可靠性基础知识讲座

第五讲 可靠性评估

廖 炯 生

上一讲介绍可靠性试验。产品抽样进行试验时要用统计学抽样方法，对于试验结果要用统计学分析与推断（评估）方法。

产品可靠性与它的失效分布有密切关系。当失效分布类型已知而参数未知时，根据子样试验数据对失效分布的参数进行估计的方法分为点估计和区间估计。点估计方法有矩法、极大似然法等。

不同抽样的样本的点估计值一般是不同的，因为抽样有随机性。点估计没有给出估计值$\hat{\theta}$和真值θ的误差是多少。因此工程上更多的采用区间估计

$$P_r(|\theta-\hat{\theta}|\leqslant\delta)=1-\alpha$$

即真值θ在$\theta_L=\hat{\theta}-\delta$和$\theta_U=\hat{\theta}+\delta$两界限之间的概率为$1-\alpha=\gamma$。$[\theta_L,\ \theta_U]$叫置信区间，γ叫置信度，α叫风险率。

γ反映区间估计的把握性这一侧面，而另一侧面是区间估计的精确性。在计算方法和置信度要求相同的条件下，子样（试验成本）越大，则置信区间越窄，估计精度越高。当子样大小不变，若要求置信度越高，则置信区间越宽，估计精度越低。这是互相矛盾的两个侧面。关键在于充分利用子样试验信息，作出尽可能可信的和精确的估计。

在工程上，主要关心产品可靠性下限或失效率上限，这叫单侧区间估计。上面说的是双侧区间估计。

单元产品可靠性评估

在可靠性技术中最常见的三种失效分布类型是成败型（二项分布）、寿命型（一般为指数分布，还有威布尔分布等）、性能和结构可靠性模型（多为正态分布）。下面分别介绍。

成败型

如果产品每次试验（检验）只可能成功或失败（合格或不合格），假设在一个独立试验序列中每次试验成功的概率R恒定不变（贝努利概型）。由二项分布可知，n次试验中成功s次的概率（似然函数）为

$$L(R)=\binom{n}{s}R^s(1-R)^{n-s}$$

令 $$\frac{d}{dR}L(R)=\frac{d}{dR}\ln L(R)=0,$$

解得 $\hat{R}=S/n$

这就是极大似然估计。它是在R的一切可能取值中使观测结果（n，s）出现概率最大的一个取值，我们就把它作为R的估计值。它容易求得，在一定条件下具有一些最优性质，但不总是无偏的。而且试验无失效或全失效时不能用。

矩法估计结果同上。估计的方差为

$$\hat{\sigma}^2_{\hat{R}}=\frac{\hat{R}(1-\hat{R})}{n}$$

当子样n不大时，方差不小。所以小子样时慎用。

实际上更有用的是区间估计。单侧置信下限估计：

$$P_r(R_L\leqslant R)\geqslant\gamma$$

R_L由下式确定$(F=n-s)$：

$$\sum_{X=0}^{F}\binom{n}{X}R_L^{n-X}(1-R_L)^X=1-\gamma$$

根据试验结果（n,F）及要求置信度γ查国家标准GB 4087.3—85《二项分布可靠度置信下限》可得R_L。

双侧置信区间估计：

$$P_\gamma(R_L\leqslant R\leqslant R_U)\geqslant\gamma,$$

其中R_L、R_U满足

$$\left.\begin{aligned}\sum_{X=0}^{F}\binom{n}{X}R_L^{n-X}(1-R_L)^X&=\frac{1}{2}(1-\gamma)\\ \sum_{X=F}^{n}\binom{n}{X}R_U^{n-X}(1-R_U)^X&=\frac{1}{2}(1-\gamma)\end{aligned}\right\}$$

上式可化为

$$\left.\begin{aligned}\sum_{X=0}^{F}\binom{n}{X}R_L^{n-X}(1-R_L)^X&=1-\frac{1+\gamma}{2}\\ \sum_{X=0}^{F-1}\binom{n}{X}R_U^{n-X}(1-R_U)^X&=\frac{1+\gamma}{2}\end{aligned}\right\}$$

则可由GB 4087.3—85，用（n，F）及$1-\frac{1+\gamma}{2}$查出R_L；

用（n,F－1）及$\frac{1+\gamma}{2}$查出R_U。

以上求得非随机化最优置信上限R_U和下限R_L。

正态型

设产品某项性能指标X是正态分布的随机变量$N(\mu,\sigma)$。随机抽取n个样品，测得$x_1,x_2,\cdots,x_n$；据此评估未知参数μ、σ，则矩法估计和极大似然估计结果同为

