

文章编号:1000-8055(2011)02-0370-06

同条件多组定数截尾试验整体数据融合方法

林逢春, 康 锐

(北京航空航天大学 可靠性与系统工程学院, 北京 100191)

摘 要: 针对正态分布、对数正态分布、极值分布和 Weibull 分布等位置-尺度分布, 提出一种同条件多组定数截尾试验的整体数据融合方法, 能够对相同试验条件下的多组定数截尾试验数据进行综合分析, 给出位置参数和尺度参数的整体融合估计, 并在此基础上建立百分位值的点估计和区间估计. 与传统方法相比, 该方法具有信息量大、精度高的特点, 能够进行极小子样的可靠性评定, 解决了一个组别只有一个失效数据时产品可靠性的评定难题.

关 键 词: 可靠性; 定数截尾试验; 数据融合; 整体估计; 小子样

中图分类号: TB114.3

文献标识码: A

Integral data fusion method for multigroup type-II censored tests under the same condition

LIN Feng-chun, KANG Rui

(School of Reliability and Systems Engineering,

Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: An integral data fusion method for multigroup type-II censored tests under the same condition was presented for normal distribution, lognormal distribution, extreme value distribution, Weibull distribution and other location-scale distributions. This method could integrally infer censored data in multiple groups under the same conditions, and give the integral fusion estimate of the location and the scale parameter. And then, the estimator and the confidence interval of population percentile were established. Compared with traditional methods, the presented method not only has higher precision but also solves the problem of reliability assessment with very small sample.

Key words: reliability; type-II censored test; data fusion; integral estimate; small sample

在工程实际中, 常常会遇到同一试验条件下先后进行多组截尾试验的情况, 此时需要根据多组截尾数据对该试验条件下产品的可靠性进行综合评估. 由于试验条件相同, 可将不同组别下的试验数据看作来自同一母体的多个独立样本. 对于指数分布的情况, 通过计算总试验时间, 可以直接给出母体均值的参数估计. 对于正态分布、对数正态分布、极值分布和 Weibull 分布等位置-尺度分布的情况, 一般采用基于线性估计的两步融合方

法^[1-3]. 两步融合方法首先对各个组别下的截尾数据分别进行分析, 得到各个组别下分布参数的线性无偏估计, 然后采用加权估计方法得到母体位置参数和尺度参数的融合估计. 由于两步融合方法在数据融合过程中会丢失一些信息, 所以存在误差积累和参数估计精度不高的问题, 当试样样本量和截尾数较小时尤为显著. 特别地, 当部分组别下只有一个失效数据时, 两步融合方法根本无法对其进行分析处理.

收稿日期: 2009-12-30; 修订日期: 2010-05-03

作者简介: 林逢春(1984-), 女, 山东临沂人, 博士后, 主要从事小样本信息技术、寿命数据分析和可靠性统计研究.

文献[4-5]针对位置-尺度分布提出一种截尾数据回归分析方法,能够对不同试验条件下的截尾数据进行整体推断,给出回归曲线的置信限和百分位值的置信区间估计.在此基础上,本文提出一种同条件多组定数截尾试验的数据融合方法,能够将同一试验条件下的多组截尾数据作为一个整体进行统计分析,给出位置参数、尺度参数的整体融合估计,进而给出百分位值的点估计和区间估计.本文还对整体融合方法和两步融合方法进行了对比分析,分析结果表明,整体融合估计优于两步融合估计,并且能够解决部分组别下只有一个失效数据的统计分析问题.文中最后给出了一个应用实例.

1 定数截尾数据的整体融合估计

设 X 为位置-尺度分布随机变量,即

$$\begin{aligned} X &\sim \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) \\ -\infty &< x < +\infty, \\ -\infty &< \mu < +\infty, \quad \sigma > 0 \end{aligned} \quad (1)$$

其分布函数为 $F[(X-\mu)/\sigma]$, 其中 μ 和 σ 分别为位置参数和尺度参数. 令 $X^\circ = (X-\mu)/\sigma$, 则 X° 被称为标准位置-尺度分布随机变量, 具有不含任何未知参数的概率密度函数 $f(x^\circ)$ 和分布函数 $F(x^\circ)$. 正态分布、对数正态分布、两参数 Weibull 分布和极值分布等均属于位置-尺度分布.

