



2024 机械可靠性学术年会 — 培训

可靠性统计分析

汤银才，徐安察

华东师范大学，浙江工商大学



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-刻度参数分布的统计推断
- 7 几个与可靠性相关的 R 程序包

现代数学基础

7 可靠性统计

■ 苏诗松 汤银才 王玲玲 编著

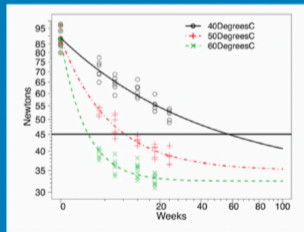
高等教育出版社
HIGHER EDUCATION PRESS

Wiley Series in Probability and Statistics

Second Edition

STATISTICAL METHODS FOR RELIABILITY DATA

WILLIAM Q. MEEKER | LUIS A. ESCOBAR | FRANCIS G. PASCUAL



WILEY



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-尺度参数分布的统计推断
- 7 常用可靠性相关软件程序包

产品的质量特性

☞ 产品 (元器件, 系统, 结构, 材料, 软件等) 质量特性概括起来主要有以下 5 个方面: 性能、寿命、**可靠性**、安全性和经济性。

☞ 硬件产品的质量特性:

- 1 **可靠性** (reliability)
- 2 维修性 (maintainability)
- 3 可用性 (availability)
 - (上述三个也称为可信性)
- 4 安全性 (safety)
- 5 保障性 (supportability)

☞ 软件产品的质量特性: 功能性 (functionality), **可靠性** (reliability), 易用性 (usability), 效率 (efficiency), 维修性 (maintainability), 可移植性 (portability)



可靠、失效与故障

☞ **可靠性**: 产品在规定的条件下和规定的时间内完成规定功能的能力。

☞ **失效和故障 (Failure and Fault)**:

- 产品终止完成规定功能的能力称为**失效**。
- 引起失效的物理、化学或其它过程称为**失效机理**。
- 对于可修产品, 失效常称为**故障**。

☞ **可靠度**: 可靠度就是指产品在规定条件下、规定时间内完成规定功能的概率。

影响产品寿命的因素

- ☞ 产品在规定的使用条件下开始工作到发生故障（失效）为止所经历的时间，称为产品的失效时间或**寿命**。
- ☞ 影响产品寿命的因素：
 - ▣ 选用的材料
 - ▣ 设计方案
 - ▣ 制造工艺
 - ▣ 规定功能

失效机理

- ☞ 物理 (硬) 失效机理 (physical-failure mechanism)
- ☞ 退化 (软) 失效机理 (degradation failure mechanism)
- ☞ 单失效机理
- ☞ 竞争失效
 - 竞争硬失效 — 多个硬失效机理
 - 竞争软失效 — 多个软失效机理
- ☞ 混合失效 — 硬失效与软失效共存

产品的失效类型

☞ **硬失效 (hard failure)** [也称为物理失效 (physical failure) 或灾难性失效 (catastrophic failure)]: 产品的功能突然并完全中止称为硬失效。按失效率曲线可分为:

- ▢ 递增失效曲线
- ▢ 递减失效曲线
- ▢ 浴盆曲线型失效

☞ **软失效 (soft failure) 或退化失效**: 产品的性能随时间连续退化, 当其指标 (参数) 逐渐退化到一个临界门限值时就称为软失效。

退化失效

☞ 按退化机理分为:

- ▢ 物理退化 (physical degradation): 如材料的老化、磨损
- ▢ 性能退化 (performance degradation): 如电流、光亮度的下降

☞ 按退化轨道类型分为:

- ▢ 线性退化 (linear degradation)
- ▢ 凹退化 (concave degradation)
- ▢ 凸退化 (convex degradation)
- ▢ S-型退化 (S-shape degradation)

可靠性试验

☞ 寿命试验: 正常应力下

- ▮ 定时截尾寿命试验
- ▮ 定数截尾寿命试验
- ▮ 混合截尾寿命试验

☞ 加速寿命试验

- ▮ 恒定应力加速寿命试验 (CS-ALT)
- ▮ 步进应力加速寿命试验 (SS-ALT)
- ▮ 序进应力加速寿命试验 (PS-ALT)
- ▮ 变应力加速寿命试验 (VS-ALT),
如温度循环加速寿命试验

☞ 双应力寿命试验/加速寿命试验, 如温湿度 ALT

☞ 高加速试验 (HALT)

- ▮ 注意失效机理发生变化!

☞ 退化试验

☞ 加速退化试验 (ADT)

- ▮ 恒定应力退化试验 (CS-ADT)
- ▮ 步进应力退化试验 (SS-ADT)
- ▮ 序进应力退化试验 (PS-ADT)
- ▮ 变应力退化试验 (VS-ADT)

退化试验与加速退化试验

为什么要进行退化试验 (DT)?

- ① 产品 (特别是系统中的部件和材料) 的可靠性越来越高
- ② 产品的失效主要是用软失效 (即性能的退化) 来刻划的
- ③ 在正常使用应力下无法或很难获得产品的失效数据
- ④ **退化数据提供了产品的可靠性信息**

为什么要进行加速退化试验 (ADT)?

- ① 在正常应力下产品退化太慢, 从而提供的信息较少
- ② 在加速寿命试验下产品仍无法提供足够的失效数据

可靠性数据收集

可靠性数据的收集有二种方法.

- 👉 现场寿命数据: 由于现场环境条件复杂多变, 收集来的数据来自不同工作环境下的产品寿命, 这种数据处理起来十分困难;
- 👉 实验室寿命试验数据: 这种来自相同条件下的寿命试验数据, 处理较为容易.

可靠性数据分析的特点

- 👉 数据通常是截尾或区间型的;
- 👉 大多数可靠性数据用正的随机变量的分布, 如指数分布、威布尔分布、伽玛分布、对数正态分布进行建模, 而正态分布很少直接使用;
- 👉 统计推断与预测经常涉及外推;
- 👉 感兴趣参数: 非分布参数, 而是失效概率、寿命分布的分位数、失效率等.

数据类型 (1) — 寿命数据

☞ 完全数据 — 失效数据

☞ 不完全数据

▢ 定时截尾数据

▢ 定数截尾数据

▢ 区组数据

▢ 混合截尾数据

▢ 竞争失效数据

▢ 无失效数据

☞ 再现数据 (recurrent event data) — 可修系统, 软件

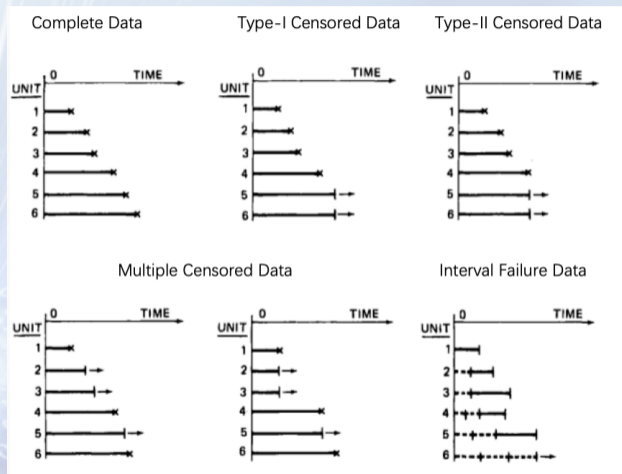


图 1.1: 失效数据类型

示例 1: 常应力下集成电路寿命试验数据

Failure Pattern of the Integrated Circuit Life Test Data Where 28 Out of 4156 Units Failed in the 1370 Hour Test at 80°C and 80% RH (Meeker 1987)

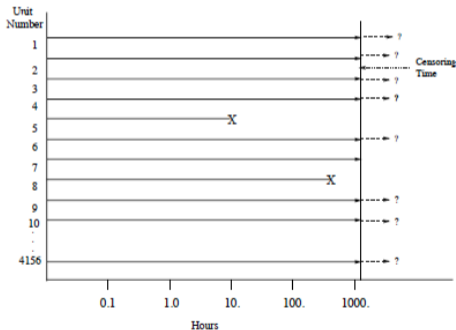


图 1.2: 常应力下集成电路寿命试验数据

示例 2: 温度 CS-ALT, 原始数据

Hours Versus Temperature Data from a ←
Temperature-Accelerated Life Test on Device-A

