卫星可靠性：统计分析与建模

在太空、水下等环境中使用的设备，一般具有维修的成本较高、甚至维修不可操作等特点，因此在实际应用中，一般要求这些设备具有较高的可靠性。这篇综述就是以这类设备的一种——卫星——为例，详细阐述关于这类设备的可靠性的统计分析和建模。对卫星进行可靠性统计分析，尤其是关于实际的在轨可靠性和多状态失效情形的分析，对卫星的设计卫星设计子系统的时候非常有用。但是，由于卫星多与政府、国防相关联，关于在轨卫星的使用数据和可靠性统计分析很缺乏。因此系统地阐述一下卫星的可靠性统计分析与建模，很有必要。

1. 卫星数据介绍

这篇综述所用的数据集由在 1990 年 1 月到 2008 年 10 月发射的 1584 颗绕地卫星的运行情况的记录组成，对于每一颗卫星，数据集记录了它的发射时间、失效时间、导致卫星失效的关键子系统以及删失时间。实际中，每一颗卫星的生产过程、材料种类和运行环境都是不同的，因此它们的失效模式不是独立同分布的。如果按照卫星的自身性质和飞行任务对卫星进行分类，在每一类中认为卫星独立同分布，用同一类的卫星进行统计分析，这样样本量会变小，参数的估计方差会较大，不利于进行统计推断。所以，我们可以认为所用的所有卫星的失效模式是独立同分布的。

2. 卫星失效数据的非参数估计

2.1. 删失数据. 当数据的失效时间未被观测到时，如在实验过程中未能记录到准确的失效时间、在实验结束时被观测对象仍未失效，数据便产生了删失。数据删失的类型有很多种，在统计分析中，若不考虑到删失数据，统计推断结果会产生较大的偏差。本文中考虑的数据主要是相异进入的右删失数据，具体来说有以下特点：（1）样本进入实验的时间（即发射时间）是不相同的，但都是已知的；（2）失效时间和删失时间是随机分布的；（3）产生删失的原因要么是卫星在未失效时就已经被召回，要么是在实验结束时卫星仍在工作。由于卫星的发射时间不同，我们可以估计卫星的可靠性函数随工作时间（从成功发射算起）的变化，这就解决了各个卫星进入实验的时间不同的问题。当数据存在删失时，统计建模与推断过程必须考虑数据删失的情况，并作出相应的修正，否则推断结果可能出现很大的偏差。对于存在删失的数据集，一种常用的方法是 Kaplan-Meier 估计。

2.2. Kaplan-Meier 估计. Kaplan-Meier 估计是一种常用于当数据存在删失值时候的估计方法，可以用于估计可靠性函数。它的使用过程如下，首先我们考虑数据集完整时候的情形。完整数据是每个样本的失效时间都观测到了的数据集，可以记那些失效时间为

$$t_{(1)} < t_{(2)} < \cdots < t_{(n)},$$

经验可靠性函数的定义为

$$R_n(t)=\frac{\text{在时刻 t 还未失效的样本数}}{\text{样本总数}n}$$

由完整数据得到的经验可靠性函数的估计为

$$R_n(t)=\begin{cases}1, & t<t_{(1)}\\ 1-\frac{i}{n}, & t_{(i)}\leq t\leq t_{(i+1)},1\leq i<n\\ 0, & t\geq t_{(n)}\end{cases}$$

显然，$R_n(t)$ 是一个阶梯函数。当存在结点（在某一时刻有多于一个样本失效）时，$R_n(t)$ 的形式稍作调整。

当数据存在删失时，Kaplan-Meier 的估计如下。记在 n 个样本中，观测到 m 个样本的准确失效时间，记为 $t_{(1)}<t_{(2)}<\cdots<t_{(m)}$，其余样本的失效时间为删失数据，可能介于记录到的某两个失效时间之间，也可能在实验结束时也未失效。记

$$\begin{aligned}n_i&=\text{在时刻}t_{(i)}\text{之前还在工作的样本数}\\ &=n-\text{在时刻}t_{(i)}\text{之前删失的样本数}-\text{在时刻}t_{(i)}\text{之前失效的样本数}\end{aligned}$$

