# 基于记录值的可靠性分布模型的统计推断研究

JIYU JILUZHI DE
KEKAOXING FENBU MOXING DE
TONGJI TUIDUAN YANJIU

周慧 任海平 著

冶金工业出版社
www.cnmip.com.cn

Metallurgical Industry Press
冶金工业出版社

基于记录值的可靠性分布模型的统计推断研究

体验更多精彩阅读
尽在冶金工业出版社微信平台

ISBN 978-7-5024-8287-9
9 787502 482879 >
定价45.00元
销售分类建议：统计学

# 基于记录值的可靠性分布模型的统计推断研究

周慧　任海平　著

北京
冶金工业出版社
2019

## 内 容 提 要

本书研究了基于记录值样本的可靠性分布模型参数的 Bayes 统计推断问题和基于记录值样本的产品寿命绩效指标的统计推断问题。研究成果可以为工程师在处理含有记录值数据信息的产品性能评估问题时提供决策参考。

本书可供从事统计分析、预测、评估工作的工程技术人员阅读，也可供高等学校相关专业师生参考。

**图书在版编目(CIP)数据**

基于记录值的可靠性分布模型的统计推断研究／周慧，任海平著．—北京：冶金工业出版社，2019. 12
ISBN 978-7-5024-8287-9

Ⅰ.①基…　Ⅱ.①周…　②任…　Ⅲ.①统计推断—研究
Ⅳ.①O212

中国版本图书馆 CIP 数据核字(2019)第 287951 号

出 版 人　陈玉千
地　　址　北京市东城区嵩祝院北巷 39 号　邮编　100009　电话　(010)64027926
网　　址　www. cnmip. com. cn　电子信箱　yjcbs@ cnmip. com. cn
责任编辑　杨　敏　美术编辑　彭子赫　版式设计　禹　蕊
责任校对　石　静　责任印制　李玉山
ISBN 978-7-5024-8287-9
冶金工业出版社出版发行；各地新华书店经销；三河市双峰印刷装订有限公司印刷
2019 年 12 月第 1 版，2019 年 12 月第 1 次印刷
169mm × 239mm；7. 5 印张；151 千字；109 页
**45. 00** 元

**冶金工业出版社　投稿电话　(010)64027932　投稿信箱　tougao@ cnmip. com. cn**
**冶金工业出版社营销中心　电话　(010)64044283　传真　(010)64027893**
**冶金工业出版社天猫旗舰店　yjgycbs. tmall. com**

# 前　言

记录值是刻画随机变量序列变化趋势的一个重要的数值，自被提出以来在遗传学、气候学、水文学、地震、保险精算等诸多领域得到广泛的应用。例如在保险业中，索赔额序列通常被假定服从某个重尾分布的正值独立同分布的随机变量序列，大额索赔的发生规律是破产理论的重要研究内容之一，其中包括对记录值分布规律的研究；在气象学中研究降雨（雪）量，我们可以由到目前为止所得到的测量值（记录值）来预测未来的降雨（雪）量等。因此，研究记录值的变化趋势以及统计推断理论，对于国民经济的发展具有重要意义。

对记录值的研究，引起很多学者的兴趣，国内主要是研究基于特定分布族的概率极限理论和随机比较问题，国外研究得较多，例如记录值和记录时间的分布函数、记录值的矩之间的关系，以及利用记录值进行社会调查、地震预测等研究。利用记录值进行模型参数的统计推断理论，最近十几年引起国外很多学者的兴趣，但大多是在经典统计理论框架下进行研究的。应用 Bayes 和经验 Bayes 统计推断理论研究记录值模型的相关文献还不多，国内也没有专门的专著出版。寿命绩效指标作为一类重要的望大型过程能力指标，在质量控制领域有着重要的应用，但关于其统计推断研究的中文文献也非常少，同时研究得也不够深入，为此本专著一方面进一步将 Bayes 统计方法应用到记录值模型，另一方面探讨基于记录值的产品寿命绩效指标的统计