Hours	Status	Number of Devices	Temperature °C	In Subexperiment	
				Units	Failures
5000	Censored	30	10	30	0/30
1298	Failed	1	40	100	10/100
1390	Failed	1	40		
⋮	⋮	⋮	⋮		
5000	Censored	90	40		
581	Failed		60	20	9/20
925	Failed		60		
1432	Failed		60		
⋮	⋮	⋮	⋮		
5000	Censored	11	60		
283	Failed	1	80	15	14/15
361	Failed	1	80		
515	Failed	1	80		
638	Failed	1	80		
⋮	⋮	⋮	⋮		
5000	Censored	1	80		

图 1.3: 温度 CS-ALT, 原始数据

示例 2: 温度 CS-ALT, 数据图

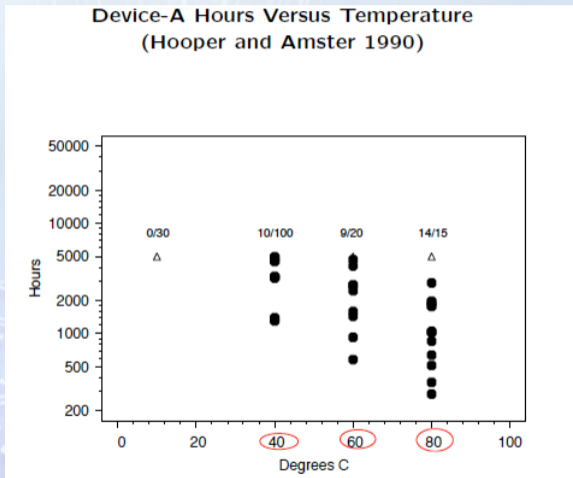


图 1.4: 温度 CS-ALT, 数据图



数据类型 (2) — 退化数据 (degradation data)

- ☞ 重复性退化数据 (repeated degradation data)
- ☞ 一次性 (毁灭性) 退化数据 (destructive degradation data)
- ☞ 其它退化数据, 如食品的货架寿命数据

示例 3: 退化试验, 裂缝的大小

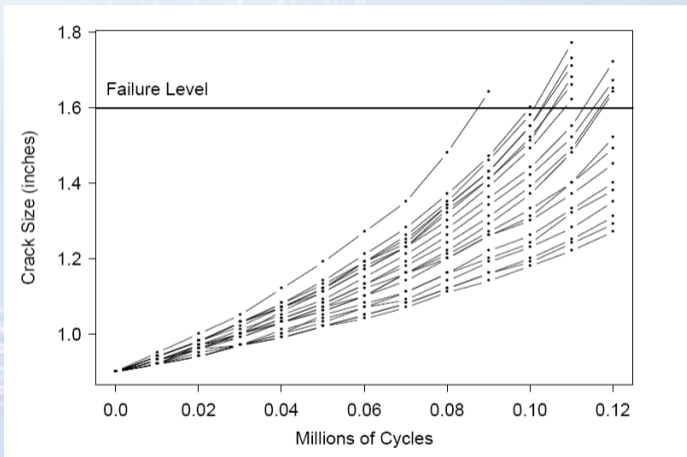


图 1.5: 退化试验, 裂缝的大小

示例 4: 三个应力下的温度 CS-ADT, 退化路径图

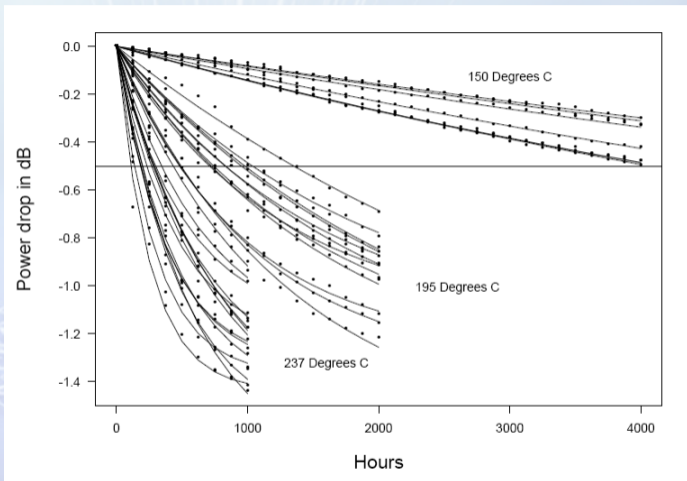


图 1.6: 温度 CS-ADT, 数据图

退化数据与失效数据的关系

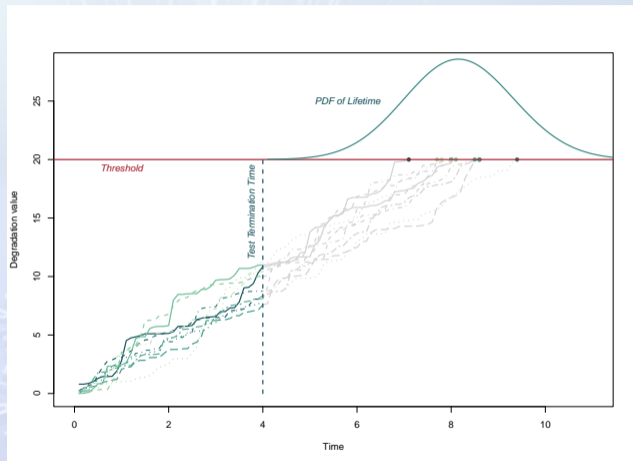


图 1.7: 退化截尾

可靠性统计分析建模

☞ 寿命试验: 正常应力下

- ▢ **失效分布**: 如 Weibull 分布, 对数正态分布.

☞ 加速寿命试验

- ▢ **失效分布**
- ▢ **加速方程**: 寿命与应力的关系
 - 逆幂律模型 (电压)
 - Arrhenius 模型 (温度), 广义 Arrhenius 模型
 - Eyring 模型 (电压与温度), 广义 Eyring 模型
 - Coffin-Manson 模型 (温度循环)
- ▢ **加速损伤机理**: 时间折算模型, 如 CE, TFR, TRV 模型.

☞ 退化试验: 正常应力下

- ▢ **退化轨道**: 如回归, 随机过程.

☞ 加速退化试验 (ADT)

- ▢ **退化轨道**:
 - 线性回归, 非线性回归
 - 维纳过程
 - 逆高斯过程: 复合泊松过程
 - 伽玛过程: 递增退化
 - 指数扩散过程
- ▢ **加速方程**: 退化率与应力的关系
- ▢ **加速损伤机理**: 时间折算模型, 如 CE 模型.



可靠性统计分析涉及的问题

- ① 寿命试验的统计分析: 点估计, 区间估计, 假设检验 (分布选择)
- ② 寿命试验的最优设计: 样品个数与试验时间分配, 应力水平的设定
- ③ 加速寿命试验的统计分析
- ④ 加速寿命试验的最优设计
- ⑤ 退化试验的统计分析
- ⑥ 退化试验的最优设计
- ⑦ 加速退化试验的统计分析
- ⑧ 加速退化试验的最优设计
- ⑨ 抽样检验



挑战与思考: 数据融合

- 👉 历史数据/经验与现有数据的融合
- 👉 实验室数据与现场数据融合
- 👉 寿命试验数据与退化试验数据的融合 (混合数据)
- 👉 测试数据与维修数据的融合
- 👉 小数据与大数据融合
- 👉 可靠性数据与用户行为数据的融合
- 👉 部件信息与系统结构信息的整合 (复杂系统)

挑战与思考: 方法的融合

- 👉 物理分析与数据分析的融合
- 👉 似然方法与贝叶斯方法的融合
- 👉 统计方法与机器学习方法的融合
- 👉 独立性方法与相依性方法的融合
- 👉 精确方法与近似方法的融合
- 👉 兼顾:
 - 🔺 精度、性能 (performance) 及扩展性 (scalability)
 - 🔺 开源与创新
 - 🔺 理论与应用



