上海市研究生教学用书

GONGCHENG CANSHU DE

工程参数的

ZUIYOUHUA SHEJI

最优化设计

郁崇文
汪　军　编著
王新厚

东华大学出版社

本书得到上海市研究生教育专项经费资助

工程参数的最优化设计

郁崇文
汪　军　编著
王新厚

东华大学出版社

内 容 提 要

本书结合大量的纺织工程应用实例，对在工程应用中经常遇到的试验方案设计、试验数据处理、回归方程的建立与分析，以及基于回归方程的优化数学模型的建立与求最优化解等方面作了理论阐述和实例分析，并在书末配有相关的函数表和应用程序，便于教学、自学和实际应用。本书可作为纺织工程专业的硕士研究生学位课教材，也可用作其他工程类专业的研究生参考书，还可用作有关科研和工程技术人员的参考书。

图书在版编目(CIP)数据

工程参数的最优化设计/郁崇文等编著. —上海:东华大学出版社,2003.9
ISBN 7-81038-703-0

Ⅰ. 工... Ⅱ. 郁... Ⅲ. 设计参数-应用-纺织工业-工程设计:最优设计 Ⅳ. TS101.1

中国版本图书馆 CIP 数据核字(2003)第 083860 号

工程参数的最优化设计
郁崇文　汪　军　王新厚　编著
东华大学出版社出版
(上海市延安西路 1882 号　邮政编码:200051)
新华书店上海发行所发行　无锡春远印刷厂印刷
开本:787×1092　1/16　印张:20　字数:495 千字
2003 年 9 月第 1 版　2003 年 9 月第 1 次印刷
印数:0001—3000
ISBN 7-81038-703-0/TS · 161
定价:45.00 元

前　　言

纺织工程领域在新原料应用、新产品试制、新工艺研究中经常需要解决诸如具有各种不同性能指标原料的最佳混配比例，工艺参数的最佳组合，关键零、部件设计参数的组合优化，浆料、化学助剂的最佳配方等最优化问题，且这类优化问题所追求的目标往往是被加工的纤维制品或半制品各项质量指标的总体优化。对这类非线性问题，就我们目前对工程问题的实质理解，尚难建立起上述这些工程参数与制品质量之间精确的数学物理模型，进而在模型的基础上进行优化；现代优化计算方法，如神经网络、遗传算法等，虽不要求建立精确的数学物理模型，但新原料新产品、新工艺等研究由于缺乏同类问题的历史资料积累也难以直接应用，因而只能通过科学的试验研究方法予以解决。科学的试验研究方法首先要求以最少的试验工作量来获得尽可能多的关于研究对象的信息；其次是在研究的工程参数变化范围内建立起优化对象各指标与工程参数之间精度较高的函数关系；第三是通过对优化对象各指标间科学地协调与折中，达到指标的总体优化。

十多年来，东华大学纺织工程系新型纺纱研究室在连续承担原纺织工业部、纺织总会的重大科研项目中，应用回归试验设计和最优化设计相结合的方法，较好地解决了此类优化问题。通过回归试验设计，科学地安排试验方案，不仅减少了试验工作量，且在此基础上得到了工程参数与各质量指标之间的高精度回归方程；以各单一指标的回归方程为分目标函数，应用多目标优化方法，建立优化数学模型，通过对模型的约束或非约束求解，即可获得体现各分目标的总体优化的工程参数的最优化组合，且这种组合在实验研究中均得到了较好的验证。为此，纺织工程系在总结科研心得的基础上，为纺织工程专业的硕士研究生开设了“纺织优化设计”选修课，受到研究生的普遍欢迎，某些非纺织类专业的博士研究生对此课程也颇感兴趣，不久，“纺织优化设计”又先后被列为我校纺织工程专业的工学硕士和工程硕士研究生的学位课程。

本教材是在总结长期教学实践经验，并在多届应用的“纺织优化设计”学位课讲义的基础上对教材内容进行修改、补充、更新而成。本教材既沿用了若干优秀教材（见参考文献）中的一些基本理论内容，又结合编者长期教学、科研的心得积累，对教材内容的结构体系进行了改革，把回归设计和最优化设计有机地结合在一本教材内，书内还附有较多的纺织工程的综合应用实例及相应的算法程序，因而具有较强的实际应用意义，对那些目前尚难精确建立数学物理模型，有待通过实验研究进行解决的工程参数优化问题有一定的参考价值。