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X},\ \hat{\sigma}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X}_i)^2$$

随机变量$\overline{X}$的期望值为μ，随机变量$\hat{\sigma}^2$的期望值为$\frac{n-1}{n}\hat{\sigma}^2$，令$S^2=\frac{n}{n-1}\hat{\sigma}^2$，则$S^2$的期望值为$\sigma^2$。在n不大时以$\overline{X}$、$S^2$作为μ、$\sigma^2$的估计值是存在一定误差的。

下面讨论性能可靠性的区间估计。

定义 $P_\gamma(X\geqslant a)=R$

则有 $a=\mu-U_{1-R}\sigma$，U_{1-R}是正态分布分位点，用n次观测值$x_1,x_2,\cdots,x_n$算出μ、σ的估计值$\overline{X}$、S，如选系数K使得

$$P_\gamma(\overline{X}-KS\leqslant\mu-U_{1-R}\sigma)=\gamma,$$

则可认为$X\geqslant\overline{X}-K_S$的概率至少为R，置信度为γ。

令性能参数下限$a=\overline{X}-K_S$，即容许限系数$K=\frac{\overline{X}-a}{S}$，由n、K、γ反查国家标准GB4885—85，可得到R值，即$X\geqslant a$的性能可靠性。

若给定性能参数X的上限M，则$K=\frac{M-\overline{X}}{S}$，同理，由n、K、γ反查GB4885—85，可得$X\leqslant M$的性能可靠性R。

若要求 $L\leqslant X\leqslant M$ 双尾控制下性能可靠性置信下限，可进一步参考有关文献。

指数分布寿命型

产品寿命试验进行到全部样品失效的情形是很少的，那样耗费太大。常见的是各种截尾寿命试验。在指数分布条件下，根据各种截尾寿命试验数据进行点估计和区间估计的公式参见本讲座第四讲《可靠性试验》一文的表3和表4。

复杂产品可靠性评估

对于由许多不同单元组成的复杂产品，例如导弹、人造卫星等大型复杂系统，虽然在原理上也可以按照上述单元产品可靠性评估方法去做；但在实际上，要用大量试验样品、很多试验次数或很长试验时间，对于这类大型系统来说是不现实的。由于人力、资金、时间限制，小子样长寿命大系统可靠性评估是统计学上一大困难问题。

系统可靠性（或其他性能）的金字塔式试验程序或综合与评定程序（“金字塔模型”）是说，在实验室对大量元器件、原材料进行模拟使用试验或加速寿命试验，取得基本可靠性数据；在此基础上进行整机可靠性试验（台数和试验量都少一些），并把大

量元器件试验信息折合上来，对整机可靠性进行综合；按此格式逐级向上综合和评定各分系统和全系统的可靠性。这样，就有可能用少量全系统试验甚至不经过全系统试验而对大型复杂系统可靠性作出评估。要点在于各级试验信息的逐级向上折合和充分利用。但因为由单元组成系统的方式多种多样（如串联、并联、贮备、表决系统及复杂可靠性网络），单元的失效分布类型可不相同，所以复杂可靠性的金字塔式综合评估方法的困难还很大。拿最简单的串联系统来说，不能按各单元产品可靠性置信下限的乘积来进行串联系统可靠性综合。那样做不但过分保守，而且给不出确切的置信度，所谓"链条模型"没有考虑各环节失效的随机性，而各单元可靠性置信下限的估计存在随机性。又拿最常见的成败型、寿命型、应力强度型或性能型单元来说，当各单元可靠性认识值的分布已知时，系统可靠性认识值的分布如何综合，一般很难求得解析解。

系统可靠性评估方法分经典法、Bayes法、信任法三类，国内外都存在着这三个学派的争论。对此我们认为在学术研究上应当贯彻百家争鸣的方针，在型号工程实践上以适用的近似置信限方法为主，同时加强严格置信限理论研究。后者一旦获得突破，多种多样近似置信限方法就有了比较和判别的标准，问题也迎刃而解了。

下面扼要介绍经典法严格置信限理论及其困难，然后介绍几种近似置信限方法。

1957年Buehler提出经典法严格置信限理论的要点是：考虑m个成败型单元串联之系统，第i单元试验n_i次，成功x_i次；要求在各单元试验数据基础上以置信度γ推断系统可靠性下限$L\gamma(X)$，其中试验向量$X=\{x_1, x_2, \cdots, x_m\}$。

