at pre-assigned time T_2 (Case III).

Case I: $\{X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}, X_{m+1:n}, \dots, X_{d_1:n}\}$, if $X_{m:m:n} < T_1, R_m = R_{m+1} = \dots = R_{d_1} = 0$,

Case II: $\{X_{1:m:n}, X_{2:m:n}, \dots, X_{d_1:m:n}, \dots, X_{m:m:n}\}$, if $T_1 < X_{m:m:n} < T_2$,

Case III: $\{X_{1:m:n}, X_{2:m:n}, \dots, X_{d_2:m:n}, \dots, X_{m:m:n}\}$, if $T_2 < X_{m:m:n}$.

2. Model

❖ $X_i = \min\{X_{i1}, X_{i2}, \dots, X_{i1}\}$, X_{ij} denotes the life-time of the i -th item under the j -th RisF with probability density function (PDF) and cumulative distribution function (CDF) such as $g_j(x) = \exp(-x/\theta_j)/\theta_j$, and $G_j(x) = 1 - \exp(-x/\theta_j)$, respectively.

❖ Also, we suppose that there are two RisFs for the failure of items. Then, it is to obtain the PDF and CDF of life-time as

$$F(x; \theta) = 1 - \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)x\right],$$

$$f(x; \theta) = \left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right) \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)x\right], \quad x > 0, \quad \theta_1 > 0, \quad \theta_2 > 0,$$

where $\theta = (\theta_1, \theta_2)$.

❖ It is well known that each failure observation is composed of failure lifetime and the cause of failure under the CoRiM. Let $X = \{x_{1:m:n}, x_{2:m:n}, \dots, x_{m:m:n}\}$ denote ordered Pr2CS data of n items, and $Z = \{z_1, z_2, \dots, z_2\}$ denote the indicator of risk cause corresponding to the ordered Pr2CS data.

❖ Here $z_i = 1, i = 1, 2, \dots, m$, denotes the failure of the i -th unit caused by first RisF. On the other hands, $z_i = 0$ denotes that second RisF is responsible for the i -th failure. Based on above assumption, the joint PDF of life-time and corresponding factor (X, Z) is given by

$$f_{X,Z}(x, j) = \left(\frac{1}{\theta_j}\right) \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)x\right], \quad j = 1, 2$$

❖ From GenIIPrHyCS data, therefore, we have the following data $(x_{1:m:n}, z_1), (x_{2:m:n}, z_2), \dots, (x_{u:m:n}, z_u)$, where $u = D_1$ for Case I, $u = m$ for Cases II, $u = D_2$ for Case III.

❖ Based on the three scenarios as discussed above, the likelihood function is created and the MLE of θ_j is obtained as

$$\hat{\theta}_j = \frac{1}{n_j} [\sum_{j=1}^u (1 + R_j) x_{j:m:n} + W], \quad (1)$$

where $W = R'_1 T_1$ for case I, $W = 0$ for case II, $W = R'_2 T_2$ for case III. Here, we denote the total failure number of units due to the RisF j by $n_j, j = 1, 2$, then it is easy to obtain $n_1 = \sum_{i=1}^u z_i$ and $n_2 = \sum_{i=1}^u (1 - z_i) = u - n_1$.

❖ Noting from (1), the MLEs do not exist when $n_j = 0, j = 1, 2$. In order to estimate θ_j , we have to observe at least one failure caused by each RisF. That is

$$\xi^{(u)} = \{n_1 \geq 1, n_2 \geq 2, n_1 + n_2 = u\}.$$

3. Exact conditional inference for MLE

❖ The following Lemma 1 established due to Cho and Lee (2017) is used to derive the explicit expression of the Conditional MGF(ConMGF) of MLE.

❖ **Lemma 1.** Let $v_j > 0$ where $j = 1, 2, \dots, m$, and let X denote the absolutely continuous random variable with PDF $f(x)$ and CDF $F(x)$. Then for $m \geq 1$, we have

$$\int_{-\infty}^{x_{m+1}} \dots \int_{-\infty}^{x_3} \int_{-\infty}^{x_2} \prod_{j=1}^m f(x_j) \{1 - F(x_j)\}^{v_j-1} dx_1 dx_2 \dots dx_m = \sum_{i=0}^m \zeta_{i,m}(v_m) \{1 - F(x_{m+1})\}^{\varphi_{i,m}(v_m)},$$

where $v_m = (v_1, v_2, \dots, v_m)$; $\zeta_{i,m}(v_m) = \frac{(-1)^i}{\{\prod_{j=1}^m \sum_{k=m-i+1}^{m-i+j} v_k\} \{\prod_{j=1}^{m-i} \sum_{k=j}^{m-i} v_k\}}$, $\varphi_{i,m}(v_m) = \sum_{k=m-i+1}^m v_j$ with the usual conventions that $\prod_{j=1}^0 \xi_j \equiv 1$ and $\sum_{j=1}^0 \xi_j \equiv 0$.