本教材共分三篇，十二章。第一篇的主要内容是回归试验设计及试验数据的分析、处理；第二篇主要介绍了对试验设计所得到的回归方程，建立完整的优化数学模型，并进行数值计算优化的内容；第三篇则对模糊决策及灰色近优评判等作了初步介绍。由于在纺织工程领域，对有些零、部件的设计参数优化问题，考虑到成本等问题，往往不是通过严格的试验设计进行优化，而是根据现有的各种零、部件，在满足基本问题要求的前提下进行优选。解决此类优选问题，应用模糊数学和灰色理论中的若干原理是很有效的。为此，在第三篇中，编者结合自己的

研究实践，着重向读者介绍了模糊决策和灰色近优评判的基本原理和应用。

在本教材的编写过程中，东华大学新型纺纱研究室的黄秀宝教授对整个教材的结构体系、应用实例的选取等，提出了指导性的意见。

华东师范大学的茆诗松教授、北京科技大学的陈立周教授先后审阅了书稿，并提出了许多中恳的意见和建议。

在本教材的编写、出版过程中，得到了上海市学位委员会和东华大学研究生院的大力支持；研究生史春霞、常英健、周玲、步红刚等参与了文稿的整理、打印和程序的编写。

对上述为本教材编写提供帮助的单位和个人，作者在此表示衷心的感谢。

由于我们水平有限，书中的错误与不足在所难免，恳请有关专家和读者批评指正。

2003 年 5 月

目　　录

第一篇 试验设计

在工程实际及其他许多应用场合，为了探索某些指标受其他因素的影响程度（或两者的关系），在尚未有成熟的理论解析式能直接表达这种关系、或者由于影响因素实在太多，难以全面考虑时，往往需要通过安排一系列的试验，以观察在已知的特定影响因素条件下，该指标的变化情况，并通过对试验结果的统计分析，找出指标和影响因素之间的回归关系式。从而为对相关指标和因素进行控制和预测提供依据。显然，试验结果的合理性和正确性，是得出正确的回归关系的前提，此外，由于方差统计分析对数据的随机性、正态分布性、方差一致性等有特定的要求，因此，在进行回归分析前，必须对试验所得的结果进行数据的检验和处理，以确保满足以后统计分析的要求。

同时，由于人力、物力和其他条件的限制，如何以尽可能少的试验次数来使得试验结果尽可能全面、准确，从而获取尽可能多的有关信息，是每一个试验者所期望的。在统计分析和建立回归方程的过程中，如何使计算尽量简单、准确、方便、快速，也是人们关心的问题。本篇所叙述的就是如何根据实际需要，通过对试验方案的合理设计，达到以较少的试验工作量，获取尽可能全面和准确的结果，并尽量减少和简化计算，从而得出尽可能准确全面的回归方程，为其用作工程实际中的（产品质量等）结果预报和（工艺参数等）因素的选择提供基础。

第一章 试验数据的处理

试验中,由于随机性、仪器、设备和操作中的问题,会不可避免地导致试验结果中存在一些系统或随机性的误差,从而影响分析的准确性。此外,在应用回归分析、方差分析等统计分析手段时,对试验数据也有一定的要求,即:样本对总体的代表性和样本遵从特定的检验方法的要求。因此,必须在对试验数据进行统计分析之前,对试验数据本身进行预处理,使其符合统计分析的要求。

例如,在方差分析时,因其分析是建立在一些重要的假定的基础上的,也就是说,只有当所采集的数据符合有关假定时,才能使用方差分析。这些假定就是:个体的独立性、误差的正态性和方差的一致性。

独立性是指每个样品个体的采集不受上一个样本的影响,通常可以通过采样的随机性来保证,所谓采样的随机性意味着总体中每个个体被采集的机会均等,这实际上是样本对总体的代表性问题。非随机采样可能导致一系列问题,如方差的不一致性和误差的非正态性等。样本的正态性是常用数理统计方法的前提条件。尽管这一假定至关重要,但一个样本是否具有随机性根本无法用统计方法加以检验,只能靠实验者在进行观测和试验设计时加以注意。