1.1 位置参数和尺度参数的整体融合估计

设在第 i 组下取 n_i 个产品进行失效数为 r_i 的定数截尾试验, $X_{i1} \leq X_{i2} \leq \dots \leq X_{ir_i}$ 为前 r_i 个产品失效时间, 试验到 X_{ir_i} 时刻截止 ($i=1, 2, \dots, m$), 则位置参数 μ 和尺度参数 σ 的融合估计 $\hat{\mu}$ 和 $\hat{\sigma}$ 分别为

$$\hat{\mu} = \sum_{i=1}^m \sum_{j=1}^{r_i} a_{ij} X_{ij} \quad (2)$$

$$\hat{\sigma} = \sum_{i=1}^m \sum_{j=1}^{r_i} b_{ij} X_{ij} \quad (3)$$

式中

$$a_{ij} = \frac{c_2 \sum_{k=1}^{r_i} v^{ijk} - d \sum_{k=1}^{r_i} u_{ik} v^{ijk}}{c_1 c_2 - d^2} \quad (i=1, 2, \dots, m; j=1, 2, \dots, r_i) \quad (4)$$

$$b_{ij} = \frac{c_1 \sum_{k=1}^{r_i} u_{ik} v^{ijk} - d \sum_{k=1}^{r_i} v^{ijk}}{c_1 c_2 - d^2} \quad (i=1, 2, \dots, m; j=1, 2, \dots, r_i) \quad (5)$$

$$c_1 = \sum_{i=1}^m \sum_{j,k=1}^{r_i} v^{ijk} \quad (6)$$

$$c_2 = \sum_{i=1}^m \sum_{j,k=1}^{r_i} u_{ij} u_{ik} v^{ijk} \quad (7)$$

$$d = \sum_{i=1}^m \sum_{j,k=1}^{r_i} u_{ij} v^{ijk} \quad (8)$$

$$(v^{ijk})_{r_i \times r_i} = (v_{ijk})_{r_i \times r_i}^{-1} \quad (i=1, 2, \dots, m) \quad (9)$$

其中 $X_{ij}^\circ = (X_{ij} - \mu)/\sigma$ ($i=1, 2, \dots, m; j=1, 2, \dots, r_i$) 为标准位置-尺度分布 $F(x^\circ)$ 的顺序统计量; $u_{ij} = E(X_{ij}^\circ)$ ($j=1, 2, \dots, r_i$) 为分布 $F(x^\circ)$ 样本量为 n_i 的第 j 个顺序统计量的均值, $v_{ijk} = \text{Cov}(X_{ij}^\circ, X_{ik}^\circ)$ ($j, k=1, 2, \dots, r_i$) 为分布 $F(x^\circ)$ 样本量为 n_i 的第 j 个与第 k 个顺序统计量的协方差 ($i=1, 2, \dots, m$), 它们仅与 n_i 和 $F(x^\circ)$ 有关, 而与 μ 和 σ 无关, 均可查表^[6-9]或通过专门程序计算得到^[10].

可以证明, 式(2)和(3)给出的 $\hat{\mu}$ 和 $\hat{\sigma}$ 是 μ 和 σ 的最佳线性无偏估计, 其方差和协方差分别为

$$D(\hat{\mu}) = \frac{c_2}{c_1 c_2 - d^2} \sigma^2 \quad (10)$$

$$D(\hat{\sigma}) = \frac{c_1}{c_1 c_2 - d^2} \sigma^2 \quad (11)$$

$$\text{Cov}(\hat{\mu}, \hat{\sigma}) = -\frac{d}{c_1 c_2 - d^2} \sigma^2 \quad (12)$$

并且 $\hat{\mu}$ 和 $\hat{\sigma}$ 渐近服从二维正态分布. 大量 Monte Carlo 模拟表明, 当数据总量 $r = \sum_{i=1}^m r_i$ 稍大时, $\hat{\mu}$ 和 $\hat{\sigma}$ 即可近似为二维正态分布.