挑战与思考: 全生命周期的管理

- 👉 在线故障诊断
- 👉 预后健康诊断 (PHM)
- 👉 个性化质量评估与服务



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介**
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-尺度参数分布的统计推断
- 7 常用可靠性相关软件程序包

二大主流学派

👉 统计学任务：利用数据

- ▢ **推断** (点估计, 区间估计, 假设检验/模型选择与比较)
- ▢ 对模型进行**检验**(变量/模型的选择)
- ▢ **预测**

👉 统计学二大学派

- ▢ 经典统计学 (频率学派): 只利用总体与样本信息, 如最大似然估计
- ▢ 贝叶斯 (Bayes): 利用一切可以利用的信息, 得到贝叶斯估计

$$\text{信息} = \text{总体信息} + \text{样本信息} + \text{先验信息} \quad (2.1)$$

👉 争论长期存在, 但更多是哲学层面, 我们应该更务实。

👉 经典统计的局限: p -值不合理, 不遵从似然原理, 等等。

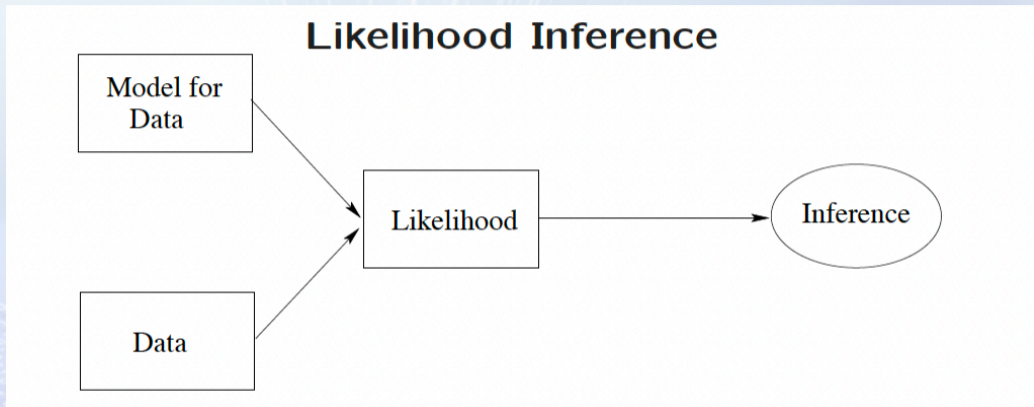


图 2.1: 似然推断

最大似然估计

数据: $y_1, \dots, y_n \stackrel{iid}{\sim} f(y|\theta), \theta \in \Theta^k$, 参数: $\theta = (\theta_1, \dots, \theta_k)$. 求 MLE(假定存在且唯一)的步骤为:

- 1 写出似然函数和对数似然函数

$$L(\theta) = \prod_{i=1}^n f(y_i|\theta) \quad (2.2)$$

$$l(\theta) = \ln L(\theta) \quad (2.3)$$

- 2 求解似然方程组 (假定密度函数微)

$$\frac{\partial}{\partial \theta_i} \ln L(\theta) = 0, i = 1, 2, \dots, k. \quad (2.4)$$

求解方法: Newton-Raphson 方法; (加速) EM 算法.

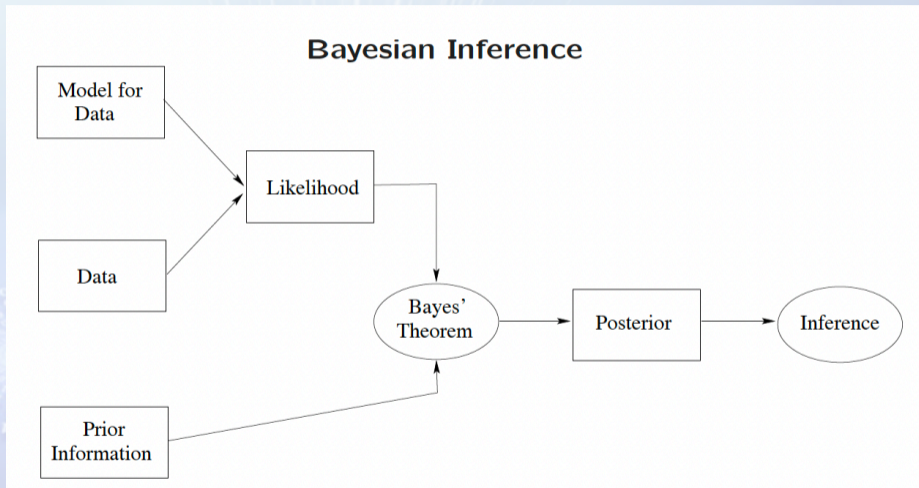


图 2.2: 贝叶斯推断

贝叶斯推断

- 👉 参数是随机的
- 👉 使用样本以外的额外信息

两类信息:

- 👉 关于参数的 (先验) 信息: $\theta \sim \pi(\theta)$, $\theta \in \Theta \subset \mathbb{R}^k$;
- 👉 关于分布的 (样本) 信息: $y \sim p(y | \theta) = L(\theta)$.

后验推断方法: 当 $p(y | \theta)$ 过于复杂时

- 👉 高效的直接抽样: ARS
- 👉 MCMC 方法:
 - 👉 Gibbs 抽样
 - 👉 MH 抽样
- 👉 贝叶斯近似: ABC, INLA

- 👉 在没有数据 y 时基于先验分布 $\pi(\theta)$ 进行推断;
- 👉 一旦得到了数据 y , 就根据贝叶斯公式基于后验分布进行推断

$$\pi(\theta | y) = \frac{\pi(\theta)L(\theta)}{\int_{\Theta} \pi(\theta)L(\theta)d\theta} \propto \pi(\theta)L(\theta)$$

后验分布 \propto 先验分布 \times 似然函数.

- 👉 贝叶斯公式反映了人的学习过程.

$$\pi(\theta) \xrightarrow{y_1} \pi(\theta | y_1) \xrightarrow{y_2} \pi(\theta | y_1, y_2) \dots$$
$$\pi(\theta | y_1, y_2) = \pi(\theta)p(y_1 | \theta)p(y_2 | \theta) = \pi(\theta | y_1)p(y_2 | \theta).$$

先验的选择

常用先验分布

- 👉 主观先验: 基于领域专家的经验或历史数据;
- 👉 共轭先验;
- 👉 无信息 (客观) 先验:
 - 👉 均均先验: $\pi(\theta) \propto 1$.
 - 👉 Jeffreys 先验: $\pi(\theta) \propto \sqrt{I(\theta)}$.
 - 👉 reference 先验
 - 👉 概率匹配先验

1. 点估计

- 👉 后验均值: $\hat{\theta} = E(\theta | y) = \int_{\Theta} \theta \pi(\theta | y) d\theta$
- 👉 后验中位数: $\tilde{\theta} = P(\theta < \tilde{\theta} | y) = 0.5.$
- 👉 后验众数: $\check{\theta} = \max_{\theta \in \Theta} \pi(\theta | y).$

2. 区间估计

- 👉 等尾可信区间
- 👉 最高概率 (HPD) 可信区间

贝叶斯检验

👉 检验假设:

$$H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1$$

👉 支持 H_0 的后验概率:

$$\alpha^* = P(\theta \in \Theta_0 | y) = \int_{\Theta_0} \pi(\theta | y) d\theta$$

👉 若 α^* 足够小 (< 0.05), 就拒绝 H_0 .

👉 α^* 为拒绝 H_0 犯错误的概率.

贝叶斯预测

👉 数据: $y = (y_1, y_2, \dots, y_n)$, $y_i \sim p(y | \theta)$.

👉 y_{n+1} : 未来待预测观测值

👉 设 y_{n+1} 与 y_1, y_2, \dots, y_n 独立

👉 y_{n+1} 的后验预测分布:

$$P(y_{n+1}|y) = \int_{\Theta} p(y_{n+1}|\theta)\pi(\theta|y)d\theta.$$

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标**
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-尺度参数分布的统计推断
- 7 常用可靠性相关软件 R 程序包

失效率与浴盆曲线

典型的失效率函数为图 3.1 的曲线, 称为浴盆曲线.