第一节 独立性检验

保证个体的独立性就是要使所采集的每个样本与上一个样本没有关联。数据是否符合独立性假定可以用游程检验加以判断。例如,在小区栽培实验中,常常会因为相邻地块的性质相似而造成数据的非独立性特征。如果不将同一处理水平的样点放在相邻位置,即将试验地块作随机化安排,就可以保证获得具有独立性的样本。

一个样本是否能代表它的总体,关键在于采样的随机性。非随机采样一方面可能导致样本特征与总体特征的不一致,另一方面也可能反映在个体的非独立性方面,样本中个体的独立性是指它们的采集过程完全不受其他个体的干扰。对那些分布在某一时间或空间范围内的个体而言,个体间的自相关特性是导致采样非独立性的重要原因,而样本的非独立性则会造成统计分析的错误结论。例如,为了比较两个同类城区的噪声污染程度,先在第一区测定若干数据,再到第二区进行测定。由于噪声水平往往有随时间变化的趋势,这样得到的数据就有一定程度的自相关性,即一个测定值与其邻近测定值有某种相似性,直接比较上述两组数据显然不能正确反映两个城区间的差别。

采样的独立性首先应当在采样过程中加以注意。例如在上述两城区噪声比较研究中,可以考虑在两个城区同步采样以克服时间因素的影响。本节提供的游程检验虽然可用于样本非独立性检验,但由于方法本身仅适用于二分类类型数据,从方法学角度出发,这类检验的效率很低,不能指望仅仅由这样的检验来保证采样的可靠性。

游程检验是一种假设检验的方法,使用时先假定采集的个体具有独立性特征,然后通过适当方法加以证实或加以否定。游程检验基于样本显示的游程数(r)。一个游程意味着一个连

续的、具有相同取值的数据串，在它前后与其相邻的则是不同取值。波动游程检验是游程检验的基本方法，该方法使用二分类类型数据，对原始数据就是二分类者，可直接进行检验，若原始数据并非二分类类型，那么必须在检验前作数据变换(类型转换)，使成为二分类数据。例如，在原始数据是连续量的情况下，可根据每个数据与中位数的大小关系，分别用正号和负号代表大于和小于中位数的观测值，如果样本量为奇数，则删去等于中位数的个体，这样就得到一组个体取值为正号或负号的二分类样本。这种基于中位数的检验又叫中位数上下游程检验，对于这种方法，一个游程就是一连串正号或一连串负号，该数据串前后均为另一种符号，对其他定量数据，如离散量或顺序量，也可以参照此法演变出类似的游程检验来。

将总样本量记为 n，分别用 n_1 和 n_2 代表两个类别的数据个数，对原始数据不是类型变量的中位数上下游程检验或类似变通办法，n_1 一般等于 n_2，检验时首先将 n 个观测值按出现顺序(采样顺序)排列，然后求得游程数 r。例如，对下列二分类样本，其游程数为 6。

M	M	M	M	F	F	M	M	M	F	M	F	F	F

对以下数据，由于属于连续量，先按中位数 9.5 转换为正号或负号，大于 9.5 者取正号，反之取负号，则可知，其游程数等于 7。

16	12	10	9	5	8	15	5	2	1	8	16	20	3	20	12
+	+	+	−	−	−	+	−	−	−	−	+	+	−	+	+

当将数据按大小次序排列后，其中位数可按下式计算

$$\text{中位数} = \begin{cases} x_{k+1}, & (n = 2k+1) \\ \dfrac{1}{2}(x_k + x_{k+1}), & (n = 2k) \end{cases}$$

如果 n_1 和 n_2 都小于 20，可以直接从波动游程检验临界值表(附表 A-1)中得到一对检验临界值 $r_{1,0.05}[n_1, n_2]$ 和 $r_{2,0.05}[n_1, n_2]$，限于篇幅，表中仅仅列举了 0.05 显著性水平的临界值，如果计算游程数 r 落在两个临界值之外，即

$$r \leqslant r_1 \quad \text{或} \quad r \geqslant r_2$$

那么可以认为该样本中的个体不具备独立性。

在样本量较大时 ($n_1 > 20$ 或 $n_2 > 20$)，r 的抽样分布近似于正态，可以先计算

$$t_s = \frac{\left| r - \dfrac{2n_1 n_2}{n_1 + n_2} - 1 \right| - 0.5}{\sqrt{\dfrac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}} \tag{1-1}$$