❖ **Theorem 1.** Conditional on $\xi^{(u)}$, the ConMGF of $\hat{\theta}_1$ is given by

$$M_{\hat{\theta}_1}(t) = E\left(e^{t\hat{\theta}_1} | \xi^{(u)}\right)$$

$$= \sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(\xi^{(d_1)} | D_1 = d_1)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^m} \left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)^{-d_1} \sum_{j=0}^{d_1} \zeta_{j,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \times q_1^{\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) R'_{d_1-i+1}}$$

$$+ \sum_{d_1=0}^{m-1} \sum_{i=1}^{m-1} \frac{\zeta'_m}{P(\xi^{(m)} | D_1 = d_1, D_2 = m)} \binom{m}{i} \frac{\theta_1^{m-i} \theta_2^i}{(\theta_1 + \theta_2)^m} \left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)^{-m} \sum_{i_1=0}^{d_1} \sum_{i_2=0}^{m-d_1} \zeta_{i_1,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \zeta_{i_2,d_2}$$

$$\times (1_{m-d_1-1}, R_m + 1) q_1^{\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) \sum_{j=d_1-i+1}^{m-i_2} (R_j + 1)} \frac{1}{q_2} \sum_{j=m-i_2+1}^m (R_j + 1) + \sum_{d_2=1}^{m-1} \sum_{i=1}^{d_2-1} \frac{\zeta'_{d_2}}{P(\xi^{(d_2)} | D_2 = d_2)}$$

$$\times \binom{d_2}{i} \frac{\theta_1^{d_2-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_2}} \left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)^{-d_2} \sum_{j=0}^{d_2} \zeta_{j,d_2}(R_1 + 1, \dots, R_{d_1} + 1, 1_{d_2-d_1}, R'_{d_2} + 1) q_2^{\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) R'_{d_2-i+1}}.$$

where $q_1 = \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)T_1\right]$ and $q_2 = \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)T_2\right]$.

❖ **Proof.** Conditional on $\xi^{(u)}$, the MGF of $\hat{\theta}_1$ is given by

$$M_{\hat{\theta}_1}(t) = E\left(e^{t\hat{\theta}_1} | \xi^{(u)}\right)$$

$$= E\left(e^{t\hat{\theta}_1} | D_1 = d_1, \xi^{(d_1)}\right) P(D_1 = d_1) + \sum_{d_2=0}^{m-1} E\left(e^{t\hat{\theta}_1} | D_1 = d_1, D_2 = m, \xi^{(m)}\right) P(D_1 = d_1, D_2 = m) \quad (2)$$

$$+ \sum_{d_2=1}^{m-1} E\left(e^{t\hat{\theta}_1} | D_2 = d_2, \xi^{(d_2)}\right) P(D_2 = d_2).$$

For convenience, let us denote the subset of indicator of failure causes as Q_u^* , where

$$Q_u^* = \{Z = (z_1, \dots, z_u): z_i = 0 \text{ or } 1; i = 1, \dots, u\}.$$

For Case I ($D_1 = d_1$): Conditional on $D_1 = d_1, n_1 = i$, the joint distribution of order statistics $x_{1:m:n} < \dots$

$< x_{m:m:n} < T_1$ has the form

$$f(x_{1:m:n}, \dots, x_{m:m:n} | D_1 = d_1, n_1 = i)$$

$$= \frac{\zeta'_{d_1}}{P(D_1 = d_1, n_1 = i)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)^{d_1} \exp\left[-\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right) \sum_{j=1}^{d_1} (1 + R_j) x_{j:m:n}\right].$$

Upon the conditional PDF obtained above, we can readily have

$$E\left(e^{t\hat{\theta}_1} | D_1 = d_1, \xi^{(d_1)}\right)$$

$$= \sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(D_1 = d_1, n_1 = i)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \int_0^{T_1} \dots \int_0^{x_{2:m:n}} \prod_{j=1}^{d_1} f(x_{j:m:n}) [1 - F(x_j)]^{v_j-1} q_1^{\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) R'_{d_1-j+1}} dx_{1:m:n} \dots dx_{d_1:m:n}.$$

From **Lemma 1** with $v_j = (1 + R_j) \left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)$ and then factor $\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)$ out of all of the v_j 's, the above expression can be easily simplified as

$$\sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(D_1 = d_1, n_1 = i)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right)^{-d_1} \sum_{j=0}^{d_1} \zeta_{j,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \times q_1^{\left(1 - \frac{t}{i} \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) R'_{d_1-j+1}}. \quad (3)$$