- 👉 早期失效期: 开始时, $h(t)$ 处于逐渐下降的过程, 主要是由于设计错误、制造缺陷、工艺缺陷、管理不善等因素, 使一部分产品不可避免地存在一些潜在缺陷和弱点.
 - 🔺 对策: 我们可以通过筛选技术, 将早期失效的产品剔除.
- 👉 偶然失效期: 经过一段时间后, $h(t)$ 逐渐趋于平稳, 这时 $h(t)$ 基本上是一个常数.
- 👉 耗损失效期: 当产品使用到后期时, 由于耗损、老化等原因, 失效率又迅速上升.
 - 🔺 对策: 对耗损失效期的产品可采用维修、更换等手段.

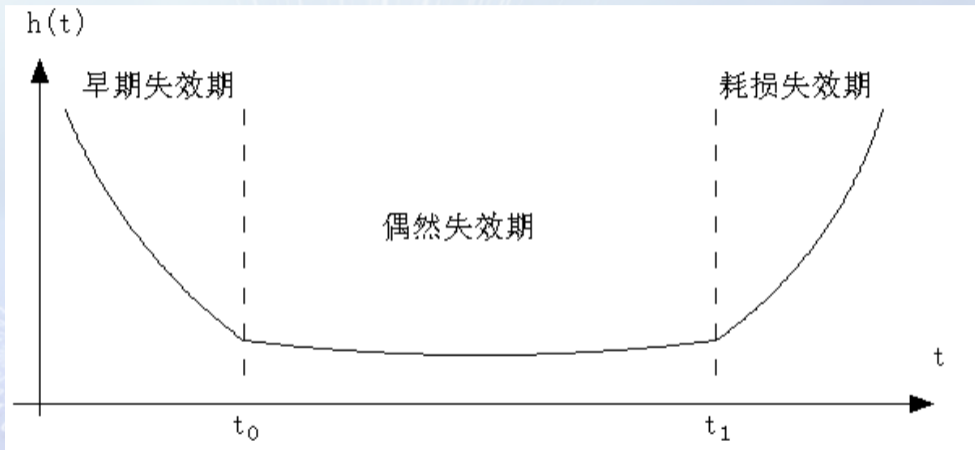


图 3.1: 浴盆曲线

可靠寿命

👉 可靠寿命: 使可靠度等于给定值 R 的时间 $t_{(R)}$ 称为可靠度为 R 的可靠寿命, 即由方程

$$R(t_{(R)}) = R \quad (3.1)$$

求得.

👉 当 $R = 0.5$ 时的可靠寿命 $t_{(0.5)}$ 称为中位寿命.

👉 当 $R = e^{-1}$ 时的可靠寿命称为特征寿命.

👉 $R = 0.9, 0.95$ 也是工程师最为关系的可靠寿命.

平均寿命

👉 产品的平均寿命为

$$E(T) = \int_0^{\infty} t dF(t) = \int_0^{\infty} R(t) dt. \quad (3.2)$$

👉 对于不可修产品来说平均寿命通常记为 **MTTF**.

👉 对于可修产品, 通常除了考虑首次故障前的平均时间, 还要研究平均无故障工作时间 (记为 **MTBF**).

平均剩余寿命

☞ 产品在工作到时刻 t 仍然正常的条件下, 它的剩余寿命分布为

$$F_t(x) = P[T - t \leq x | T > t] = \begin{cases} \frac{F(x+t) - F(t)}{1 - F(t)}, & \text{当 } x \geq 0 \\ 0, & \text{当 } x < 0 \end{cases}. \quad (3.3)$$

☞ 相应的可靠度函数为

$$R_t(x) = \frac{R(x+t)}{R(t)}, \quad \text{当 } x \geq 0. \quad (3.4)$$

☞ 剩余寿命分布 $F_t(x)$ 的均值称为**平均剩余寿命**, 计算公式如下:

$$m(t) = \int_0^{\infty} x dF_t(x) = \int_0^{\infty} R_t(x) dx = \frac{1}{R(t)} \left[E(T) - \int_0^t R(y) dy \right]. \quad (3.5)$$



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布**
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-尺度参数分布的统计推断
- 7 常用可靠性相关软件程序包

寿命分布表示 (1)

失效时间 (寿命) T 是一个非负随机变量, 通常是连续的.

T 的概率分布可以用下面任一函数刻画.

累积分布函数 (CDF)

概率密度函数 (pdf)

生存函数/可靠度函数 (sf)

失效率函数 (hf).

图 4.1 给出了一个典型失效时间分布的四个函数.

寿命分布表示 (2)

👉 累积分布函数 $F(t)$:

$$F(t) = P(T \leq t), \quad t \geq 0, \quad (4.1)$$

表示产品在时间 t 之前失效的概率 (比例). 在可靠性中, 分布函数 $F(t)$ 常称为失效分布函数或寿命分布函数 (简称失效分布或寿命分布).

👉 概率密度函数 $f(t)$:

$$f(t) = \frac{dF(t)}{dt}.$$

表示产品失效关于时间的相对频率.

寿命分布表示 (3)

- 👉 可靠度函数 $R(t)$: 产品在规定的条件下, 在规定的时间内 t 内完成规定功能的概率, 简称可靠度.

$$R(t) = P(X > t) = 1 - F(t) = \int_t^{\infty} f(x)dx. \quad (4.2)$$

它给出了产品在时刻 t 仍未失效 (生存着) 的概率.

- 👉 失效率函数 $h(t)$: 已工作到时刻 t 的产品, 在时刻 t 后单位时间内发生失效的概率称为产品在时间 t 的失效率函数, 简称失效率.

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{R(t)\Delta t} = \frac{f(t)}{R(t)}. \quad (4.3)$$

失效率函数表示产品在 t 之前还未失效的条件下, 在下一个小的时间段发生失效的倾向性.

示例: 指数分布

设某种产品的失效时间具有

☞ 累积分布函数: $F(t) = 1 - \exp(-t^{1.7})$

☞ 概率密度函数: $f(t) = F'(t) = 1.7t^{0.7} \exp(-t^{1.7})$

☞ 可靠度函数: $R(t) = 1 - F(t) = \exp(-t^{1.7})$

☞ 失效率函数: $h(t) = f(t)/R(t) = 1.7 \times t^{0.7}$

见图 4.1 (a)—(d) 所示.

示例: 指数分布

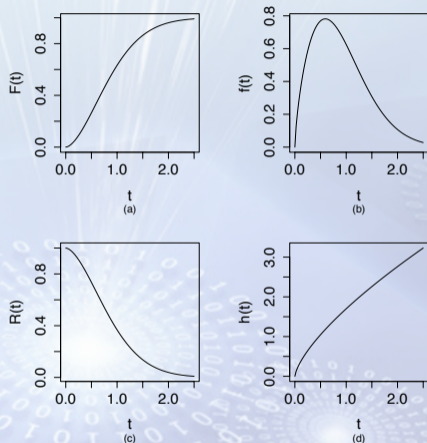


图 4.1: 失效时间 cdf,pdf,sf 和 hf

位置-刻度参数分布

$$F(y; \mu, \sigma) = \Phi\left(\frac{y - \mu}{\sigma}\right), \quad (4.4)$$

其中 $\Phi(\cdot)$ 不依赖于任何未知参数, $-\infty < \mu < \infty$ 为位置参数, $\sigma > 0$ 为刻度参数.

👉 位置-刻度参数分布在统计分析中占据了非常重要的地位。

特例

- 👉 指数分布
- 👉 对数正态分布 \Leftarrow 正态分布
- 👉 威布尔分布 \Leftarrow 极值分布
- 👉 对数 logistic 分布 \Leftarrow loglogistic 分布

双参数指数分布

☞ 寿命 T 服从二参数指数分布, 记为 $T \sim \text{EXP}(\theta, \gamma)$, 其 cdf, pdf 和 hf 分别为

$$F(t; \theta, \gamma) = 1 - \exp\left(-\frac{t - \gamma}{\theta}\right),$$

$$f(t; \theta, \gamma) = \frac{1}{\theta} \exp\left(-\frac{t - \gamma}{\theta}\right),$$

$$h(t; \theta, \gamma) = \frac{1}{\theta}, \quad t > \gamma,$$

其中 $\theta > 0$ 为尺度参数, γ 为位置参数, 又称为门限参数.