将计算值与附表 A-2 中的 $t_\alpha[\infty]$ 相比，如果

$$t_s > t_\alpha[\infty]$$

则可以认为该样本中的个体不具有独立性，这里 α 一般取 0.05。

除波动游程检验外，还有一些适合于检验个体独立性的方法，如利用一级自回归系列模型的参数方法以及非参数的 Von Neumann 比计算。由于这些方法都不太常用，故不再作详细

介绍。

【例 1】 检验以下样本采集的独立性：

9	8	5	3	1	8	9	7	8	5	2	8	1	2	2	9	9	8	9	5	4·	7	2	1

根据取值大于还是小于它们的中位数 6，将之转换为二分类数据：

+	+	−	−	−	+	+	+	+	−	−	+	−	−	−	+	+	+	+	−	−	+	−	−

对此组数据，有：$n_1 = 12,\ n_2 = 12,\ r = 10$。

由于 n_1 和 n_2 均在 20 以下，可直接从附表 A-1 中查到临界值，有

$$r_{1,0.05}[12,\ 12] = 7 \quad 和 \quad r_{2,0.05}[12,\ 12] = 19$$

计算 r 值落在两临界值之间，不能认为样本中个体不具独立性。

第二节　异常值检验

在一批试验数据中，如混杂有异常数据，则必然会歪曲实验结果。因此，必须正确地剔除异常数据(或称坏值)。另一方面，由于在特定条件下进行实验测量的随机波动性，致使测量数据有一定的分散性，如果人为地丢掉一些误差较大的、但不属于异常的数据，这样会造成虚假的高精度，这也是不正确的。

人们对异常数据的判别与剔除，往往采用两种方法：

1. *物理判别法*

即在试验过程中，人们根据常识或经验，判别由于仪器或设备的震动、数据的误读等原因造成的坏值，随时发现，随时剔除。

2. *统计判别法*

统计判别法的基本思想在于，给定一个置信概率(例如 0.99)，并确定一个置信限，凡超过此限的误差，就认为它不属随机误差范围，系属异常数据，应予剔除。

一般情况下从一个总体中抽样时，取值越接近分布中心，其值出现的可能性就越大。相反，那些距分布中心远的取值出现的概率就较小。例如，从一个均值等于 0，方差等于 1 的标准正态分布总体中抽样时，取值落在 0 附近的可能性远远大于取值在 5 左右的概率。一个样本中出现概率很小的值叫做异常值。无论异常值的出现纯属偶然还是与观测过程中的某些失误有关，其存在都会对统计分析结果产生影响。这当然不是实验者所希望的。为了克服少数异常值带来的干扰，常常有必要在进行统计分析之前检验并剔除样本中的异常值。常用的异常值检验方法包括 t-检验、Grubbs 检验、Dixon 检验和 Walsh 检验等。

试验者可以在选定的正确性概率下根据上述四种方法作出某个或某些观测值是否属于异常的判断，可以主观确定的这一最大允许错误率记作显著性水平 α，用以表示某观测值并非异常，而检验结果却将它判断为异常的可能性，通常取 $\alpha = 0.05$。这意味着如果检验结果认为某值是异常，该结论不正确的概率不会大于 5%，由此可见，试验者可以通过改变 α 值来调整检验方法的严格程度。假如宁可错误地剔除非异常数据也不愿意放过可能的异常值，那么应当选择大一些的 α 值，反之，如果要求尽量不作出错误剔除，那么可用较小的 α 值进行检验。

除 Walsh 法可同时检验若干个可疑值外，以下介绍的方法都是针对一个可疑值进行的。如果实验者怀疑的观测值不止一个，就必须逐个加以检验并决定取舍。一般步骤是先将所有数据从小到大排列，以两端极值作为可疑值，然后分别加以检验。如果发现最大值或最小值是异常，那么可以将其剔除并进一步检验次大值或次小值，直至剩余数据的最大和最小值都不再是异常为止。