❖ **Corollary 1.** The first moment of $\hat{\theta}_1$ is given by

$$E_{\theta_1}(\hat{\theta}_1) = M'_{\hat{\theta}_1}(0)$$

$$= \sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(\xi^{(d_1)} | D_1 = d_1)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \sum_{j=0}^{d_1} \zeta_{j,d_1}(R_1 + 1, \dots, R_{d_1} + 1) q_1^{R'_{d_1-j+1}} \frac{1}{i} \left[d_1 \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} + T_1 R'_{d_1-j+1}\right]$$

$$+ \sum_{d_1=0}^{m-1} \sum_{i=1}^{m-1} \frac{\zeta'_m}{P(\xi^{(m)} | D_1 = d_1, D_2 = m)} \binom{m}{i} \frac{\theta_1^{m-i} \theta_2^i}{(\theta_1 + \theta_2)^m} \sum_{i_1=0}^{d_1} \sum_{i_2=0}^{m-d_1} \zeta_{i_1,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \zeta_{i_2,d_2}(1_{m-d_1-1}, R_m + 1)$$

$$\times q_1^{\sum_{j=d_1-i+1}^{m-i_2} (R_j + 1)} \frac{1}{q_2} \sum_{j=m-i_2+1}^m (R_j + 1) + \frac{1}{i} \left[m \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} + T_1 \sum_{j=d_1-i+1}^{m-i_2} (R_j + 1) + T_2 \sum_{j=m-i_2+1}^m (R_j + 1)\right] +$$

$$\sum_{d_2=1}^{m-1} \sum_{i=1}^{d_2-1} \frac{\zeta'_{d_2}}{P(\xi^{(d_2)} | D_2 = d_2)} \binom{d_2}{i} \frac{\theta_1^{d_2-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_2}} \sum_{j=0}^{d_2} \zeta_{j,d_2}(R_1 + 1, \dots, R_{d_1} + 1, 1_{d_2-d_1}, R'_{d_2} + 1) q_2^{R'_{d_2-j+1}}$$

$$\times \frac{1}{i} \left[d_2 \frac{\theta_1 \theta_2}{\theta_1 + \theta_2} + T_2 R'_{d_2-j+1}\right].$$

❖ **Theorem 2.** Conditional on $\xi^{(u)}$, the conditional PDF of $\hat{\theta}_1$ is given by

$$f_{\hat{\theta}_1}(x)$$

$$= \sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(\xi^{(d_1)} | D_1 = d_1)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \sum_{j=0}^{d_1} \zeta_{j,d_1}(R_1 + 1, \dots, R_{d_1} + 1) q_1^{R'_{d_1-j+1}} \gamma\left(x - \frac{T_1 R'_{d_1-j+1}}{i}; d_1, \frac{\theta_1 \theta_2}{i(\theta_1 + \theta_2)}\right)$$

$$+ \sum_{d_1=0}^{m-1} \sum_{i=1}^{m-1} \frac{\zeta'_m}{P(\xi^{(m)} | D_1 = d_1, D_2 = m)} \binom{m}{i} \frac{\theta_1^{m-i} \theta_2^i}{(\theta_1 + \theta_2)^m} \sum_{i_1=0}^{d_1} \sum_{i_2=0}^{m-d_1} \zeta_{i_1,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \zeta_{i_2,d_2}(1_{m-d_1-1}, R_m + 1)$$

$$\times q_1^{\sum_{j=d_1-i+1}^{m-i_2} (R_j + 1)} \frac{1}{q_2} \sum_{j=m-i_2+1}^m (R_j + 1) \gamma\left(x - \frac{1}{i} \left[T_1 \sum_{j=d_1-i+1}^{m-i_2} (R_j + 1) + T_2 \sum_{j=m-i_2+1}^m (R_j + 1)\right]; m, \frac{\theta_1 \theta_2}{i(\theta_1 + \theta_2)}\right)$$

$$+ \sum_{d_2=1}^{m-1} \sum_{i=1}^{d_2-1} \frac{\zeta'_{d_2}}{P(\xi^{(d_2)} | D_2 = d_2)} \binom{d_2}{i} \frac{\theta_1^{d_2-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_2}} \sum_{j=0}^{d_2} \zeta_{j,d_2}(R_1 + 1, \dots, R_{d_1} + 1, 1_{d_2-d_1}, R'_{d_2} + 1) q_2^{R'_{d_2-j+1}}$$

$$\times \gamma\left(x - \frac{T_2 R'_{d_2-j+1}}{i}; d_2, \frac{\theta_1 \theta_2}{i(\theta_1 + \theta_2)}\right),$$

where $\gamma(x - c; a, b)$ denote gamma distribution with shape parameter a, rate parameter b and shift parameter c.