1.2 位置参数和尺度参数的区间估计

位置参数 μ 和尺度参数 σ 置信水平为 γ 的单侧置信上、下限分别为

$$\mu_{\text{up}} = \hat{\mu} + \frac{\hat{\sigma}}{1 - u_\gamma^2 S_{22}} \times [u_\gamma^2 S_{12} + u_\gamma \sqrt{u_\gamma^2 (S_{12}^2 - S_{11} S_{22}) + S_{11}}] \quad (13)$$

$$\mu_{\text{low}} = \hat{\mu} + \frac{\hat{\sigma}}{1 - u_\gamma^2 S_{22}} \times [u_\gamma^2 S_{12} - u_\gamma \sqrt{u_\gamma^2 (S_{12}^2 - S_{11} S_{22}) + S_{11}}] \quad (14)$$

$$\sigma_{\text{up}} = \frac{1}{1 - u_\gamma \sqrt{S_{22}}} \hat{\sigma} \quad (15)$$

$$\sigma_{\text{low}} = \frac{1}{1 + u_\gamma \sqrt{S_{22}}} \hat{\sigma} \quad (16)$$

式中

$$S_{11} = \frac{D(\hat{\mu})}{\sigma^2} = \frac{c_2}{c_1 c_2 - d^2} \quad (17)$$

$$S_{22} = \frac{D(\hat{\sigma})}{\sigma^2} = \frac{c_1}{c_1 c_2 - d^2} \quad (18)$$

$$S_{12} = \frac{\text{Cov}(\hat{\mu}, \hat{\sigma})}{\hat{\sigma}^2} = -\frac{d}{c_1 c_2 - d^2} \quad (19)$$

区间 $[\mu_{\text{low}}, \mu_{\text{up}}]$ 和 $[\sigma_{\text{low}}, \sigma_{\text{up}}]$ 分别是 μ 和 σ 置信水平为 $2\gamma-1$ ($\gamma \geq 50\%$) 的置信区间。

1.3 百分位值的点估计和区间估计

母体 X 概率为 P 的百分位值 $x_P = \mu + F^{-1}(P)\sigma$ 的无偏估计 $\hat{X}_P$ 为

$$\hat{X}_P = \hat{\mu} + F^{-1}(P)\hat{\sigma} \quad (20)$$

其置信水平为 γ 的单侧置信上、下限分别为

$$X_{\text{Pup}} = \hat{\mu} + \frac{\hat{\sigma}}{1 - u_\gamma^2 S_{22}} [F^{-1}(P) + u_\gamma^2 S_{12} + k_P] \quad (21)$$

$$X_{\text{Plow}} = \hat{\mu} + \frac{\hat{\sigma}}{1 - u_\gamma^2 S_{22}} [F^{-1}(P) + u_\gamma^2 S_{12} - k_P] \quad (22)$$

区间 $[X_{\text{Plow}}, X_{\text{Pup}}]$ 是 x_P 置信水平为 $2\gamma-1$ ($\gamma \geq 50\%$) 的置信区间, 其中

$$k_P = u_\gamma \{ u_\gamma^2 (S_{12}^2 - S_{11} S_{22}) + S_{11} + [F^{-1}(P)]^2 S_{22} + 2F^{-1}(P) S_{12} \}^{1/2} \quad (23)$$

2 整体融合方法与两步融合方法的对比分析

2.1 定数截尾数据的两步融合估计

设在第 i 组下取 n_i 个产品进行失效数为 r_i 的定数截尾试验, $X_{i1} \leq X_{i2} \leq \dots \leq X_{ir_i}$ 为前 r_i 个产品失效时间, 试验到 X_{ir_i} 时刻截止 ($i = 1, 2, \dots, m$), 则位置参数 μ 和尺度参数 σ 的两步融合估计 $\tilde{\mu}$ 和 $\tilde{\sigma}$ 分别为

$$\tilde{\mu} = \left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \frac{\hat{\mu}_i}{A_{r_i, n_i}} \quad (24)$$

$$\tilde{\sigma} = \left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \frac{\hat{\sigma}_i}{l_{r_i, n_i}} \quad (25)$$

式中 $\hat{\mu}_i$ 和 $\hat{\sigma}_i$ 分别为第 i 组下位置参数 μ 和尺度参数 σ 的最佳线性无偏估计 (BLUE), A_{r_i, n_i} 和 l_{r_i, n_i} 分别是 $\hat{\mu}_i$ 和 $\hat{\sigma}_i$ 的方差系数. $\hat{\mu}_i$ 和 $\hat{\sigma}_i$ 可由式 (28) 和 (29) 给出

$$\hat{\mu}_i = \sum_{j=1}^{r_i} a_{ij}^* X_{ij} \quad (i = 1, 2, \dots, m) \quad (26)$$

$$\hat{\sigma}_i = \sum_{j=1}^{r_i} b_{ij}^* X_{ij} \quad (i = 1, 2, \dots, m) \quad (27)$$

其中 a_{ij}^* 和 b_{ij}^* 分别是位置参数 μ 和尺度参数 σ 的 BLUE 系数. a_{ij}^* 和 b_{ij}^* 以及 A_{r_i, n_i} 和 l_{r_i, n_i} 只与 n_i , r_i 和 $F(x^\circ)$ 有关, 而与 μ 和 σ 无关, 可以查表^[6]或通过专门程序计算得到^[1].