Grubbs 检验和 t-检验适用于采自正态分布总体的样本。这两种检验的主要差别反映在它们的严格性（其反面是保守性）程度方面。相比之下，t-检验要比 Grubbs 检验严格。这就是说，如果取相同的 α 值，并用这两种方法检验同一样本，t-检验方法剔除的异常值可能多于（至少等于）Grubbs 检验的结果。因此，实验者也可以根据特定的严格性要求在这两种方法中进行选择。

Grubbs 检验和 t-检验的检验值都等于可疑值与算术均值之差的绝对值除以标准差，分别用 G_s 和 K_s 表示其统计量。但 Grubbs 法使用的均值 $\bar{x}$ 和标准差 s 是全体数据的统计量，即，对全体数据中的每个数据 x_i，有

$$G_s = \frac{|x_i - \bar{x}|}{s} \tag{1-2}$$

而 t-检验用的是不包括可疑值在内数据的计算结果。分别用 $\tilde{x}$ 和 $\tilde{s}$ 表示不包括可疑值数据的均值和标准差，有

$$K_s = \frac{|x_i - \tilde{x}|}{\tilde{s}} \tag{1-3}$$

两种方法的检验值计算中都不应当包括已被剔除的异常值。如果检验值大于相应临界值，即当

$$G_s > G_\alpha[n] \qquad \text{(Grubbs 检验)}$$

或

$$K_s > K_\alpha[n] \qquad \text{(t-检验)}$$

时，可以判定该可疑值为异常。这一判断的可靠性为 $1-\alpha$。两种方法的检验临界值 $G_\alpha[n]$ 和 $K_\alpha[n]$ 值可分别根据 α 和样本量 n（包括可疑值在内）从异常值 Grubbs 检验临界值表（附表 A-3）和异常值 t-检验临界值表（附表 A-4）中查到。

在很多情况下，试验者对总体是否为正态分布以及样本中是否存在异常观测值都没有把握。虽然数理统计方法分别提供了这两个方面的检验手段，但对它们的相互影响却很难判断。一方面，当样本来自非正态分布总体，如对数正态分布总体时，先进行异常值检验可能导致错误剔除，从而使剔除后样本趋近正态分布。另一方面，如果样本代表的背景总体确属正态分布，而试验者在异常值剔除之前先进行分布检验，那么由于样本中个别异常值的存在造成分布向一侧偏斜，从而可能得出总体不属正态分布的错误结论。由此可见，究竟应当先用上述两种方法作异常值剔除还是先进行分布检验是一个很难回答的问题，试验者只能凭经验作出适当的选择。

Dixon 检验也适用于正态分布情形，但由于其检验值计算式因样本量不同而异，一般仅用于小样本量数据。采用 Dixon 检验时首先根据样本量 n 和 α 值从 Dixon 检验临界值与检验系

数计算式表(附表 A-5)中查得 Dixon 系数 D_s 的计算式(表中右侧)和相应的检验临界值(表中左侧)。根据计算式求出检验系数后再与临界值比较,如果

$$D_s > D_\alpha[n]$$

则该可疑值异常。

Walsh 检验是一种非参数方法,可用于任意分布对象,使用时不需要临界值表。与上述所有方法的另一重要差别在于这种检验一般只适用于大样本量数据。事实上,只有当以下关系成立时才有可能使用 Walsh 检验。

$$\mathrm{Trunc}(\sqrt{2n}) > 1 + \frac{1}{\alpha} \tag{1-4}$$

式中,Trunc 函数代表取整运算。根据此式,当 α 取值为 0.05 时,样本量必须在 220 以上。另外,Walsh 检验不必像上述方法那样对可疑值进行逐个计算,它可以同时检验若干可疑数据。如果实验者怀疑数据中最大的 r 个或最小的 r 个数据为异常值的话,首先计算

$$c = \mathrm{Trunc}(\sqrt{2n}) \tag{1-5}$$

$$k = r + c \tag{1-6}$$

$$b = \frac{1 + \sqrt{\dfrac{c - \dfrac{1}{\alpha}}{\alpha(c-1)}}}{c - \dfrac{1}{\alpha} - 1} \tag{1-7}$$

