

文章编号: 1000-8055(2003)02-0167-08

置信检验理论

傅惠民, 梁朝虎

(北京航空航天大学 固体力学研究所, 北京 100083)

摘要: 本文提出了母体特征值置信分布的概念, 建立了置信检验理论, 该理论包括强检验、弱检验和弱强检验三种检验方法, 分别满足了不同工程检验的需求。与假设检验相比, 置信检验依据的是“大概率原则”, 而显著性假设检验依据的是“小概率原则”, 因此, 置信检验能够以高概率确认工程中所关心的“条件”是否成立(如产品合格与否), 能够对产品质量进行高精度的检验和控制, 克服了显著性假设检验仅仅是一种拒绝检验, 当接受一个假设时也只是意味着不拒绝该假设的缺点。置信检验不但能够大大提高工程, 而且更加便于工程应用。

关键词: 航空、航天推进系统; 置信检验; 强检验; 弱检验; 弱强检验; 假设检验; 抽样检验; 质量控制
中图分类号: TB301; O211 **文献标识码:** A

The Theory of Confidence Tests

FU Huimin, LIANG Zhao-hu

(Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The concept of confidence distribution of population characteristic is presented in this paper. Based on the confidence distribution, a new theory of confidence tests is established, which includes three kinds of tests: strong test, weak test and weak-strong test. The confidence tests are based on The Law of Small Probability, while the significance hypothesis tests are based on The Law of Large Probability. Compared with hypothesis tests, the confidence tests can inspect and control the quality of products with higher precision, and overcome the disadvantage that to accept a null hypothesis may simply mean that it is not rejected by significance hypothesis tests. In addition, the application of the confidence tests can greatly improve the test precision and bring more convenience in engineering.

Key words: aerospace propulsion system; confidence test; strong test; weak test; weak-strong test; hypothesis test; sampling inspection; quality control

1 引言

对于一种假设, 若找到一个与该假设相符的例子, 并不能说明该假设成立, 但是, 若能找到一个反例, 则可以断定该假设不成立, 这是众所周知

的反证法。目前工程中各类检验的基础——假设检验^[1~3]采用的就是反证法, 所以假设检验在拒绝原假设时有充分的理由, 但是当不能拒绝原假设时却没有足够的理由接受原假设。工程检验的目的是要确认所关心的“条件”是否成立(如产品

收稿日期: 2003-02-08; 修订日期: 2003-02-23

基金项目: 国防科技预研项目(413200204)

作者简介: 傅惠民(1956—), 男, 浙江遂昌人, 北京航空航天大学飞行器设计与应用力学系“长江学者”特聘教授, 博士, 从事小样本信息技术、软校准技术、数据融合方法及可靠性研究

合格与否),假设检验只能确认这些“条件”不成立,而无法确认这些“条件”成立,所以不能满足工程实际的需要。目前在工程使用中,接受原假设仅仅意味着不拒绝原假设,这往往导致对同一问题,由于所作的假设不一样,得到的结论截然不同,甚至产生相互矛盾的结果,这也是将许多不合格产品误判为合格产品,酿成重大事故的原因之一。对此,国内外都进行了大量研究^[1~3],但一直未能解决这一难题。文献[4~6]在解决母体百分率和百分位值的检验问题时,提出了一种依据“大概率原则”进行检验的思想,克服了假设检验的上述缺点。本文又进一步提出了母体特征值(如母体均值、方差、百分位值和百分率等)置信分布的概念,系统建立了置信检验的理论和方法。与假设检验和抽样检验相比,置信检验可以大大提高检验精度,且更加便于工程应用。

2 置信检验理论基础

设 λ 为母体 X 的一个特征值(如母体均值、方差、百分位值和百分率等), $x = (x_1, x_2, \dots, x_n)$ 是来自母体 X 的一个样本,那么由经典数理统计方法,可以求得 Neyman 意义下 λ 的置信度为 Y 的单侧置信上限 λ_U 和下限 λ_L , 即有:

$$P(\lambda \leq \lambda_U) = Y \quad (1)$$

$$P(\lambda \leq \lambda_L) = 1 - Y \quad (2)$$

在 Neyman 意义下,式(1)和式(2)中的 λ 是一个确定的量, λ_U 和 λ_L 是样本 x 和 Y 的函数,当样本 x 随机变化时, λ_U 和 λ_L 也随机变化,所以 λ_U 和 λ_L 都是随机变量。但是,在 Fiducial 意义下,式(1)和式(2)却有全新的解释,即将样本 x 固定, λ 看作随机变量, λ_U 和 λ_L 仅是 Y 的函数,此时式(1)和式(2)都可以看作 λ 的分布函数, λ_U 和 λ_L 则是 Fiducial 意义下 λ 的置信度为 Y 的单侧置信上限和下限。有关研究表明, Neyman 意义下的置信限与 Fiducial 意义下的置信限不一定相同^[1],但是,在上述问题中它们则是相同的。Neyman 意义下的置信度可以用频率解释,但必须在反复抽样中才能实现,而 Fiducial 意义下的置信度在一次抽样中具有明确的意义,但不能用频率解释。在本文所研究的问题中,二者是相同的,因此可以充分发挥它们各自的优点,使置信度既能用频率解释,又在一次抽样中有明确的意义。