☞ 当 $\gamma = 0$ 时, 此分布变为单参数指数分布, 记为 $T \sim \text{EXP}(\theta)$.

☞ 图 4.2 给出了参数 $\gamma = 0, \theta = 0.5, 1.0, 2.0$ 时指数分布的 cdf, pdf 和 hf.

指数分布的 cdf, pdf 和 hf.

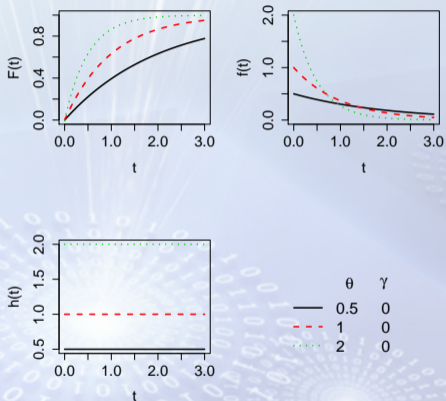


图 4.2: 指数分布的 cdf, pdf 和 hf.

对数正态分布

☞ 寿命 T 服从 (二参数) 对数正态分布, 记为 $T \sim \text{LOGNOR}(\mu, \sigma)$, 其 cdf 和 pdf 分别为

$$F(y; \mu, \sigma) = \Phi_{\text{nor}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$f(y; \mu, \sigma) = \frac{1}{\sigma t} \phi_{\text{nor}} \left(\frac{\log(t) - \mu}{\sigma} \right), \quad t > 0,$$

其中 $\phi_{\text{nor}}(\cdot)$ 和 $\Phi_{\text{nor}}(\cdot)$ 分别为标准正态分布的 pdf 和 cdf.

☞ $Y = \log(T) \sim \text{NOR}(\mu, \sigma)$.

☞ 图 4.3 给出了参数 $\mu = 0, \sigma = 0.3, 0.5, 0.8$ 时对数正态分布的 cdf, pdf 和 hf.

对数正态分布的 cdf, pdf 和 hf

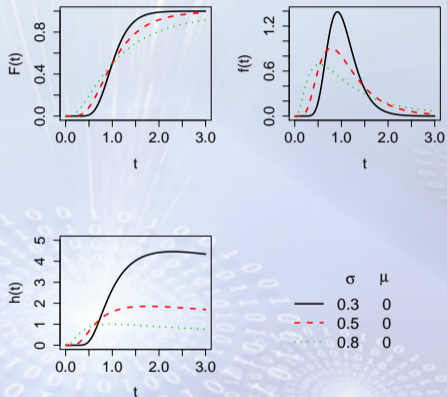


图 4.3: 对数正态分布的 cdf, pdf 和 hf.

威布尔分布

☞ 寿命 T 服从形状参数为 β , 尺度参数为 η 的 (二参数) 威布尔分布 (记为 $T \sim \text{WEIB}_1(\eta, \beta)$), 其 cdf, pdf 和 hdf 分别为

$$F(t; \eta, \beta) = 1 - \exp \left[- \left(\frac{t}{\eta} \right)^\beta \right],$$

$$f(t; \eta, \beta) = \beta \frac{t^{\beta-1}}{\eta^\beta} \exp \left[- \left(\frac{t}{\eta} \right)^\beta \right],$$

$$hf(t; \eta, \beta) = \beta \frac{t^{\beta-1}}{\eta^\beta}, \quad t > 0,$$

其中 $\beta > 0$ 为形状参数, $\eta > 0$ 为尺度参数.

☞ 图 4.4 给出了参数 $\eta = 1, \beta = 0.8, 1.0, 1.5$ 时威布尔分布的 cdf, pdf 和 hf.

威布尔分布的 cdf, pdf 和 hf.

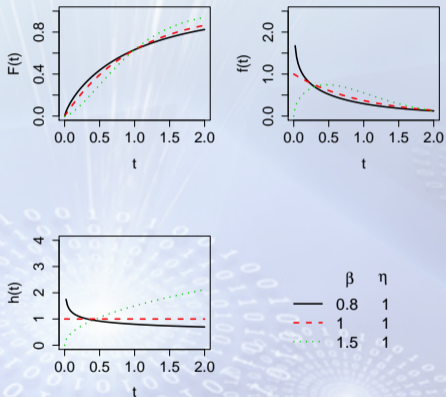


图 4.4: 威布尔分布的 cdf, pdf 和 hf.

威布尔分布与极值分布的关系

☞ 威布尔分布能描述经常出现的失效分布的形态.

▢ 当 $\beta < 1$ 时, 其 hf 是递减的;

▢ 当 $\beta > 1$ 时, 其 hf 是递增的;

▢ 当 $\beta = 1$ 时, 威布尔分布变成刻度参数为 $\theta = \eta$ 的指数分布.

☞ 如果 $T \sim \text{WEIB}_1(\eta, \beta)$, 则 $\log(T) \sim \text{SEV}(\mu, \sigma)$ (最小极值分布), 其中 $\sigma = 1/\beta$ 和 $\mu = \log(\eta)$ 为刻度与位置参数, 记为 $T \sim \text{WEIB}_2(\mu, \sigma)$, 相应的 cdf, pdf 和 hf 可分别表示为

$$F(y; \mu, \sigma) = \Phi_{\text{sev}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$f(y; \mu, \sigma) = \frac{1}{\sigma t} \phi_{\text{sev}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$hf(y; \mu, \sigma) = \frac{1}{\sigma \exp(\mu)} \left[\frac{t}{\exp(\mu)} \right]^{1/\sigma - 1}, \quad t > 0,$$

对数 Logistic 分布

☞ 随机变量 T 服从对数 Logistic 分布, 记为 $T \sim \text{LOGIS}(\mu, \sigma)$, 其 cdf, pdf 和 hf 分别为

$$F(t; \mu, \sigma) = \Phi_{\text{logis}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$f(t; \mu, \sigma) = \frac{1}{\sigma t} \phi_{\text{logis}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$hf(t; \mu, \sigma) = \frac{1}{\sigma t} \Phi_{\text{logis}} \left(\frac{\log(t) - \mu}{\sigma} \right), \quad t > 0,$$

其中 $\Phi_{\text{logis}}(z)$ 和 $\phi_{\text{logis}}(z)$ 分别为标准 Logistic 分布 $\text{LOGIS}(0, 1)$ 的 cdf 和 pdf, 分布中的参数 μ 和 σ 是 $\log(T)$ 分布的位置参数和刻度参数, 因为 $Y = \log(T) \sim \text{LOGIS}(\mu, \sigma)$.

☞ 图 4.5 给出了参数 $\mu = 0, \sigma = 0.2, 0.4, 0.6$ 时对数 Logistic 分布的 cdf, pdf, hf.

对数 Logistic 分布的 cdf, pdf 和 hf

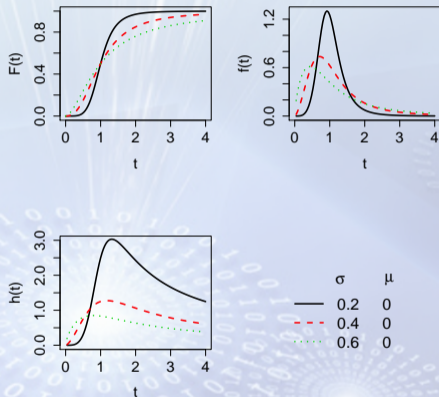


图 4.5: $\mu = 0, \sigma = 0.2, 0.4, 0.6$ 时对数 Logistic 分布的 cdf, pdf 和 hf.