如果可疑值为数据中的最小值,求

$$W = x_r - (1+b)x_{n-r} - bx_{n+1-k} \tag{1-8}$$

若可疑值为最大值,则计算

$$W = -x_{n+1-r} + (1+b)x_{n-r} - bx_{n+1-k} \tag{1-9}$$

将这 r 个可疑值判定为异常的条件是

$$W < 0$$

上述几种异常值检验方法及各自的适用条件归纳在表 1-1 中。

表 1-1　异常值检验方法

检验方法	适用范围及特点
t-检验	正态分布样本,较严格
Grubbs 检验	正态分布样本,保守性适中
Dixon 检验	正态分布样本,小样本量,较保守
Walsh 检验	正态或非正态分布样本,大样本量

【例 2】 用不同方法从下列观察数据中剔除异常值:

−4.44, −1.41, −0.78, −0.78, 0.55, 0.64, 1.05, 1.05, 1.72, 5.91, 9.30, 20.49

【解】 将上述数据按从小到大按顺序排列,即

$$X_1 = -4.44,\ X_2 = -1.41,\ \cdots,\ X_{11} = 9.30,\ X_{12} = 20.49$$

(异常值用 * 表示,非异常值记为—)

(1) 用 t-检验,分别对数据中的最大、最小值进行检验,结果如下:

可疑值	x_i	n	$\tilde{x}$	$\tilde{s}$	$\lvert x_i - \tilde{x} \rvert$	K_s	$K_{0.05[n]}$	结论
X_1	−4.44	12	3.404	6.175	7.844	1.27	2.33	—
X_{12}	20.49	12	1.137	3.471	19.353	5.58	2.33	*
X_{11}	9.30	11	0.321	2.434	8.979	3.69	2.37	*
X_{10}	5.91	10	−0.267	1.769	5.877	3.32	2.43	*
X_9	1.72	9	−0.515	1.722	2.235	1.30	2.51	—

(2) 以相同的方式用 Grubbs 检验得到的结果比 t-检验略为保守,仅将 X_{11} 和 X_{12} 确定为异常值:

可疑值	x_i	n	$\tilde{x}$	$\tilde{s}$	$\lvert x_i - \tilde{x} \rvert$	G_s	$G_{0.05[n]}$	结论
X_1	−4.44	12	2.750	6.297	7.190	1.14	2.29	—
X_{12}	20.49	12	2.750	6.297	17.740	2.82	2.29	*
X_{11}	9.30	11	1.137	3.471	8.163	2.35	2.23	*
X_{10}	5.91	10	0.321	2.434	5.289	2.17	2.18	—

(3) 用 Dixon 检验发现的异常值与 t-检验的完全一样:

可疑值	x_i	n	D_s	$D_{0.05[n]}$	结论
X_1	−4.44	12	0.266	0.546	—
X_{12}	20.49	12	0.679	0.546	*
X_{11}	9.30	11	0.708	0.576	*
X_{10}	5.61	10	0.554	0.477	*
X_9	1.72	9	0.214	0.512	—

(4) 用 Walsh 检验时,因样本量很小,α 值至少应取 0.34。在实际应用中,Walsh 检验不适合于检验如此小的数据。此例仅用来表明如何运算。先假定最大和最小值为异常,故取

$$r = 1,\ c = \mathrm{Trunc}\sqrt{24} = 4,\ k = 1 + 4 = 5$$

$$b = \frac{1 + \sqrt{\dfrac{4 - \dfrac{1}{0.34}}{(0.34) \times 3}}}{4 - \dfrac{1}{0.34} - 1} = 34.32$$