可以证明, 式 (24) 和 (25) 给出的 $\tilde{\mu}$ 和 $\tilde{\sigma}$ 是 μ

和 σ 的线性无偏估计, 其方差和协方差分别为

$$D(\tilde{\mu}) = \sigma^2 \left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \quad (28)$$

$$D(\tilde{\sigma}) = \sigma^2 \left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \quad (29)$$

$$\text{Cov}(\tilde{\mu}, \tilde{\sigma}) = \sigma^2 \left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \frac{B_{r_i, n_i}}{A_{r_i, n_i} l_{r_i, n_i}} \quad (30)$$

式中 B_{r_i, n_i} 是 $\hat{\mu}_i$ 和 $\hat{\sigma}_i$ 的协方差系数, 只与 n_i , r_i 和 $F(x^\circ)$ 有关, 而与 μ 和 σ 无关, 可以查表^[6]或通过专门程序计算得到^[1].

2.2 对比分析

传统的截尾数据分析方法在截尾数较小的情况下, 分析精度较差, 对于只有一个失效数据的情况更是无法进行分析处理. 因此, 当截尾数较小时, 两步融合方法的估计精度较低, 也无法解决部分组别只有一个失效数据的问题. 但是, 本文提出的整体融合方法能够同时对所有组别下的失效数据进行分析, 大大增加了分析时可用的信息量, 从而有效提高了融合估计的估计精度, 并且能够解决部分组别只有一个失效数据的分析难题. 对于后者, 本文在第3节给出了一个应用实例, 表明了本文方法的有效性. 本节将从估计量的优良性出发, 进行整体融合方法和两步融合方法的对比分析.

由1.1节和2.1节可知, 整体融合估计量 $\tilde{\mu}$ 和 $\tilde{\sigma}$ 以及两步融合估计量 $\hat{\mu}$ 和 $\hat{\sigma}$ 都是线性无偏的, 下面以有效性为标准评判两种估计量的优良性.

可以证明

$$D(\tilde{\mu}) = D[(\tilde{\mu} - \hat{\mu}) + \hat{\mu}] = D(\tilde{\mu} - \hat{\mu}) + 2\text{Cov}(\tilde{\mu} - \hat{\mu}, \hat{\mu}) + D(\hat{\mu}) \quad (31)$$

$$D(\tilde{\sigma}) = D[(\tilde{\sigma} - \hat{\sigma}) + \hat{\sigma}] =$$

$$D(\tilde{\sigma} - \hat{\sigma}) + 2\text{Cov}(\tilde{\sigma} - \hat{\sigma}, \hat{\sigma}) + D(\hat{\sigma}) \quad (32)$$

而 $D(\tilde{\mu} - \hat{\mu}) > 0$, $D(\tilde{\sigma} - \hat{\sigma}) > 0$, 并且

$$\text{Cov}(\tilde{\mu} - \hat{\mu}, \hat{\mu}) = 0 \quad (33)$$

$$\text{Cov}(\tilde{\sigma} - \hat{\sigma}, \hat{\sigma}) = 0 \quad (34)$$

因此

$$D(\tilde{\mu}) = D(\tilde{\mu} - \hat{\mu}) + D(\hat{\mu}) > D(\hat{\mu}) \quad (35)$$

$$D(\tilde{\sigma}) = D(\tilde{\sigma} - \hat{\sigma}) + D(\hat{\sigma}) > D(\hat{\sigma}) \quad (36)$$

即整体融合估计量 $\tilde{\mu}$ 和 $\tilde{\sigma}$ 优于两步融合估计量 $\hat{\mu}$ 和 $\hat{\sigma}$. 下面给出式 (33) 和 (34) 的证明.