推断问题，以便为管理者进行决策时提供参考。同时也希望能够通过本著作吸引更多国内学者参与到记录值模型的统计推断理论与应用研究中。

本专著的第1章、第2章和第5章由周慧撰写，第3章、第4章由周慧和任海平撰写，全书由周慧定稿。

感谢宜春学院及宜春学院数学与计算机科学学院的领导和同事在本书出版过程中给予的帮助和支持。本书内容涉及的研究得到江西省教育厅科学研究项目（GJJ180829）、宜春一流学科“管理科学与工程”、国家自然科学基金（71661012、11561068、11771382）的资助，另外，在撰写过程中参考了有关文献，在此一并表示感谢。

由于作者水平有限，书中不当之处在所难免，敬请读者批评指正。

作　者

2019年5月

# 目 录

# 1 绪 论

## 1.1 研究背景及意义

记录值是刻划随机变量序列变化趋势的一个重要的数值，它可当作顺序统计量的一种，而且记录值又依照不同的选取方式可分为上记录值（upper record values）与下记录值（lower record values），其定义最早由 Chandler（1952）给出。记录值的定义可以简要描述如下[1]：假设 $X_1, X_2, \cdots$ 为一组来自独立同分布的随机变量组成的无穷序列，其概率密度和累积分布函数分别为 $f(x)$ 与 $F(x)$。一观测值 $X_k$ 若其值高于（低于）在其之前的观测值，则称 $X_k$ 为上记录值（下记录值），或简称记录值。设 $X_1, X_2, \cdots, X_n, \cdots$ 为互相独立且来自相同分布的随机样本序列，则定义 $X_{U(1)}, X_{U(2)}, \cdots, X_{U(n)}$ 为前 $n$ 个上记录值样本，其中假设发生这些上记录值的时间为：

$$U(1)=1,\ U(n)=\min\{j \mid j>U(n-1),\ X_j>X_{U(n-1)}\},\qquad n\geqslant 2$$

另外，也定义 $X_{L(1)}, X_{L(2)}, \cdots, X_{L(n)}$ 为前 $n$ 个下记录值样本，其中这些下记录值发生的时间分别为 $L(1)=1$，$L(n)=\min\{j \mid j>U(n-1), X_j<X_{U(n-1)}\}$，$n\geqslant 2$。例如，若观察到观测值序列为 $\{3, 2, 2.5, 2.6, 1, 3.7, 2.2, 5.4, 2.7, 6.8, 0.5, \cdots\}$，则 $X_{U(1)}=3, X_{U(2)}=3.7, X_{U(3)}=5.4, X_{U(4)}=6.8$ 为前 4 个上记录值样本；而 $X_{L(1)}=3$，$X_{L(2)}=2, X_{L(3)}=1, X_{L(4)}=0.5$。又如碎石机辗碎的石头比在其之前被辗碎的石头还要大时，就要重新调整碎石机。下述资料是从一开始到第三次调整碎石机这段期间石头的大小[2]：

9.3, 0.6, 24.4, 18.1, 6.6, 9.0, 14.3, 6.6, 13.0, 2.4, 5.6, 33.8

如果只在需要调整碎石机时才记录石头的大小，此时得到上记录值样本为 9.3、24.4、33.8。

在现实生活中，常常出现记录值，如工业上压力的测试、气温的变化、运动竞赛成绩、经济学以及寿命检测等，其中最典型的例子就是金氏世界纪录。如前所述，如果碎石机辗碎的石头比在其之前被辗碎的石头还要大时，就要重新调整碎石机，因此，只有在需要调整碎石机时才记录石头的大小，此时得到上记录值（upper record values）样本；又或者对产品进行寿命检测实验时，实验者只在产品的寿命比在其之前发生故障的产品长时，才记录其寿命资料，当统计人员要分析资料时，只有上记录值样本可以分析。如上所述，记录值的应用十分广泛，已