定义 1 当母体特征值 λ 在 Fiducial 意义下的置信限与 Neyman 意义下的置信限相同时, λ 的 Fiducial 分布被称为置信分布。

定理 1 若 λ_U 和 λ_L 分别是 Neyman 意义下母体特征值 λ 的单侧置信上、下限,令

$$F_U(\lambda_U) = P(\lambda \leq \lambda_U) \quad (3)$$

$$F_L(\lambda_L) = P(\lambda \leq \lambda_L) \quad (4)$$

那么 $F_U(\lambda)$ 和 $F_L(\lambda)$ 均为母体特征值 λ 的置信分布函数,分别被称为 λ 的上置信分布函数和下置信分布函数。对于连续母体, λ 的上、下置信分布函数是相同的,即 $F_U(\lambda) = F_L(\lambda)$, 统一记作 $F(\lambda)$ 。对于离散母体, λ 的上、下置信分布函数一般不相同,当讨论有关上限问题时,用上置信分布函数,当讨论有关下限问题时,用下置信分布函数,因此也可以统一记作:

$$F(\lambda) = \begin{cases} F_U(\lambda) & \text{当 } F_U(\lambda) \geq 0.5 \\ F_L(\lambda) & \text{当 } F_L(\lambda) \leq 0.5 \end{cases} \quad (5)$$

定理 2 设 λ 为母体 X 的一个特征值, λ_0 , λ_U 和 λ_L 为规定的指标值。对于来自该母体的一个样本 $x = (x_1, x_2, \dots, x_n)$, 可以相应地求得 λ 的置信分布函数 $F(\lambda)$, 若

$$Y = F(\lambda_0) = P(\lambda \leq \lambda_0) \quad (6)$$

则可在置信度 Y 下(或以 Y 的把握)断定 $\lambda \leq \lambda_0$ 成立;若

$$Y = 1 - F(\lambda_0) = P(\lambda \geq \lambda_0) \quad (7)$$

则可在置信度 Y 下(或以 Y 的把握)断定 $\lambda \geq \lambda_0$ 成立;若

$$Y = F(\lambda_U) - F(\lambda_L) = P(\lambda_U \leq \lambda \leq \lambda_L) \quad (8)$$

则可在置信度 Y 下(或以 Y 的把握)断定 $\lambda_U \leq \lambda \leq \lambda_L$ 成立。其中 $1 - Y$ 为对式 $\lambda \leq \lambda_0$, $\lambda \geq \lambda_0$ 或 $\lambda_U \leq \lambda \leq \lambda_L$ 判断错误的概率。

此外,将上面诸式中的 $\lambda \leq \lambda_0$, $\lambda \geq \lambda_0$ 和 $\lambda_U \leq \lambda \leq \lambda_L$ 分别换成 $\lambda < \lambda_0$, $\lambda > \lambda_0$ 和 $\lambda_U < \lambda < \lambda_L$, 定理 2 仍然成立。

定理 2 是置信检验的理论基础,置信检验依据的是“大概率原则”,而显著性假设检验依据的是“小概率原则”,所以置信检验与假设检验有着本质的区别,它能够有效地克服假设检验的缺点和不足。

根据工程检验问题的需求,置信检验理论包括强检验、弱检验和弱强检验三种方式。

3 强检验理论

3.1 置信概率表征的强检验

3.1.1 单侧指标检验

对于规定的指标值 λ_0 和给定的置信度 Y , 根据定理 2 可按如下步骤对母体特征值 $\lambda \leq \lambda_0$ 还是

$\lambda \geq \lambda_0$ 进行检验:

(1) 首先从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量, 通常为 2 个试样。

(2) 由所获得的样本确定 λ 的置信分布函数 $F(\lambda)$, 并计算 $F(\lambda_0)$ 。

(3) 若 $F(\lambda_0) \geq \gamma$, 则可在置信度 γ 下断定 $\lambda \leq \lambda_0$ 。

(4) 若 $1 - F(\lambda_0) \geq \gamma$, 则可在置信度 γ 下断定 $\lambda \geq \lambda_0$ 。

(5) 若 $1 - \gamma < F(\lambda_0) < \gamma$, 则样本信息量太小, 无法在置信度 γ 下作出判断, 需要继续从母体中随机抽取一个试样, 并重复步骤(2)~(5)。