对最小值的检验值

$$W_1 = -4.44 - (1 + 34.32) \times (-1.41) + 34.32 \times (0.55) = 64.24$$

相应地,对于最大值的检验值

$$W_2 = -20.49 + (1 + 34.32) \times (9.30) - 34.32 \times (1.05) = 271.95$$

对最小和最大值的计算值 W 都大于零,可见它们都不是异常值。用 Walsh 检验得到的结果

与其他方法完全不同的主要原因是此例样本量太小，故所取α值偏大。

第三节　正态性检验

分布的正态性是数理统计方法中最重要的假定，包括参数检验、方差分析、相关分析和回归分析在内的大多数统计分析都要求数据符合这一假定。造成试验结果非正态性的原因有两种，一是总体本身就不是正态分布的，二是样本中包含个别异常值。对于非正态分布的总体，做适当数据变换（如对对数正态分布总体做对数变换，对左偏或右偏数据作 Box-Cox 变换等），或者改用非参数检验方法都是可行的选择。在有异常值存在的情况下，则应将其先剔除再作方差分析。

正态分布检验不仅可用来判断原始变量是否服从正态分布，还常常用于检验非正态分布总体经某种数学变换后是否成为正态分布形式。关于对数正态分布和指数正态分布的检验就是常见的两例。从属对数正态分布的总体经对数变换后即成为正态分布，而指数变换可以将指数正态分布总体转换为正态分布形式。对这些分布类型的检验与正态检验方法雷同，只要在作正态检验之前对原始数据进行适当变换即可。

一、几种正态分布的检验方法及比较

将在本节介绍的正态分布检验方法主要包括偏度-峰度检验，Shapiro-Wilk 检验和 Lillifors 检验。其中偏度-峰度检验由于对数据信息利用充分，且可分别描述分布的偏斜程度和峰形陡缓，因此是最常用的检验方法。相比之下，Shapiro-Wilk 正态分布检验更适用于小样本量的场合。

本节还包括几种不太严格的，但快速、简便的正态检验方法。这些快速检验方法虽然比较粗糙而且极易受个别异常值的干扰，但在用作对某些样本作大致判断，以确定能否使用特定的参数统计方法时仍有其实用价值，本节将要介绍的此类方法包括，D'Agostino 法以及正态分布的经验判断法则和 David 快速检验。有关方法及各自的特点如下表所列。

表 1-2　正态性检验方法

检验方法	适用性及特点
偏度-峰度检验	仅对偏度或峰度敏感
Shapiro-Wilk 检验	小样本数据
David 检验	简便但易受异常值干扰
D'Agostino 检验	简便但不严格，仅对峰态敏感

1. 正态分布的偏度-峰度检验

偏度系数和峰度系数是两个描述总体中个体分布形式的重要统计量。在理论正态分布总体中的这两个参数都应当为零，对偏度系数而言，该值大于零说明分布峰右偏，而小于零则表示左偏。另一方面，峰度系数大于零和小于零的两种情形分别反映了分布峰的尖峰态和平坦峰态。鉴于这两个统计量的上述特征可以利用偏度系数和峰度系数检验一个总体的分布类型是否为正态，亦即检验总体的这两个参数是否等于零，这就是偏度-峰度检验。

与其他正态分布检验方法相比，偏度-峰度检验的优点是不言而喻的。作为典型的参数方

法，其检验功效高，并有明确的概率意义，甚至能利用 t-分布函数得到拒绝原假设时的准确相伴概率。除此之外，这种方法的另一优点在于可以从偏度和峰度两个不同角度来判断总体对正态分布的偏离，这在对研究对象分布形式的描述和解释方面都非常有用。譬如一个右偏的尖峰分布很可能说明数据分布接近对数正态形式。当然总体对正态分布的偏离有时也会反映在其他方面，比如双峰态就是其中一种。偏度-峰度检验对判断这样的分布特性当然是无能为力的。

偏度-峰度检验实质上是分别针对总体偏度系数和峰度系数进行 t-检验的两种独立方法。由于分布的偏斜形式有左偏和右偏两种可能，而峰度同样有尖峰和平坦峰两种状态，因此关于偏度系数和峰度系数的统计检验与许多其他检验一样有单、双侧之分。正因为如此，偏度检验和峰度检验各有两种不同的对立假设。分别用 γ_1 和 γ_2 表示总体偏、峰度系数，偏度系数 t-检验的原假设为

$$\mathrm{H}_0: \gamma_1 = 0$$

而对立假设区别单、双侧两种情况分别是

$$\begin{aligned} &\mathrm{H}_1: \gamma_1 \neq 0 \quad &(\text{双侧检验}) \\ &\mathrm{H}_1: \gamma_1 > 0 \quad &(\text{单侧检验}) \\ &\mathrm{H}_1: \gamma_1 < 0 \quad &(\text{单侧检验}) \end{aligned}$$