❖ **Corollary 2.** Conditional on $\xi^{(u)}$, the tail probability of PDF of $\hat{\theta}_1$ can be expressed as

$$P_{\hat{\theta}_1}(\hat{\theta}_1 > k)$$

$$= \sum_{i=1}^{d_1-1} \frac{\zeta'_{d_1}}{P(\xi^{(d_1)} | D_1 = d_1)} \binom{d_1}{i} \frac{\theta_1^{d_1-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_1}} \sum_{j=0}^{d_1} \zeta_{j,d_1}(R_1 + 1, \dots, R_{d_1} + 1) q_1^{R'_{d_1-j+1}} \Gamma\left(d_1, \frac{\theta_1 \theta_2}{(d_1-i)(\theta_1 + \theta_2)} \left(k - \frac{T_1 R'_{d_1-j+1}}{d_1-i}\right)\right)$$

$$+ \sum_{d_1=0}^{m-1} \sum_{i=1}^{m-1} \frac{\zeta'_m}{P(\xi^{(m)} | D_1 = d_1, D_2 = m)} \binom{m}{i} \frac{\theta_1^{m-i} \theta_2^i}{(\theta_1 + \theta_2)^m} \sum_{i_1=0}^{d_1} \sum_{i_2=0}^{m-d_1} \zeta_{i_1,d_1}(R_1 + 1, \dots, R_{d_1} + 1) \zeta_{i_2,d_2}(1_{m-d_1-1}, R_m + 1)$$

$$\times q_1^{\sum_{j=d_1-i+1}^{m-i_2} (R_j + 1)} \frac{1}{q_2} \sum_{j=m-i_2+1}^m (R_j + 1) \Gamma\left(m, \frac{\theta_1 \theta_2}{(m-i)(\theta_1 + \theta_2)} \left(k - \frac{1}{m-i} \left[T_1 \sum_{j=d_1-i+1}^{m-i_2} (R_j + 1) + T_2 \sum_{j=m-i_2+1}^m (R_j + 1)\right]\right)\right)$$

$$+ \sum_{d_2=1}^{m-1} \sum_{i=1}^{d_2-1} \frac{\zeta'_{d_2}}{P(\xi^{(d_2)} | D_2 = d_2)} \binom{d_2}{i} \frac{\theta_1^{d_2-i} \theta_2^i}{(\theta_1 + \theta_2)^{d_2}} \sum_{j=0}^{d_2} \zeta_{j,d_2}(R_1 + 1, \dots, R_{d_1} + 1, 1_{d_2-d_1}, R'_{d_2} + 1) q_2^{R'_{d_2-j+1}}$$

$$\times \Gamma\left(d_2, \frac{\theta_1 \theta_2}{(d_2-i)(\theta_1 + \theta_2)} \left(k - \frac{T_2 R'_{d_2-j+1}}{d_2-i}\right)\right).$$

where k is an arbitrary constant, $\langle x \rangle = \max\{x, 0\}$ and $\Gamma(a, b) = \int_b^\infty (1/(a-1)!) x^{a-1} e^{-x} dx$.

4. Data Analysis

❖ In order to analyzed the real data, we use the estimators in the above section. The real data were from some small electronic appliances exposed to the automatic test machine by Lawless, J (2011). This data was analyzed by References Mao, S et al. (2014) and Cho, Y and Lee, K (2017).

❖ Table 1 presents the 95% CIs for $\hat{\theta}_1$ and $\hat{\theta}_2$, and we have contained the standard error (SE) and MSE calculated from Corollary 1. Also, the PDFs of $\hat{\theta}_1$ and $\hat{\theta}_2$ based on the example data is shown in Figure 1.

<Table 1. Inference of parameters for example>

T_1	T_2	n_1	n_2	$\hat{\theta}_1$	SE($\hat{\theta}_1$)	95% C.I	$\hat{\theta}_2$	SE($\hat{\theta}_2$)	95% C.I
7000	8000	12	16	7144.417	2062.415	(4057.341, 12580.330)	5358.312	1339.578	(3282.644, 8746.460)
3000	7000	12	16	8294.250	2394.344	(4710.336, 14605.030)	6220.688	1555.172	(3810.957, 10154.130)
3000	5000	11	16	7970.455	2403.182	(4057.341, 14392.450)	5479.688	1369.922	(3357.001, 8944.583)

< Figure 1. The conditional PDFs of $\hat{\theta}_1$ and $\hat{\theta}_2$ >