证明:

$$1) \quad \text{Cov}(\tilde{\mu} - \hat{\mu}, \hat{\mu}) = \text{Cov}(\tilde{\mu}, \hat{\mu}) - D(\hat{\mu}) =$$

$$\begin{aligned} & \text{Cov} \left[\left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j=1}^{r_i} \frac{a_{ij}^*}{A_{r_i, n_i}} X_{ij}, \sum_{i=1}^m \sum_{j=1}^{r_i} a_{ij} X_{ij} \right] - \frac{c_2}{c_1 c_2 - d^2} \sigma^2 = \\ & \left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{a_{ij}^* a_{ik}}{A_{r_i, n_i}} \text{Cov}(X_{ij}, X_{ik}) - \frac{c_2}{c_1 c_2 - d^2} \sigma^2 = \\ & \left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{a_{ij}^* a_{ik}}{A_{r_i, n_i}} \sigma^2 v_{ijk} - \frac{c_2}{c_1 c_2 - d^2} \sigma^2 = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{a_{ij}^*}{A_{r_i, n_i}} \left(c_2 \sum_{l=1}^{r_i} v^{jkl} - d \sum_{l=1}^{r_i} u_{il} v^{jkl} \right) v_{ijk} - c_2 \right] = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \left(c_2 \sum_{i=1}^m \sum_{j,k,l=1}^{r_i} \frac{a_{ij}^*}{A_{r_i, n_i}} v^{jkl} v_{ijk} - d \sum_{i=1}^m \sum_{j,k,l=1}^{r_i} \frac{a_{ij}^*}{A_{r_i, n_i}} u_{il} v^{jkl} v_{ijk} \right) - c_2 \right] \end{aligned}$$

由于

$$\sum_{j,k=1}^{q_i} a_{ij}^* v_{ijk} v^{jkl} = a_{il}^*, \quad \sum_{j,k=1}^{q_i} a_{ij}^* v_{ijk} v^{jkl} u_{il} = a_{il}^* u_{il}, \quad \sum_{j=1}^{r_i} a_{ij}^* = 1, \quad \sum_{j=1}^{r_i} a_{ij}^* u_{ij} = 0$$

所以

$$\begin{aligned} \text{Cov}(\tilde{\mu} - \hat{\mu}, \hat{\mu}) &= \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \left(c_2 \sum_{i=1}^m \sum_{l=1}^{r_i} \frac{a_{il}^*}{A_{r_i, n_i}} - d \sum_{i=1}^m \sum_{l=1}^{r_i} \frac{a_{il}^* u_{il}}{A_{r_i, n_i}} \right) - c_2 \right] = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right)^{-1} \left(c_2 \sum_{i=1}^m \frac{1}{A_{r_i, n_i}} \right) - c_2 \right] = 0 \end{aligned}$$

$$2) \quad \text{Cov}(\tilde{\sigma} - \hat{\sigma}, \hat{\sigma}) = \text{Cov}(\tilde{\sigma}, \hat{\sigma}) - D(\hat{\sigma}) =$$

$$\begin{aligned} & \text{Cov} \left[\left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j=1}^{r_i} \frac{b_{ij}^*}{l_{r_i, n_i}} X_{ij}, \sum_{i=1}^m \sum_{j=1}^{r_i} b_{ij} X_{ij} \right] - \frac{c_1}{c_1 c_2 - d^2} \sigma^2 = \\ & \left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{b_{ij}^* b_{ik}}{l_{r_i, n_i}} \text{Cov}(X_{ij}, X_{ik}) - \frac{c_1}{c_1 c_2 - d^2} \sigma^2 = \\ & \left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{b_{ij}^* b_{ik}}{l_{r_i, n_i}} \sigma^2 v_{ijk} - \frac{c_1}{c_1 c_2 - d^2} \sigma^2 = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \sum_{i=1}^m \sum_{j,k=1}^{r_i} \frac{b_{ij}^*}{l_{r_i, n_i}} \left(c_1 \sum_{l=1}^{r_i} u_{il} v^{jkl} - d \sum_{l=1}^{r_i} v^{jkl} \right) v_{ijk} - c_1 \right] = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \left(c_1 \sum_{i=1}^m \sum_{j,k,l=1}^{r_i} \frac{b_{ij}^*}{l_{r_i, n_i}} u_{il} v^{jkl} v_{ijk} - d \sum_{i=1}^m \sum_{j,k,l=1}^{r_i} \frac{b_{ij}^*}{l_{r_i, n_i}} v^{jkl} v_{ijk} \right) - c_1 \right] \end{aligned}$$