被广泛应用到诸如水文学、气候学、地震、保险精算、机械工程以及体育等领域[3~7]。如在保险业中，通常假定索赔额序列是服从某个重尾分布的正值独立同分布的随机变量序列，根据破产理论，导致保险公司破产的往往是那些以小概率发生的大额索赔，因此，大额索赔的发生规律是破产理论的重要研究内容之一，其中包括对记录值分布规律的研究；在气象学中研究降雨（雪）量，可以由到目前为止所得到的测量值（记录值）来预测未来的降雨（雪）量等。因此研究记录值的变化趋势以及统计推断理论，对于国民经济的发展具有重要意义。

## 1.2 国内外研究现状

对记录值的研究，国内主要是研究基于特定分布族的概率极限理论和随机比较问题，国外研究的较多，例如记录值和记录时间的分布函数、记录值的矩之间的关系等，以及利用记录值进行社会调查、地震预测等研究。利用记录值进行模型参数的统计推断理论，最近十几年引起国外很多学者的兴趣。

在国内，基于记录值样本，胡治水等[8]研究了对数正态总体的记录值的部分和序列的渐近分布问题。苏淳等[9]研究了 Logistic 分布记录值序列部分和中心极限定理，并指出这一工作不仅具有概率论的极限理论方面的研究价值，而且在金融、保险等领域也具有相当重要的应用前景。王琪和黄文宜[10]讨论了广义指数分布参数的 Bayes 估计问题，在贝塔先验分布和平方误差损失、LINEX 损失函数下，导出了参数的 Bayes 和经验 Bayes 估计。王亮等[11]在对称和非对称损失函数下讨论了 Burr Ⅻ分布可靠性指标的 Bayes 估计问题，并针对超参数未知情形给出了一种确定超参数值的新方法。王琪和任海平[12]在平方误差损失函数下研究了 Rayleigh 分布参数的最小风险同变估计和 Bayes 估计，并讨论了一类线性形式估计的可容许性问题。高小琪和韦程东[13]研究了两参数指数威布尔分布的参数估计问题，基于记录值样本得到了参数的最大似然估计并在不同损失函数下推导了参数的 Bayes 估计，最后通过蒙特卡洛模拟进行比较，发现在合适的先验下 Bayes 估计比最大似然估计的精度更高。邢建平[14]研究了熵损失函数下指数分布参数的最小风险同变估计、Bayes 估计和经验 Bayes 估计。黄文宜[15]在一类加权损失函数下研究了几何分布可靠度参数估计问题，首先导出了可靠度的最大似然估计和最小方差无偏估计；其次在可靠度的贝塔先验分布的假定下导出了可靠度参数的 Bayes 估计，最后通过统计模拟试验发现加权损失函数下得到的 Bayes 估计的结果更精确。王亮和师义民[16]在平衡损失函数下研究了 Cox 模型参数的 Bayes 统计推断问题以及未来失效样本的预测问题。通过 ML-Ⅱ方法估计超参数，导出了相关可靠性指标的经验 Bayes 估计，并给出未来记录值样本的经验 Bayes 预测区间。韩雪和张青楠[17]研究了广义 Logistic 分布的位置参数和尺度参数的最佳线性无偏估计以及预测问题，并通过蒙特卡洛统计模拟发现，当样本容量不同

且形状参数也取不同值时，参数估计和预测的偏差和均方误差都较小，表明所提出的方法是实用和有效的。罗嘉成和陈勇明[18]讨论了极值分布参数的极大似然估计和逆矩估计量并利用卡方分布构造尺度参数的准确置信区间，并进一步研究了尺度和位置参数的联合置信域问题。