$L\gamma(X)$必须满足以下一组条件：

1，精确性　$P_\gamma\{R\geqslant L\gamma(X)\}\geqslant\gamma$，$0\leqslant R\leqslant 1$；

2．正则性　若$X_1<X_2$，则$L\gamma(X_1)\leqslant L\gamma(X_2)$；

3．最优性　$L\gamma(X)$应尽可能大。

试验向量的总数为$N=\prod\limits_{i=1}^{m}(n_i+1)$，假设N个试验向量已按正则性条件排好序，对于观察到的某试验向量$X_j=\{x_{1j}, x_{2j}, \cdots, x_{mj}\}$，$L\gamma(X_j)$服从下式：

$$L_\gamma(X_j)=\inf\left\{R\,\middle|\,\sum_{k}^{N}{}_{j} B_k=1-\gamma\right\}$$

其中　$B_k=\prod\limits_{i=1}^{m}\binom{n_i}{x_{ik}}p_i^{x_{ik}}(1-p_i)^{n_i-x_{ik}}$

p_i为单元i的可靠性。

严格置信限理论困难在于：

1．试验向量的排序，现有的Buehler、UMVE、LP和MLP排序法都会出现非正常点[1]；

2．上述方程的求解，它是一个非线性规划问题。

关键的困难是排序问题，三十一年来未能突破。因此，考虑实践需要，提出了多种近似置信限方法。

修正极大似然估计法(MML)[2]

成败型单元串联系统可靠性的极大似然估计为

$$\hat{R}=\prod_{i=1}^{m}x_i/n_i$$

估计的方差为

$$\hat{\sigma}^2_{\hat{R}}=\prod_{i=1}^{m}\left(\frac{x_i}{n_i}\right)^2\sum_{i=1}^{m}\left(\frac{n_i-x_i}{n_i x_i}\right)$$

设系统等效试验次数为n，成功次数为x，按成败型（二项分布）应有

$$\hat{R}=x/n$$

$$\hat{\sigma}^2_{\hat{R}}=[\hat{R}(1-\hat{R})]/n$$

以上四式联立解得

$$n=\frac{\prod_{i=1}^{m}\frac{n_i}{x_i}-1}{\sum_{i=1}^{m}\frac{1}{x_i}-\sum_{i=1}^{m}\frac{1}{n_i}},$$

$$x=n\prod_{i=1}^{m}\frac{x_i}{n_i}$$

由折合数据(n,x)，给定γ,查GB4087.3即得系统可靠性近似置信下限R_L。

逐次压缩法(SR)[3]

文献[3]对一些近似置信限方法作了比较，认为MML法计算简单而准确，但它不能考虑无失效单元的影响。如果试验次数最少的单元无失效，不考虑它将使MML法的结果偏于冒进。针对这一局限性，[3]提出SR法，步骤如下：

a.将$n_i(i=\overline{1,m})$从大到小排序；

b.$i=1$

c.若$x_i\geqslant n_{i+1}$,取$X_{i+1}=X_{i+1}$，$n_{i+1}=n_in_{i+1}/x_{i+1}$；

若$x_i\leqslant n_{i+1}$,取$X_{i+1}=X_iX_{i+1}/n_{i+1}$，$n_{i+1}=n_i$；

d.$i=i+1$，若$i<m$，返回c；否则转e；

e.取$x=x_i$，$n=n_i$，即系统等效试验结果。据(n,x)及γ可求R_L。

文献[4]指出,SR法按照点估计不变的原则压缩单元试验数据的结果总是损失信息，使估计方差增大，因为

若$\hat{p}_i=\frac{x_i}{n_i}=\frac{x'_i}{n'_i}$，　当$n'_i<n_i$，

则$\sigma^2_{\hat{p}_i}=\frac{\hat{p}_i(1-\hat{p}_i)}{n_i}<\frac{\hat{p}_i(1-\hat{p}_i)}{n'_i}$

所以得到的置信限偏于保守，逐次压缩势必造成系统可靠性置信限更加保守。

L—M法 (Lindstrom—Madden)

$$\hat{R}=\prod_{i=1}^{m}x_i/n_i$$

取$\hat{R}=y_1/n_1$，其中$n_1=\min_i\{n_i\}$

再根据系统等效试验数据(n_1,y_1)和γ求R_L。

LM法和SR法都是按照点估计不变的原则压缩单元试验数据，所以求得近似置信限都偏于保守。但LM法压缩得更多,所以其结果比SR法更差，不宜推荐使用。试看二单元串联系统：