由于

$$\sum_{j,k=1}^{q_i} b_{ij}^* v_{ijk} v^{jkl} = b_{il}^*, \quad \sum_{j,k=1}^{q_i} b_{ij}^* v_{ijk} v^{jkl} u_{il} = b_{il}^* u_{il}, \quad \sum_{j=1}^{r_i} b_{ij}^* = 0, \quad \sum_{j=1}^{r_i} b_{ij}^* u_{ij} = 1$$

所以

$$\begin{aligned} \text{Cov}(\tilde{\sigma} - \hat{\sigma}, \hat{\sigma}) &= \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \left(c_1 \sum_{i=1}^m \sum_{l=1}^{r_i} \frac{b_{il}^* u_{il}}{l_{r_i, n_i}} - d \sum_{i=1}^m \sum_{l=1}^{r_i} \frac{b_{il}^*}{l_{r_i, n_i}} \right) - c_1 \right] = \\ & \frac{\sigma^2}{c_1 c_2 - d^2} \left[\left(\sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right)^{-1} \left(c_1 \sum_{i=1}^m \frac{1}{l_{r_i, n_i}} \right) - c_1 \right] = 0 \end{aligned}$$

证毕!

3 应用实例

在与某合作单位开展的电机耐久性试验研究中,先后进行了两次试验.每次试验各投入5个产品,采取定数截尾方式,截尾数分别为2和1,观测到的失效时间列于表1最后一列.该型电机两次试验的样品来自同一批产品,试验条件相同,可以看作来自同一母体.由于第2组试验只有一个失效数据,传统的两步融合估计方法无法对其进行分析.下面采用本文提出的整体融合方法对表1数据进行处理,给出电机寿命的融合估计.

表1 某型电机寿命试验数据
Table 1 Life data of motors

组别 i	样本量 n_i	序号 j	失效时间 t_{ij}/h
1	5	1	6 355
		2	10 102
		3	>10 102
		4	>10 102
		5	>10 102
2	5	1	7 967
		2	>7 967
		3	>7 967
		4	>7 967
		5	>7 967

该型电机寿命 t 服从对数正态分布,令 $X = \ln t$, 则 X 服从正态分布,即有

$$F(x_i) = \Phi\left(\frac{x_i - \mu}{\sigma}\right) \quad i = 1, 2$$

式中 μ 和 σ 为对数均值和对数标准差.采用本文方法得到 μ 和 σ 的整体融合估计 $\hat{\mu}$ 和 $\hat{\sigma}$ 分别为

$$\hat{\mu} = 9.579, \quad \hat{\sigma} = 0.610$$

并且

$$S_{11} = \frac{D(\hat{\mu})}{\hat{\sigma}^2} = 0.607$$

$$S_{22} = \frac{D(\hat{\sigma})}{\hat{\sigma}^2} = 0.571$$

$$S_{12} = \frac{\text{Cov}(\hat{\mu}, \hat{\sigma})}{\hat{\sigma}^2} = 0.497$$

该型电机的中值寿命估计为

$$\hat{t}_{0.5} = \exp(\hat{\mu}) = \exp(9.579) = 14456.0(h)$$

可靠度为 $R = 1 - P = 0.9$ 的可靠寿命估计为

$$\hat{t}_{0.9} = \exp(\hat{X}_{0.1}) = \exp[\hat{\mu} + \Phi^{-1}(0.1)\hat{\sigma}] = 6619.3(h)$$

4 结 论

提出的多组定数截尾试验数据整体融合方法能够将相同试验条件下的多组试验数据作为一个整体进行统计推断,建立了正态分布、对数正态分布、极值分布和 Weibull 分布等位置-尺度分布的位置参数、尺度参数和百分位值的整体融合估计和置信区间估计公式.

与传统方法相比,该方法能够同时利用所有组别下试验数据中包含的信息进行统计分析,避免了两步融合估计过程中的信息损失,因此,具有信息量大、精度高的特点.而且,该方法能够解决了部分组别下只有一个失效数据时的截尾试验数据融合难题.文中详细证明了整体融合估计量的优良性,并给出了一个应用实例.

此外,若引入区间数据区间统计量^[11],还可实现定时截尾数据^[12-13]和混合截尾数据的整体融合估计.若将本文公式中截尾数据顺序统计量的均值、方差和协方差换成不完全数据顺序统计量^[14]、区间数据区间统计量^[11]或不完全区间数据统计量^[15]的均值、方差和协方差,则本文方法也适用于不完全数据、区间数据和不完全区间数据的整体融合估计.