定义

$$\hat{p}_i=\frac{n_i-1}{n_i}$$

则 $\hat{p}_i$ 是概率值

$$P(T>t_{(i)}+\delta t|T>t_{(i)})$$

的一个估计，其中 T 表示一个样本正常工作的时间。又因为

$$\begin{aligned}P(T>t_{(i)})=P(T>\delta t)\times P(T>t_{(1)}+\delta t|T>t_{(1)})\times P(T>t_{(2)}+\delta t|T>t_{(2)})\\ \times\cdots P(T>t_{(i)}+\delta t|T>t_{(i)})\end{aligned}$$

以及

$$R(t)=P(T>t),$$

则可靠性函数的 Kaplan-Meier 估计估计为

$$\hat{R}(t)=\prod_{t_{(i)}\leq t}\hat{p}_i=\prod_{t_{(i)}\leq t}\frac{n_i-1}{n_i}$$

当存在结点时，如在时刻 $t_{(i)}$ 处有 m_i 个样本失效，则 $\hat{p}_i=\frac{n_i-m_i}{n_i}$。

3. 卫星可靠性的参数估计

非参数估计由于不假定模型的分布而有很大的适用性和灵活性，但有时又会因为这一灵活性而缺少必要的精度，这是就需要用到参数估计。在可靠性分析中，威布尔分布是一种经常用到的分布。威布尔分布很灵活，可以拟合各种失效情况。通过控制参数的取值范围，威布尔分布可以拟合包括失效率递减、失效率恒定和失效率递增等情形，它的失效率函数为

$$\lambda(t) = \frac{\beta}{\theta}(\frac{t}{\theta})^{\beta-1}, \theta > 0, \beta > 0, t \geq 0,$$

则可靠性函数

$$R(t) = exp(-(\frac{t}{\theta})^{\beta})$$

对可靠性函数取对数有

$$\ln[-\ln R(t)] = \beta \ln t - \beta \ln \theta$$

可以看到 $\ln[-\ln R(t)]$ 与 $\ln t$ 呈线性关系。这提供了一种检验威布尔分布这种假定是否正确的方法。具体来说，我们可以对数据首先用非参数的方法进行分析，得到的 $\hat{R}(t)$, 画 $(t, R(t))$ 形成的散点图，若散点呈线性关系，就可以认为参数方法中用威布尔分布拟合可靠性函数是合理的，用构成散点图的这些点作线性回归可以得到 β, θ 的估计值。

除了上面提到的 Kaplan-Meier 估计的非参数方法以及指定威布尔分布的参数方法，还可以用样条回归来非参数地分析数据，可以假定可靠性函数的分布为指数分布的混合分布。

4. 置信区间的估计

当用非参数的方法估计出卫星的可靠性函数时，我们也想要得到可靠性函数的置信区间，来检验估计的准确性。Kaplan-Meier 估计方法得到的置信区间为

$$v\hat{a}r[R(t_i)] = \sigma^2(t_i) = [\hat{R}(t_i)]^2 \sum_{j \leq i} \frac{m_j}{n_j(n_j - m_j)}$$

则它的 95% 置信区间为

$$R_{95\%}(t_i) = \hat{R}(t_i) \pm 1.96\sigma(t_i)$$

5. 卫星子系统的可靠性分析

在观测卫星的飞行情况时，我们也可以记录卫星子系统的运行情况，将上述分析卫星可靠性的过程应用于卫星的子系统，可以得到卫星子系统 j 的可靠性函数，记为

$$\hat{R}_j(t), j = 1, 2, \cdots, k$$

则子系统 j 的失效引发卫星失效的概率函数为

$$\hat{P}_j = 1 - \hat{R}_j$$

子系统 j 对整个卫星系统失效所起的作用为

$$r_j = \frac{\hat{P}_j}{\hat{P}}, \hat{P} = 1 - \hat{R}$$

由此我们可以得到每个卫星子系统的可靠性在整个卫星可靠性中所起的作用，也有利于发现关键的子系统。

6. 结论

分析卫星的可靠性，用非参数的方法进行估计，作出散点图，然后用参数的方法拟合得到具体的表达式，由此得到卫星的失效模式。这一整个流程可以让我们对卫星的使用情况得到一个良好的认识，也有利于卫星的合理设计和使用。

7. 参考文献

1、Castet, J.-F, and Saleh, J.H, ”Statellite Reliability：Statistical Data Analysis and Modeling”, Journal of Spacecraft and Rockets, Vol.46, No.5, September-October 2009

2、Castet, J.-F, and Saleh, J.H,”Satellite and Satellite Subsystems Reliability: Statistical Data Analysis and Modeling”, European Safety and Reliability Conference 2009, 7-10 September 2009