在国外，基于记录值样本，Nadar 和 Kizilaslan[19]研究了 Kumaraswamy 分布的应力强度干涉模型可靠度 $P(X<Y)$ 的 Bayes 估计问题，在假定参数的先验分布为共轭先验分布和无信息先验分布，损失函数为平方误差和 LINEX 损失函数下得到了可靠度的 Bayes 估计。Solimana[20]在平方误差损失、LINEX 损失和熵损失函数下，研究了 Rayleigh 分布的一些寿命参数，如可靠性和危险率函数的 Bayes 估计问题。Barakat 和 Elgawad[21]研究了上记录值和下记录值的极限分布问题，得到了弱收敛的充分条件。Qomi 和 Kiapour[22]研究了指数分布参数的最短置信区间估计问题。Mirmostafaee 等[23]基于 k-记录值研究了 Topp-Leone 分布下的形状参数、生存函数和危险率函数的 Bayes 点估计和 Bayes 区间估计问题。在对称和非对称损失函数下得到了 Bayes 估计，并讨论了应力 - 强度干涉模型可靠度参数的 Bayes 估计。Khan 和 Arshad[24]研究了比例危险率分布族的可靠度函数和应力 - 强度干涉模型可靠度的一致最小方差无偏（UMVU）估计，并对几种特殊的比例危险率模型——幂函数分布、指数化威布尔分布、广义指数分布、广义瑞利分布和 Topp-Leone 分布的估计结果进行了简化，最后通过仿真比较了 UMVU 估计量和最大似然估计量。更多关于记录值的统计推断研究可参见文献［25～30］。

但大多文献在研究基于记录值样本的统计推断问题时，都是在经典统计理论框架下进行研究的。Houchens[31]指出对于含有记录值的模型的统计推断问题的研究，由于获得的记录值的样本量通常很小，因而 Bayes 方法是一个很好的选择。最近基于记录值模型参数的 Bayes 估计问题引起了国外很多学者的兴趣，但大多数 Bayes 推断程序都是在平方损失函数下讨论，由于在估计可靠性及失效率函数时，高估会比低估带来更严重的后果，因此在这种情况下选择对称损失函数是不合实际的。且基于记录值样本数据的 Bayes 统计推断研究还只局限于平方误差损失函数、LINEX 损失、熵损失、对称熵等常见损失函数，在其他一些损失函数，如刻度平方误差损失、预警损失（Precautionary Loss）、有界误差损失函数等的 Bayes 统计推断的研究结果还很少，研究得也不够深入。

另外，最近十多年来，一类特殊的过程能力指数——寿命绩效指标[31]的统计推断问题得到了很多学者的关注[32～41]。基于记录值样本的产品寿命绩效指标的统计推断问题也引起了部分学者的兴趣。如基于记录值样本，Wu 等[42,43]在产品寿命分别服从指数分布和 Burr Ⅻ分布条件下，探讨了寿命绩效指标的一致最小方差不变估计量（uniformly minimum variance unbiased estimator，UMVUE），并提出了一套检验产品寿命性能的检验程序用以评估具有上记录值寿命数据的产品

寿命性能。Lee 等[44]在产品寿命服从 Weibull 分布的假定下，利用数据转换技巧导出了寿命绩效指标的 UMVUE，并提出了检验产品的寿命性能的检验程序，用以评估具有上记录值寿命数据的产品寿命性能。

目前有许多过程能力指数被用来评估过程的能力，但是实际应用中，以下几种过程能力指数仍然是现在最常用的[45~47]：

（1）$C_{\mathrm{P}}=\dfrac{\mathrm{USL}-\mathrm{LSL}}{6\sigma}$，该指标最早由 Juran（1974）提出，也是第一个过程能力指标，其中 USL 和 LSL 分别为过程的规格上限和规格下限，$\sigma$ 是过程的标准差。

（2）$C_{\mathrm{pk}}=\dfrac{d-|\mu-M|}{3\sigma}=\dfrac{\min\{\mathrm{USL}-\mu,\mu-\mathrm{LSL}\}}{3\sigma}$，由于 $C_{\mathrm{P}}$ 未能考虑过程平均值是否偏离规格中心。因此为了改善 $C_{\mathrm{P}}$ 指标的缺点，Kane（1986）提出了过程能力指标 $C_{\mathrm{pk}}$，其中 $\mu$ 是过程均值。