参考文献:

- [1] 戴树森,费鹤良,王玲玲,等.可靠性试验及其统计分析[M].北京:国防工业出版社,1983.
DAI Shusen, FEI Heliang, WANG Lingling, et al. Reliability test and its statistical analysis[M]. Beijing: National Defense Industry Press, 1983. (in Chinese)
- [2] 菲诗松,汤银才,王玲玲.可靠性统计[M].北京:高等教育出版社,2008.
MAO Shisong, TANG Yincui, WANG Lingling. Reliability statistics[M]. Beijing: Higher Education Press, 2008. (in Chinese)
- [3] 张志华.加速寿命试验及其统计分析[M].北京:北京工业大学出版社,2002.
ZHANG Zhihua. Accelerated life test and its statistical analysis[M]. Beijing: Beijing Industry University Press, 2002. (in Chinese)
- [4] 傅惠民,黄伟.最佳线性无偏整体估计方法[J].机械强度,2003,25(3):319-324.
FU Huimin, HUANG Wei. Best linear unbiased integral estimate method[J]. Journal of Mechanical Strength, 2003, 25(3): 319-324. (in Chinese)
- [5] 黄伟,傅惠民.截尾数据线性回归分析方法[J].航空动力学报,2003,18(4):465-469.
HUANG Wei, FU Huimin. Linear regression analysis method for censored data[J]. Journal of Aerospace Power,

- 2003,18(4):465-469. (in Chinese)
- [6] 中国电子技术标准化研究所. 可靠性试验用表[M]. 北京:国防工业出版社,1987.
- The Chinese Research Institute for Electron Technology. Handbook of table for reliability test[M]. Beijing:National Defense Industry Press,1987. (in Chinese)
- [7] Balakrishnan N, Chan P S. Order statistics from extreme value distribution: Part I—Table of means, variances and covariances[J]. Commun. Statist. -Simula., 1992, 21(4): 1199-1217.
- [8] Parrish R S. Computing expected values of normal order statistics[J]. Commun. Statist. -Simula., 1992, 21(1): 57-70.
- [9] Parrish R S. Computing variances and covariances of normal order statistics[J]. Commun. Statist. -Simula., 1992, 21(1):71-101.
- [10] 傅惠民, 林逢春. 大样本顺序统计量均值、方差和协方差计算与验证[J]. 机械强度, 2007, 29(1):48-52.
- FU Huimin, LIN Fengchun. Computation and verification of means, variances and covariances for large sample order statistics[J]. Journal of Mechanical Strength, 2007, 29(1): 48-52. (in Chinese)
- [11] 傅惠民. 区间统计量及其分布[J]. 机械强度, 2005, 27(6): 752-757.
- FU Huimin. Interval statistic and its distribution[J]. Journal of Mechanical Strength, 2005, 27(6):752-757. (in Chinese)
- [12] 傅惠民, 岳晓蕊. 定时截尾数据最佳线性无偏估计方法[J]. 机械强度, 2009, 31(6):905-909.
- FU Huimin, YUE Xiaorui. Best linear unbiased estimate method for type-I censored data[J]. Journal of Mechanical Strength, 2009, 31(6):905-909. (in Chinese)
- [13] 傅惠民, 岳晓蕊. 定时截尾数据回归分析方法[J]. 航空动力学报, 2010, 25(1):1-6.
- FU Huimin, YUE Xiaorui. Regression analysis method for type-I censored data[J]. Journal of Aerospace Power, 2010, 25(1):1-6. (in Chinese)
- [14] 傅惠民, 王华胜. 不完全数据最佳线性无偏估计方法[J]. 航空动力学报, 2004, 19(2):169-173.
- FU Huimin, WANG Huasheng. Best linear unbiased estimate method for incomplete data[J]. Journal of Aerospace Power, 2004, 19(2):169-173. (in Chinese)
- [15] 傅惠民, 敖亮. 不完全区间数据最佳线性无偏估计方法[J]. 航空动力学报, 2007, 22(7):1025-1029.
- FU Huimin, AO Liang. Best linear unbiased estimate method for incomplete interval censored data[J]. Journal of Aerospace Power, 2007, 22(7):1025-1029. (in Chinese)