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)**
- 6 位置-尺度参数分布的统计推断
- 7 常用可靠性相关软件程序包

非参数估计方法概述 (Meeker, 2022)

☞ 图估计法

☞ $F(t)$ 的非参数区间估计:

- ☞ 单侧截尾区间数据
- ☞ 单侧截尾精确失效数据
- ☞ 区间数据与多重截尾数据
- ☞ 精确失效数据

☞ 可靠度函数 $R(t)$ 的非参数区间估计 (任意截尾场合):

- ☞ Peto/Turnbull 估计

☞ 累积风险函数 $\Lambda(t) = -\log(R(t))$ 的非参数区间估计:

- ☞ Nelson-Aalen 估计

图估计方法

- 通过概率图的方法对寿命分布进行直观的推断，获得寿命分布的失效概率、可靠度、分位数等特征量的图估计。
- 概率图主要用在位置-刻度参数模型中，包括常用的指数分布、威布尔分布和对数正态分布等。
- 图估计法的作用 (Meeker, 2022)
 - 评估寿命分布模型的推断精度
 - 识别多种失效模型或多个分布的混合
 - 通过概率图上拟合的直线获得模型参数的估计
 - 展示寿命分布极大似然估计对数据的拟合效果
 - 通过数据点的光滑拟合获得失效概率与分位数的半参数估计

图估计法的思想

- 👉 给定的寿命分布 $F(t)$, 通过对 $(t, F(t))$ 的线性化变换后具有线性关系, 相应的数据点 $(t_i, \widehat{F}(t_i))$ 在概率图上落在一条直线上.
- 👉 横轴: 时间轴, 用原来的线性轴或对数轴, 通常为分布的分位数或对数分位数;
- 👉 纵轴: 概率轴, 用双轴区分:
 - 🔺 右侧: 画出标准分布 $(\widehat{F}_0(t))$ 的分位数
 - 🔺 左侧: 画出相应的累积失效概率 $\widehat{F}(t)$

示例: 位置-刻度参数分布 (全样本)

👉 位置-刻度参数分布:

$$F(t) = F_0\left(\frac{t - \mu}{\sigma}\right).$$

👉 设 $x_{(1)} < x_{(2)} < \dots < x_{(n)}$ 为来自 $F(t)$ 的容量为 n 的次序统计量。

🔺 在 $y_i = F^{-1}(a_i)$ 处描点 $(x_{(i)}, y_i), i = 1, 2, \dots, n$, 其中 a_i 为 $F_0((x_{(i)} - \mu)/\sigma)$ 的一个估计, 通常取为经验分布估计, 此图即为概率图。

🔺 如果寿命分布模型是正确的, 那么 $(x_{(i)}, y_i), i = 1, 2, \dots, n$ 应该落在直线 $x = \mu + \sigma y$ 附近。

🔺 从概率图上读取截距与斜率就得到参数 μ 和 σ 的 (图) 估计。

两参数指数分布的概率图

由

$$p = F(t; \theta, \gamma) = 1 - \exp\left[-\frac{(t - \gamma)}{\theta}\right],$$

$$t \geq \gamma$$

得 $t_p = \gamma - \theta \log(1 - p)$.

$(t_p, -\log(1 - p))$ 落在一条直线上.

γ 是横轴 (时间轴) 上的截距

θ 是斜率 ($-\log(1 - p)$ 变动一个单位对应的值)

Plot with Exponential Distribution Probability Scales
Showing Exponential cdfs as Straight Lines for
Combinations of Parameters $\theta = 100, 300$ and $\gamma = 0, 200$

$$t_p = \gamma - \theta \log(1 - p)$$

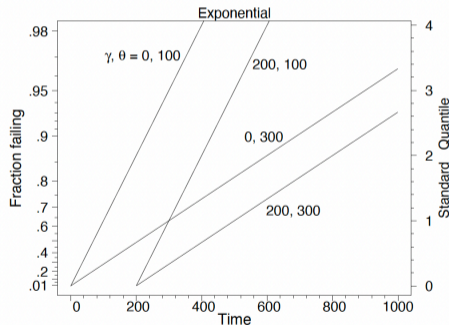


图 5.1: 指数分布概率图

正态分布的概率图

由

$$p = F(y; \mu, \sigma) = \Phi_{\text{nor}} \left(\frac{y - \mu}{\sigma} \right),$$

$$-\infty < y < \infty$$

得 $y_p = \mu + \sigma \Phi_{\text{nor}}^{-1}(p)$.

$(y_p, \Phi^{-1}(p))$ 落在一条直线上.

- ▢ μ 是横轴 (时间) 上的截距
- ▢ σ 是斜率 ($\Phi^{-1}(p)$ 变动一个单位对应的值)

Plot with Normal Distribution Probability Scales
Showing Normal distribution cdfs as Straight Lines for
Combinations of Parameters $\mu = 40, 80$ and $\sigma = 5, 10$
 $y_p = \mu + \sigma \Phi_{\text{norm}}^{-1}(p)$

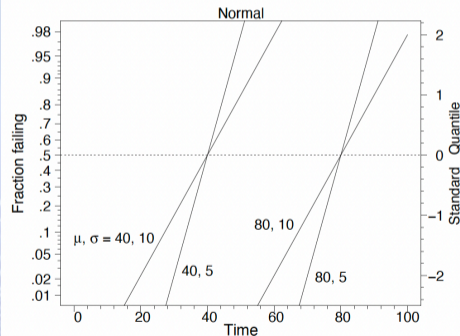


图 5.2: 正态分布概率图

对数正态分布的概率图

由

$$p = F(t; \mu, \sigma) = \Phi_{\text{nor}} \left(\frac{\log(t) - \mu}{\sigma} \right),$$

$$-\infty < y < \infty$$

得 $t_p = \exp[\mu + \sigma\Phi_{\text{nor}}^{-1}(p)]$.

$(\log(t_p), \Phi^{-1}(p))$ 落在一条直线上。

- μ 是横轴 (对数时间轴) 上的截距
- σ 是斜率 ($\Phi^{-1}(p)$ 变动一个单位对应的值)

Plot with Lognormal Distribution Probability Scales
Showing Lognormal Distribution cdfs as Straight Lines
for Combinations of $\exp(\mu) = 50, 500$ and $\sigma = 1, 2$

$$\log(t_p) = \mu + \sigma\Phi_{\text{norm}}^{-1}(p)$$

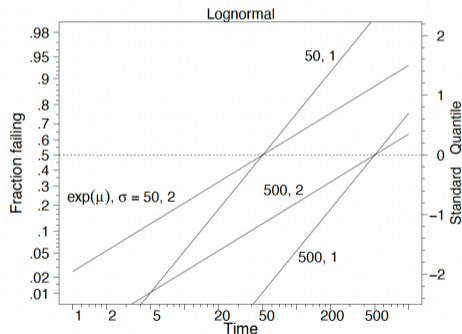


图 5.3: 对数正态分布概率图

威布尔分布的概率图



由

$$p = F(t; \mu, \sigma) = \Phi_{\text{sev}} \left[\frac{\log(t) - \mu}{\sigma} \right], t > 0$$

得 $t_p = \exp [\mu + \sigma \Phi_{\text{sev}}^{-1}(p)] =$

$\eta [-\log(1-p)]^{1/\beta}$, 其中

$\eta = \exp(\mu), \beta = 1/\sigma$.

$(\log(t_p), \log[-\log(1-p)])$ 落在一条直线上.