（3）$C_{\mathrm{pm}}=\dfrac{d}{6\sqrt{\sigma^2+(\mu-T)^2}}=\dfrac{d}{6\sqrt{E[(X-T)^2]}}$，这个指标由 Boyles（1991）提出，Boyles 发现 $C_{\mathrm{P}}$ 和 $C_{\mathrm{pk}}$ 这两个指标没有考虑到过程平均值偏离目标 $T$ 所带来的影响，于是根据 Chan 等（1988）提出的田口损失函数的概念，提出了一类新的过程能力指标 $C_{\mathrm{pm}}$。当过程平均偏离了目标值时，过程会有一个平方损失，因此过程指标 $C_{\mathrm{pm}}$ 更适用于各种不同规格界限的情况。

（4）此外，Pearn 等（1992）结合了 $C_{\mathrm{pk}}$ 和 $C_{\mathrm{pm}}$ 的观点，依据规格上下界与目标值的非对称性提出了 $C_{\mathrm{pmk}}$ 指标，其定义如下：

$$C_{\mathrm{pmk}}=\min\left\{\frac{\mathrm{USL}-\mu}{3\sqrt{\sigma^2+(\mu-T)^2}},\ \frac{\mu-\mathrm{LSL}}{3\sqrt{\sigma^2+(\mu-T)^2}}\right\}$$

以上四个过程能力指标都是评估在双边规格下具有望目型品质特性（the target-the-best type quality characteristic）的过程能力指标。

对于与产品寿命相关的产品，一般来说，顾客都希望产品的寿命越长越好，而且产品的寿命越长表示其品质越好，所以产品寿命的品质特征是属于望大型（the larger-the-better type）的品质特征。于是 Montgomery（1985）提出使用一种特殊的单边规格过程能力指标——寿命绩效指标$\left(C_L=\dfrac{\mu-L}{\sigma}\right)$来衡量产品的寿命绩效，其中 $L$ 是规格下界。

在大多数文献中，使用过程能力指标的统计推断研究均假设产品的品质特性服从正态分布，但实际生活中的产品如电子元件、引擎、变电器等的寿命往往服从的是指数分布、Pareto 分布、威布尔分布等非正态分布。Clement（1989）[48]、Pearn 和 Chen（1997）[49]、Liu 等（2006）[50]等人利用百分位数来估计过程平均数和标准差，进而对非正态分布过程能力指标进行推导。由于利用百分位数求解

会涉及大量的计算并且是基于正态分布的近似计算得到，结果并不准确，因此针对寿命型产品，特别是寿命越长越好的产品，Montgomery（1985）建议采用寿命绩效指标 $C_L$ 来衡量产品品质绩效。Tong 等[51]在电子元件寿命服从指数分布情形下，利用完全样本数据导出了 $C_L$ 的最小方差无偏估计。Wu 等[52]探讨了产品寿命服从 Rayleigh 分布下寿命绩效指标 $C_L$ 的最大似然估计，并通过此估计发展出了评估产品绩效的检验程序。何桢等[53]利用 Bootstrap 法得到了过程为正态分布和指数分布时过程能力指数 $C_{pk}$ 的置信区间。Hsu 等[54]研究了产品寿命数据为模糊数据情形下的 Parteo 分布的寿命绩效指标的统计推断问题，发展出了模糊环境下的寿命绩效指标 $C_L$ 的最大似然估计、区间估计以及基于 P 值的假设检验程序。晏爱君和刘三阳[55]提出了产品寿命服从指数分布的定数截尾数据情形下寿命绩效指标的 P 值检验程序，并通过实例说明了方法的有效性和可行性。Wu 等[56]在逐步递增的Ⅱ型截尾寿命试验下，讨论了 Rayleigh 分布产品寿命绩效指标 $C_L$ 的最大似然估计、最小方差无偏估计并进而发展了相应的产品寿命绩效检验程序。Laumen 和 Cramer[57]基于逐步递增的Ⅱ截尾寿命数据，讨论了一类特殊的指数分布族产品寿命绩效的最大似然估计和检验程序法。Lee 等[58]基于定数截尾样本，讨论了来自正态分布但是样本数据为模糊数情形的产品寿命绩效指标 $C_L$ 的最大似然估计以及假设检验问题，并给出了该类模型的产品品质绩效的检验程序。以上都是基于经典统计框架下产品寿命绩效指标的 Bayes 统计推断问题的研究。