设单元试验数据(n_1,x_1),(n_2,x_2)且$n_1>n_2$,$n_1>x_1$；若$x_1\geqslant n_2$,按SR法,则折合试验次数$n_{SR}=\frac{n_1n_2}{x_1}>n_2$，$x_{SR}=x_2$；

按LM法，则$n_{LM}=\min\{n_1,n_2\}=n_2$，$x_{LM}=\frac{x_1x_2}{n_1}$,若$x_1\leqslant n_2$,按SR法，有$n_{SR}=n_1>n_2$，$x_{SR}=x_1x_2/n_1$；

按LM法，$n_{LM}=n_2$，$x_{LM}=x_1x_2/n_1$

所以恒有　$n_{SR}>n_{LM}$，$\frac{x_{LM}}{n_{LM}}=\frac{x_{SR}}{n_{SR}}$

即LM法和SR法点估计相同,但前者的估计方差较大。由GB 4087.3不难验证，对于给定置信度γ,点估计相同而方差加大时，求得R_L减小。

CMSR法[4]

CMSR法的要旨是把MML和SR法结合起来使用，取其优点而克服各自的缺点。CMSR法是在试验次数最少的单元无失效时只进行一次压缩和折合（不同于SR法要进行$m-1$次），然后应用MML法公式求等效试验数据(n,x)。如果串联系统中不存在试验次数最少的无失效单元，则CMSR直接等同于MML法。

设$n_1\geqslant n_2\geqslant\cdots\geqslant n_m$，其中$j$个单元无失效：

$n_k=x_k$，$k=m-j+1$，$m-j+2$，$\cdots$，m；

$n_l\neq x_l$，$l=1$，2，$\cdots$,$m-j$；

CMSR法的步骤是：

a．根据有关定理，仅考虑后j个无失效单元试验结果等效于最后一个单元进行了试验

$n'_m=n_m$，$x'_m=x_m$，而且$x_m=n_m$。

b．将(n'_m,x'_m)和(n_{m-j},x_{m-j})按SR

法进行一次压缩和综合，得到(n'_{m-j},x'_{m-j})：

当$x_{m-j}\geqslant n_m$时取$x'_{m-j}=x_m$，$n'_{m-j}=n_{m-j}n_m/x_{m-j}$；

当$x_{m-j}\leqslant n_m$时取$x'_{m-j}=x_{m-j}$，$n'_{m-j}=n_{m-j}$

c.数据列$(n_1,x_1),(n_2,x_2),\cdots(n_{m-j-1},x_{m-j-1}),(n'_{m-j},x'_{m-j})$已满足MML法应用条件，可用MML公式求n,x。给定置信度γ,查GB4087.3即得R_L。

举例

四个成败型单元串联系统

$(n_i,x_i)=(120,119),(100,99),(45,45),(30,30)$

其中存在试验次数最少的无失效单元，不宜用MML法。假如硬用MML公式折合得

$n=108.685,\qquad x=106.701$

取置信度$\gamma=0.9$查GB4087.3得$R_L=0.9519$,它偏于冒进。

用SR法将全部单元试验数据压缩综合成系统等效试验数据 $n=30.5577$，$x=30$

取$\gamma=0.9$ 查GB4087.3得$R_L=0.8998$,它偏于保守。

再用LM法

$$y_1=\frac{119}{120}\times\frac{99}{100}\times\frac{45}{45}\times\frac{30}{30}\times30=29.4525$$

其中 $n_1=\min\{n_i\}=30$

取$\gamma=0.9$ 查GB4087.3得$R_L=0.8989$,它更加保守。

最后用CMSR法，按规定步骤将（100,99),(30,30)压缩一次综合成为（30.303,30)；根据有关定理，中间无失效单元(45,45)可略去。

此等效结果与第一单元试验结果组合为（120,119），(30.303,30)；它符合MML法应用条件，用MML公式折合得

$n=46.2115,\qquad x=45.3726$

取$\gamma=0.9$ 查GB4087.3得$R_L=0.9235$,因为它毕竟也把单元试验结果压缩了一次，试验信息有所损失，所以其结果可能还稍保守，但它比SR法、LM法都更不保守。