$\mu = \log(\eta)$ 是横轴 (对数时间轴) 上的截距

$\beta = 1/\sigma$ 是斜率 $(\log[-\log(1-p)])$ 变动一个单位对应的值)

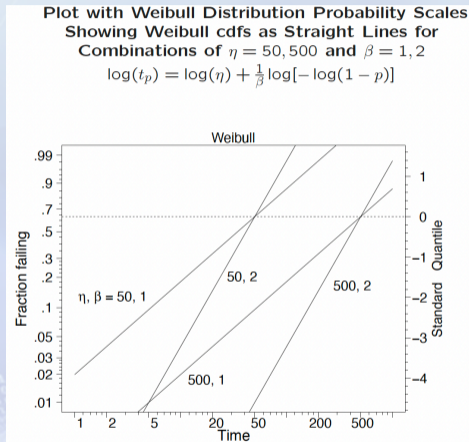


图 5.4: Weibull 分布概率图

示例: 喷气发动机引气系统故障 (Meeker, 2022)

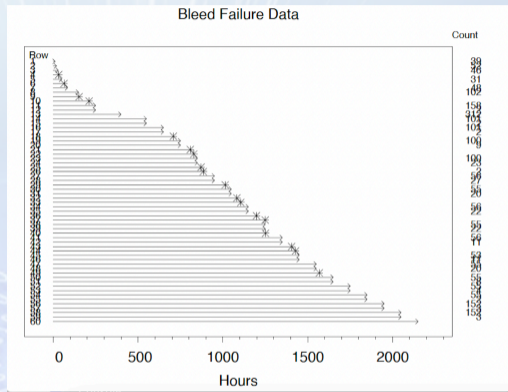


图 5.5: 系统故障事件图; 2256 个现场系统; 多重删失

示例 (续): 所有基地故障数据的 Weibull 分布概率图

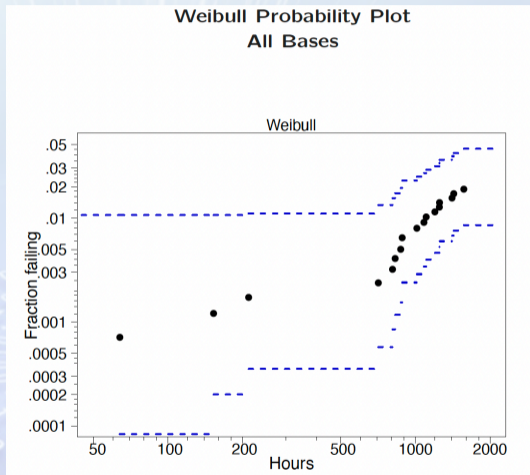


图 5.6: Weibull 分布概率图—所有基地数据

示例 (续): 两类基地数据分开的 Weibull 分布概率图

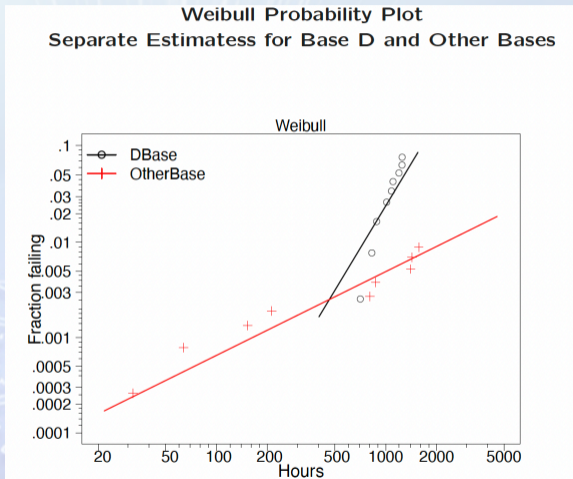


图 5.7: 两类基地数据分开的混合 Weibull 分布概率图



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-刻度参数分布的统计推断**
- 7 可靠性统计的R语言程序包

Weibull 分布的极大似然函数: 右截尾数据

☞ CDF

$$\Pr(T \leq t) = F(t; \mu, \sigma) = \Phi_{\text{sev}}\{[\log(t) - \mu]/\sigma\}.$$

☞ 似然函数

$$\begin{aligned} L(\mu, \sigma) &= \prod_{i=1}^n L_i(\mu, \sigma; \text{data}_i) \\ &= \prod_{i=1}^n [f(t_i; \mu, \sigma)]^{\delta_i} [1 - F(t_i; \mu, \sigma)]^{1-\delta_i} \\ &= \prod_{i=1}^n \left[\frac{1}{\sigma t_i} \phi_{\text{sev}}\left(\frac{\log(t_i) - \mu}{\sigma}\right) \right]^{\delta_i} \times \left[1 - \Phi_{\text{sev}}\left(\frac{\log(t_i) - \mu}{\sigma}\right) \right]^{1-\delta_i}, \end{aligned}$$

其中 $\delta_i = 1$ 表示 t_i 为精确失效; $\delta_i = 0$ 表示 t_i 为右截尾时间, $\phi_{\text{Sev}}(z)$ 和 $\Phi_{\text{Sev}}(z)$ 为标准极小值分布的 pdf 和 CDF.

对数正态分布的极大似然函数: 右截尾数据

☞ CDF

$$\Pr(T \leq t) = F(t; \mu, \sigma) = \Phi_{\text{norm}} \{[\log(t) - \mu]/\sigma\}.$$

☞ 似然函数

$$\begin{aligned} L(\mu, \sigma) &= \prod_{i=1}^n L_i(\mu, \sigma; \text{data } i) \\ &= \prod_{i=1}^n [f(t_i; \mu, \sigma)]^{\delta_i} [1 - F(t_i; \mu, \sigma)]^{1-\delta_i} \\ &= \prod_{i=1}^n \left[\frac{1}{\sigma t_i} \phi_{\text{norm}} \left(\frac{\log(t_i) - \mu}{\sigma} \right) \right]^{\delta_i} \times \left[1 - \Phi_{\text{norm}} \left(\frac{\log(t_i) - \mu}{\sigma} \right) \right]^{1-\delta_i} \end{aligned}$$

其中 $\phi_{\text{norm}}(z)$ 和 $\Phi_{\text{norm}}(z)$ 为标准极小值分布的 pdf 和 CDF.

Example: 减震器数据

Distance (km)	Failure Mode	Distance (km)	Failure Mode
6,700	Mode1	17,520	Mode1
6,950	None	17,540	None
7,820	None	17,890	None
8,790	None	18,450	None
9,120	Mode2	18,960	None
9,660	None	18,980	None
9,820	None	19,410	None
11,310	None	20,100	Mode2
11,690	None	20,100	None
11,850	None	20,150	None
11,880	None	20,320	None
12,140	None	20,900	Mode2
12,200	Mode1	22,700	Mode1
12,870	None	23,490	None
13,150	Mode2	26,510	Mode1
13,330	None	27,410	None
13,470	None	27,490	Mode1
14,040	None	27,890	None
14,300	Mode1	28,100	None

Data from O'Connor (1985, page 85).

👉 $R(\mu, \log(\sigma)) = L(\mu, \log(\sigma)) / L(\hat{\mu}, \log(\hat{\sigma})), \hat{\mu} = 10.23, \hat{\sigma} = 0.3164$

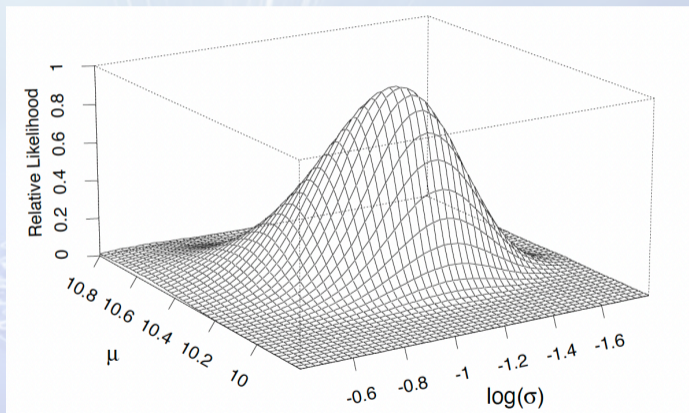


图 6.1: 相对似然—三维图

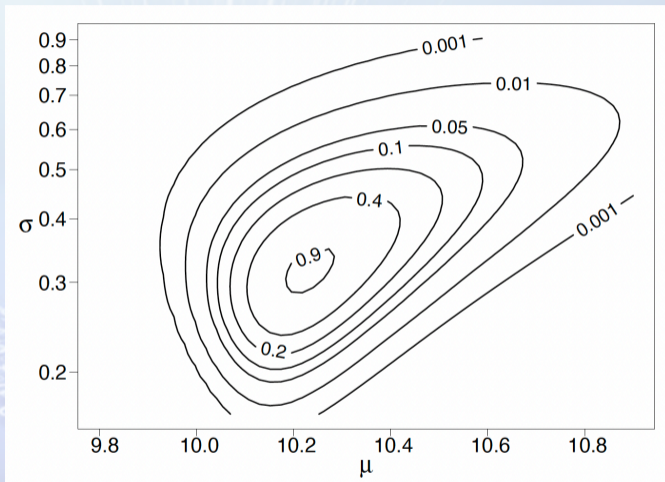


图 6.2: 相对似然—三维图

F(t) 的近似 (Wald) 逐点置信区间

$$\left[\underset{\sim}{F}(t_e), \tilde{F}(t_e) \right] = \widehat{F}(t_e) \mp z_{(1-\alpha/2)} \text{se}_{\widehat{F}},$$