双侧峰度检验的统计假设是

$$\begin{aligned} &\mathrm{H}_0: \gamma_2 = 0 \\ &\mathrm{H}_1: \gamma_2 \neq 0 \end{aligned}$$

相应的单侧检验的原假设与双侧检验相同，但对立假设却为

$$\begin{aligned} &\mathrm{H}_1: \gamma_2 > 0 \\ &\mathrm{H}_1: \gamma_2 < 0 \end{aligned}$$

如果用 g_1 和 g_2 表示样本量为 n 的样本偏度系数和峰度系数的话，偏度检验和峰度检验的计算统计量分别是

$$t_1 = \frac{|g_1|}{\sqrt{\dfrac{6n(n-1)}{(n-2)(n+1)(n+3)}}} \tag{1-10}$$

$$t_2 = \frac{|g_2|}{\sqrt{\dfrac{24n(n-1)}{(n-3)(n-2)(n+3)(n+5)}}} \tag{1-11}$$

其中

$$g_1 = \frac{\sqrt{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}} = \frac{\hat{\mu}_3}{s^3}, \quad g_2 = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2} = \frac{\hat{\mu}_4}{s^4}$$

其中，s^2 为样本的二阶中心矩即方差，$\hat{\mu}_3$、$\hat{\mu}_4$ 分别为3阶和4阶中心矩。

像其他 t-检验方法一样，可以将计算值与 t-检验临界值比较以决定是否拒绝检验的原假设。在 α 显著性水平，对双侧检验，拒绝原假设的条件是

$$t_1 > t_\alpha[v] \ (\text{偏度检验})$$
$$t_2 > t_\alpha[v] \ (\text{峰度检验})$$

自由度 $v = n-1$。

对单侧检验，则在以下条件下拒绝检验的原假设

$$t_1 > t_{2\alpha}[v] \ (\text{偏度检验})$$
$$t_2 > t_{2\alpha}[v] \ (\text{峰度检验})$$

对一个检验对象，只要偏度检验或峰度检验中有一个拒绝了原假设，就可以认为该样本来自非正态分布的总体。换言之，对于服从正态分布的总体，两项检验结果都不应当显著。

【例3】 测定了来自某地区表土82个样品中的锌含量。为作进一步统计分析，应先作正态分布检验，以确定有没有必要进行数据变换或选用非参数方法。测定值的有关统计量如下（原始数据从略）：

$$n = 82,\ g_1 = 2.71,\ g_2 = 11.18$$

分别建立两组统计假设。对于偏度检验，有

$$\mathrm{H_0}: \gamma_1 = 0$$
$$\mathrm{H_1}: \gamma_1 \neq 0$$

对峰度检验的假设为

$$\mathrm{H_0}: \gamma_2 = 0$$
$$\mathrm{H_1}: \gamma_2 \neq 0$$

据式(1-10)和(1-11)，分别计算两个检验统计量

$$t_1 = \frac{2.71}{\sqrt{\dfrac{6(82)\times(82-1)}{(82-2)\times(82+1)\times(82+3)}}} = 10.2$$

$$t_2 = \frac{11.18}{\sqrt{\dfrac{24(82)\times(82+1)}{(82-3)\times(82-2)\times(82+3)\times(82+5)}}} = 189.11$$

如果以0.05为显著性水平，自由度 $v = n-1$，从附表 A-2 中得到

$$t_{0.05}[81] = 1.99$$

临界值大大低于两个计算值，可以有把握地拒绝检验的两个原假设。可见该地区表土锌含量不属正态分布。由于偏度系数和峰度系数均明显大于零，该总体呈右偏尖峰态分布。据经验可知，土壤微量元素含量常常为对数正态分布形式（恰好呈右偏尖峰态）。下面对原始数据作对数变换，再对变换后的数据进行偏度-峰度检验，即分别对两组统计假设作偏度-峰度