当今社会的制造技术突飞猛进，产品的可靠性越来越高，通过截尾试验得到的样本数据就非常少，此时 Bayes 方法可以更好地处理小样本情形下模型参数的统计推断问题。但是目前关于寿命绩效指标的 Bayes 统计推断的研究比较少，因此有必要进行深入的研究。这方面的研究主要有：Lee 等[59]利用 Bayes 方法得到了平方误差损失下 Rayleigh 分布产品寿命绩效指标的 Bayes 估计并给出了相应的检验程序；Liu 和 Ren[60]导出了共轭先验分布下指数分布产品的寿命绩效指标的 Bayes 估计，并通过实际应用例子展示了产品品质绩效的 Bayes 检验程序。Lee 等[61]讨论了基于上记录值样本的 Weibull 分布产品的寿命绩效指标的最大似然估计估计和产品品质检验问题；Lee 等讨论了基于上记录值样本的 Rayleigh 分布产品的寿命绩效指标的 Bayes 估计和假设检验问题。

综上，本专著将进一步研究基于记录值样本的可靠性分布模型参数的 Bayes 统计推断问题和基于记录值样本的产品寿命绩效指标的统计推断问题。一方面可以丰富和发展记录值统计和 Bayes 统计推断理论，另一方面产品寿命绩效指标的检验程序可以为工程师在处理含有记录值数据信息时的产品性能评估问题时提供决策参考。

## 1.3 本书的主要研究工作

本书共分为五个章节：

第 1 章是绪论。主要介绍 Bayes 统计推断理论研究的背景意义和国内外研究现状。分析研究热点，指出本书的主要研究工作。

第 2 章是 Bayes 统计推断的基本理论知识。主要介绍 Bayes 统计的一些基本概念，如 Bayes 定理、先验分布、损失函数等概念。

第 3 章探讨基于记录值的几何分布、广义 Pareto 和比例危险率模型参数的 Bayes 估计以及估计的可容许性等问题，并初步探讨了基于记录值的指数分布参数的模糊 Bayes 点估计等问题。

第 4 章探讨基于记录值样本的指数分布和广义指数分布产品寿命绩效指标的统计推断问题。对于指数分布情形，提出了基于最小方差无偏估计和 Bayes 估计的寿命绩效指标的统计假设检验程序。

第 5 章为本书的总结和研究展望。

# 2 Bayes 统计推断的基础理论

1939 年，美籍罗马尼亚统计学家瓦尔德（A. Wald）提出，在根据观测数据对总体作出某种论断后同时考虑采用哪种决策，以及会产生怎样的后果。其将不确定意义下的决策科学也纳入到了统计学范围之内，意味着统计决策理论的建立。这一理论是数理统计学上一项重大的革新，具有极大的实际意义。

1763 年，英国学者托马斯·贝叶斯为了解决 James Bernoulli 提出的问题，在《论有关机遇问题的求解》中提出一种归纳推理的理论（Bayes' theorem），后被拉普拉斯等学者发展为一种系统的统计推断方法，这些方法构成了 Bayes 统计（Bayes statistics）的内容。Bayes 统计的基本方法是将关于未知参数的先验信息与样本信息综合，再根据贝叶斯定理得出后验信息，然后根据后验信息去推断未知参数。近半个世纪以来由于其应用的广泛性和良好的统计推断性质吸引了众多学者的关注和研究，已发展成为一套较为完整的统计推断理论体系，并被越来越多地应用于诸如可靠性工程、微型芯片制造、多元生物检测、医疗诊断和军事等各个领域。