其中

$$\widehat{F}(t_e) = F(t_e; \hat{\mu}, \hat{\sigma}) = \Phi(\hat{z}_e),$$

$$\hat{z}_e = [\log(t_e) - \hat{\mu}] / \hat{\sigma},$$

$$\text{se}_{\widehat{F}} = \frac{\phi(\hat{z}_e)}{\hat{\sigma}} \left[\widehat{\text{Var}}(\hat{\mu}) + 2\hat{z}_e \widehat{\text{Cov}}(\hat{\mu}, \hat{\sigma}) + \hat{z}_e^2 \widehat{\text{Var}}(\hat{\sigma}) \right]^{1/2}.$$

近似 (Wald) 逐点置信区间: 减震器数据

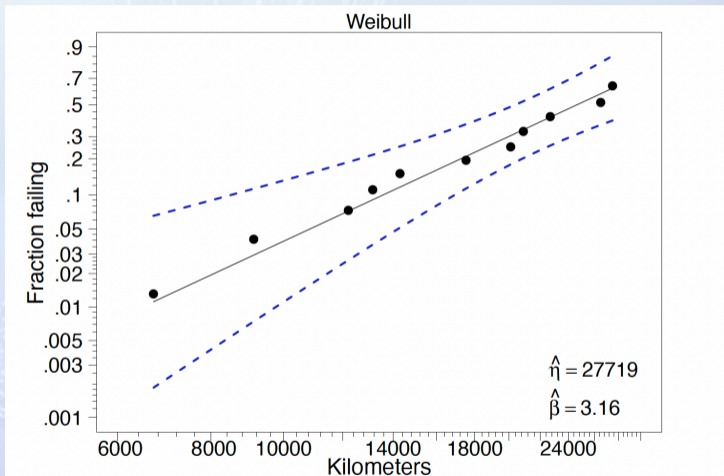


图 6.3: $F(t)$ 的 Wald 逐点置信区间

Weibull 分布与对数正态分布的比较: 减震器数据

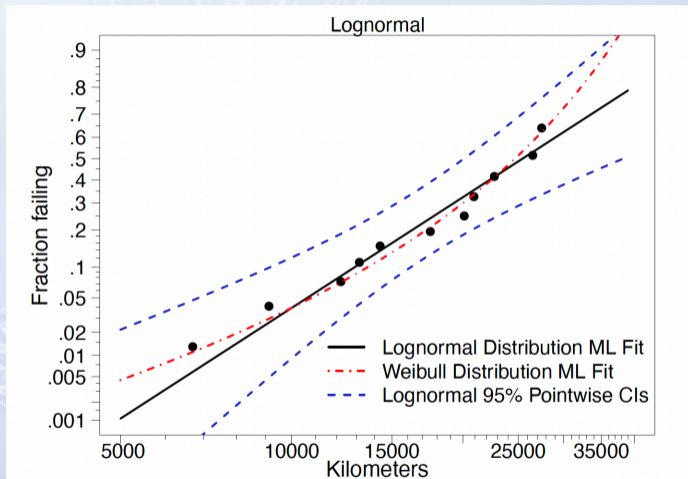



图 6.4: $F(t)$ 的 Wald 逐点置信区间—两种分布

 $R(\mu, \sigma) > \exp \left[-\chi^2_{(1-\alpha;2)}/2 \right] = \alpha.$

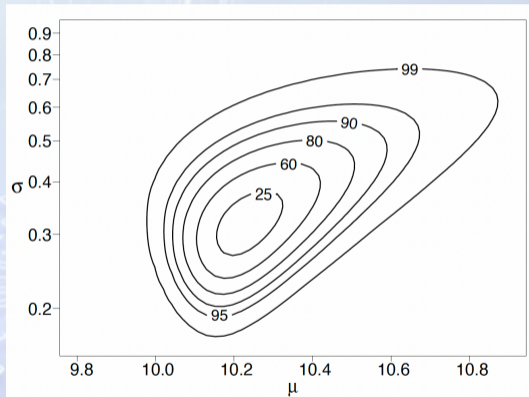


图 6.5: 参数的置信带

μ 的置信区间: $R(\mu)$ 的剖面似然

$$\text{👉 } R(\mu) = \max_{\sigma} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right].$$

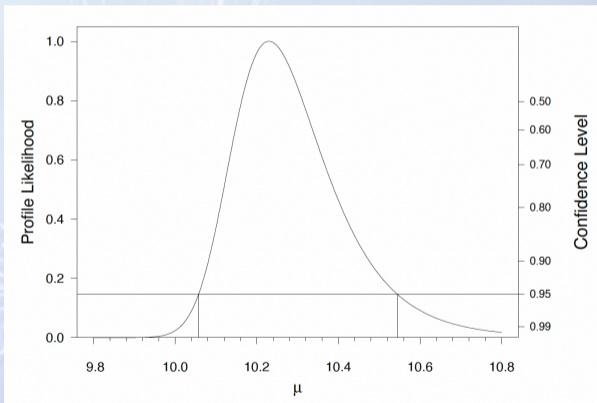


图 6.6: 基于剖面似然的参数置信区间, $R(\mu)$

σ 的置信区间: $R(\sigma)$ 的剖面似然

$$\text{👉 } R(\sigma) = \max_{\mu} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right].$$

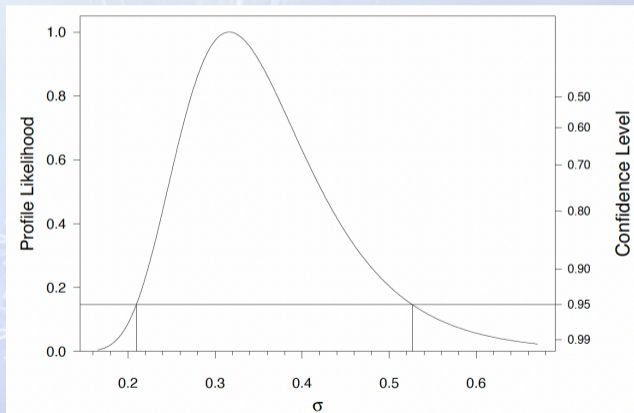


图 6.7: 基于剖面似然的参数置信区间, $R(\sigma)$



提纲

- 1 可靠性统计中的基本概念
- 2 统计推断方法简介
- 3 常用的可靠性指标
- 4 常用的寿命分布
- 5 可靠性中的非参数统计推断 (图估计)
- 6 位置-尺度参数分布的统计推断
- 7 几个与可靠性相关的 R 程序包

几个与可靠性相关的 R 程序包

👉 分布检验与函数

- 1 EWGoF: 全样本场合指数分布与二参数威布尔分布的检验与估计
- 2 weibullness: 威布分布的检验

👉 分布的拟合/极大似然估计

- 1 fitdistr/MASS: 一元分布的极大似然估计
- 2 mle/stats4: 一元分布的极大似然估计
- 3 lifetime.mle/SPREDA: 寿命分布的极大似然估计
- 4 lifedata.MLE/SPREDA: 位置-刻度分布的似然拟合
- 5 summary.Lifedata.MLE/SPREDA: 汇总统计量中的极大似然估计

👉 两个强大的 R 包

- 🔗 weibullR: Weibull 分析, 共 18 个函数
- 🔗 weibulltools: 寿命数据统计分析方法, 包括图估计、非参数估计及参数估计

谢谢!
Thank you!