以上就经典法严格置信限理论和近似置信限方法作了简略介绍。限于篇幅，对Bayes法和信任法就不作介绍了。Bayes观点认为应当考虑历史的经验（表现为待估参数的验前分布），再综合当前抽样试验的结果来进行统计推断（评估）。对于小子样复杂系统可靠性综合来说，Bayes观点是很有吸引力的。关键困难在于如何准确地稳健地把工程历史经验反映到验前分布上，还需要深入研究。

宋健同志说过，航天产品可靠性评估的意义在于“提供量的概念，发现薄弱环节、进行科学指导”。

在可靠性工程中，可靠性评估是一道“后工序”，是一个总结性环节。型号产品研制从指标论证、方案论证开始，要求同时确定可靠性指标、建立可靠性模型、进行可靠性指标分配和预计，即开展可靠性设计工作。相应地，要求开展故障模式与效应分析、故障树分析、元件应力与参数容差分析、潜在电路或软件潜在通路分析等可靠性分析工作。还要进行可靠性设计评审和可靠性试验。在这整个过程中，要不断地揭露薄弱环节，采取纠正措施，以获得预期的可靠性增长。最后，完整收集和充分利用产品的金字塔式各层次上的试验信息，作出系统可靠性综合评估。它是对于原订可靠性指标的评价和检验。对于导弹武器和卫星来说，发射前的可靠性评估可以提供量的概念，做到心中有数，为领导决策提供信息。型号产品可靠性评估大纲要求也是推动可靠性设计和分析，计划可靠性试验，推动试验数据收集的一个手段。

航天产品的继承性是极重要的。我们的许多新型号是在老型号产品的基础上发展而来的。在老型号上飞行试验成功的元器件、原材料、部件和设备，用于新型号就更有把握。尽可能采用经过飞行考验的元部件是航天产品的一条设计原则。因此，现有型号产品各级单元的可靠性评估结果加上实际飞行情况对于后续型号产品研制来说是极为宝贵

的验前信息，也是论证确定后续型号可靠性指标的重要依据。

当然，对于导弹、卫星这样的小子样（长寿命）大型复杂系统来说，统计推断的困难极大，不要把可靠性评估结果那个数值绝对化，国外对于大系统高可靠性不作最后评估的情形也是不少的，那就在分系统或部件级去作必要的验证。不管怎样，对于可靠性指标论证与设计分析，和研制全过程中的可靠性保证工作，都要求极严格极认真地去做。

参 考 文 献

[1] 黄柏琴、金振宏，宇航学报，1987，№.2，62—67

[2] Easterling R.G.,Journal of American Statistical Association,Vol.67(1972), 220—222

[3] Preston P.F., ADAO47533,1976

[4] 朱晓波、廖炯生，系统可靠性评估的CMSR方法，宇航学报

（上接第42页）

七、综合效果

1． 改进后本体组件产品质量，生产效率明显提高，合格率由73%提高到99.5%，生产效率提高 25.58%，使本体组件成本由414.36元下降到289.65元，下降30.1%。

2． 以87年年产1026件，计算一年可节省127952.46元。

3． 阀芯槽口测量技术成果，通过了部级技术鉴定。

八、几点体会

1． 价值工程与QC活动相结合，分析问题清晰、准确，措施精练。

2． 价值工程是提高产品质量，降低成本的有效管理技术。

3． 院校与工厂合作搞科研效果好。

九、今后的打算

1． 围绕班组科学管理，民品开发，继续开展QC小组活动。

2． 运用价值工程搞好民品开发。

3． 继续学习全面质量管理技术，提高班组的质量管理水平。

（上接第23页）

任务时间t得 $(Z, \eta) = (Z, \tau/t) = (3.5, 32.435)$。于是由(17)～(19)式得 $Au = 2.5729E-4$，$\theta L = 1/Au = 3886.7$（小时），$RL = EXP(-tAu) = 0.8309$。

参 考 文 献

[1] Martz & Waller, Bayesian Reliability Analysis, Wiley, 1982

[2] Springer & Thompson, Bayesian Confidence Limits for the Reliability of Cascade Exponential Subsystems, IEEE Trans. on Reliability Vol.R—16 No.2, 1967

[3] D.R.Smith, An Analysis Regarding the Determination of Bayesian Confidence Limits for the Reliability of Distribution-Free parallel Subsystems, the Theroy &Application of Reliability Vol.II, 1977

[4] E.S.Pearson & H.O.Hartley, Biometrika Tables for Statisticians, Vol.II, Appendix II, Table 3, 1972