本章在学习研究文献［62～104］的基础上，将主要介绍 Bayes 统计推断的基础理论。

## 2.1 统计决策问题的定义

### 2.1.1 决策问题的三要素

（1）样本空间和分布族。

**定义 2.1** 设总体 $X$ 的分布函数为 $F(x;\theta)$，$\Theta$ 是未知参数，$X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的一个样本，则样本所有可能值组成的集合称为样本空间，记为 $\mathcal{X}$，且有联合分布函数

$$F(x;\theta) = \prod_{i=1}^{n} F(x_i;\theta), \qquad \theta \in \Theta$$

称 $F^*$ 为样本的概率分布族，

$$F^* = \left\{ \prod_{i=1}^{n} F(x_i;\theta), \qquad \theta \in \Theta \right\}$$

（2）行动空间（决策空间）。一个统计问题中可能选取的全部决策组成的集合称为决策空间 $\mathcal{A} = \{a\}$，一个决策空间最少有两个决策。对于任何参数估计，每一个具体的估计值 $a$ 被称为一个决策。

(3) 损失函数 $L(\theta,a)$。统计决策假定每采取一个决策必有一定的后果，并将不同决策以数量的形式表示出来。若知道样本分布参数 $\theta$ 的值，行动空间的每一个行动 $a$ 所导致的损失函数记为 $L(\theta,a)$。$L(\theta,a)$ 是定义在 $\Theta\times\mathcal{A}$ 上的非负实值函数，表示当状态为 $\theta$ 采取的行动为 $a$ 时所产生的后果使决策者遭受的损失。

### 2.1.2 判决函数与风险函数

在损失函数中，一个特殊的 $\theta$ 和 $a$，是固定的，而样本信息 $X$ 是一个随机变量（或向量），人们通常考虑在 $X$ 的各种可能值下，求损失函数 $L(\theta,a)$ 的"平均值"。因此，根据样本选定动作构成一个纯策略的方法，就是下面定义的判决法则：

**定义 2.2** 任意 $x\in\mathcal{X}$，$a\in\mathcal{A}$，使得 $\delta(x)=a$，则称 $\delta(x)$ 为判决函数。$\delta(x)$ 为从 $\mathcal{X}$ 到 $\mathcal{A}$ 上的一个映射。对非数据问题，决策函数就是一个行动，若无特别声明；本章所讨论的都是非随机化的判决问题。

任何一个判决函数 $\delta(x)$ 都可以作为所给判决问题的解。但是，由于判决函数的选择都是与损失函数的多少相联系，有的可以大些，有的可以小些，故不能判断决策的好坏。因此，为了比较其优劣，在使用判决函数时，需看它们造成的损失是多少。这样，我们必须对损失函数 $L(\theta,a)$ 取平均值，用来从总体上评价、比较判决函数。

**定义 2.3** 设样本空间为 $\mathcal{X}$，分布族为 $F^*$，决策空间为 $\mathcal{A}$，$\delta(x)$ 为判决函数，则损失函数 $L(\theta,a)$ 对样本分布 $f(x\mid\theta)$ 的期望值：

$$R(\theta,\delta)=E_{X\mid\theta}[L(\theta,\delta(x))]=\begin{cases}\sum\limits_{x\in\mathcal{X}}L(\theta,\delta(x))f(x\mid\theta) & (\text{离散型})\\ \int_{\mathcal{X}}L(\theta,\delta(x))f(x\mid\theta)\,\mathrm{d}x & (\text{连续型})\end{cases}\tag{2.1}$$

为判决函数 $\delta(x)$ 的风险函数，$R(\theta,\delta)$ 表示采取判决法则 $\delta(x)$ 时所蒙受的平均损失（$L(\theta,a)$ 的数学期望）。由定义可知，风险就是平均损失，平均损失越小越好。

**例 2.1** 某投资公司购买了商业购物中心的五楼店铺，该商业店铺有两种状态：$\theta_1$——盈利，$\theta_2$——亏损；有三种决策：$a_1$——自己经营，$a_2$——全部出租，$a_3$——部分出租。如果盈利，选择自己经营，损失为 0，如果亏损而采取 $a_1$，则未损失为 12，等等。其损失函数见表 2.1。

**表 2.1 $L(\theta_i,a_j)$**

| | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $\theta_1$ | 0 | 10 | 5 |
| $\theta_2$ | 12 | 1 | 6 |