如果将 $\lambda \leq \lambda_0$ 和 $\lambda \geq \lambda_0$ 分别换成 $\lambda < \lambda_0$ 和 $\lambda > \lambda_0$, 那么上述强检验理论仍然成立。

3.1.2 双侧指标检验

对于规定的指标值 $\lambda_1, \lambda_2 (\lambda_1 < \lambda_2)$ 和给定的置信度 γ , 根据定理 2 可按如下步骤对母体特征值 λ 是否位于区间 $[\lambda_1, \lambda_2]$ 之内进行检验:

(1) 首先从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量。

(2) 由所获得的样本确定 λ 的置信分布函数 $F(\lambda)$, 并计算 $F(\lambda_1)$ 和 $F(\lambda_2)$ 。

(3) 若 $F(\lambda_2) - F(\lambda_1) \geq \gamma$, 则可在置信度 γ 下断定 λ 位于区间 $[\lambda_1, \lambda_2]$ 之内。

(4) 若 $1 - F(\lambda_2) + F(\lambda_1) \geq \gamma$, 则可在置信度 γ 下断定 λ 位于区间 $[\lambda_1, \lambda_2]$ 之外。

(5) 若 $1 - \gamma < F(\lambda_2) - F(\lambda_1) < \gamma$, 则样本信息量太小, 无法在置信度 γ 下作出判断, 需要从母体中再随机抽取一个试样, 并重复步骤(2)~(5)。

如果将区间 $[\lambda_1, \lambda_2]$ 换成区间 (λ_1, λ_2) , 那么上述强检验理论仍然成立。

3.2 置信限表征的强检验

对于规定的指标值 λ_0 和给定的置信度 γ , 根据定理 2 可按如下步骤对母体特征值 $\lambda \geq \lambda_0$ 还是 $\lambda \leq \lambda_0$ 进行检验:

(1) 首先从母体中随机抽取一个样本, 样本大小为确定 λ 的置信限所需的最小样本容量, 通常为 2 个试样。

(2) 由所获得的样本确定 λ 的置信度为 γ 的单侧置信上限 λ_u 和下限 λ_l 。

(3) 若 $\lambda_u \leq \lambda_0$, 则可在置信度 γ 下断定 $\lambda \leq \lambda_0$ 。

(4) 若 $\lambda_l \geq \lambda_0$, 则可在置信度 γ 下断定 $\lambda \geq \lambda_0$ 。

(5) 若 $\lambda_l < \lambda_0 < \lambda_u$, 则样本信息量太小, 无法在置信度 γ 下作出判断, 需要继续从母体中随机抽取一个试样, 并重复步骤(2)~(5)。

如果将 $\lambda_u \leq \lambda_0, \lambda_l \geq \lambda_0, \lambda_l < \lambda_0 < \lambda_u$ 和 $\lambda \leq \lambda_0, \lambda \geq \lambda_0$ 分别换成 $\lambda_u < \lambda_0, \lambda_l > \lambda_0, \lambda_l \leq \lambda_0 \leq \lambda_u$ 和 $\lambda < \lambda_0, \lambda > \lambda_0$, 那么上述强检验理论仍然成立。

4 弱检验理论

强检验是确认工程中所关心的“条件”是否成立的检验, 而弱检验则是确认所关心的“条件”是否被拒绝的检验。工程实践中, 常常根据所研究问题的物理性质、理论推导和实际经验等已基本确认某一“条件”成立, 现在只是希望验证一下该“条件”是否会被拒绝, 当不被拒绝时, 隐含着接受该“条件”; 此种情况可采用弱检验。单从检验的角度来说, 强检验是以高概率做出判断, 其结论是可信的, 而弱检验当某一“条件”不被拒绝时就接受它, 则缺乏足够的说服力, 这也是“弱”字的含义所在, 后面将看到它与传统的假设检验是等价的。

4.1 置信概率表征的弱检验

4.1.1 单指标弱检验

对于规定的指标值 λ_0 和给定的置信度 γ , 单指标弱检验可按如下步骤进行:

(1) 由随机抽取的样本确定 λ 的置信分布函数 $F(\lambda)$, 并计算 $F(\lambda_0)$ 。

(2) 若 $F(\lambda_0) > \gamma$, 则在置信度 γ 下拒绝条件 $\lambda \geq \lambda_0$ (即认为 $\lambda < \lambda_0$); 否则不能在置信度 γ 下拒绝条件 $\lambda \geq \lambda_0$ 。

(3) 若 $1 - F(\lambda_0) > \gamma$, 则在置信度 γ 下拒绝条件 $\lambda \leq \lambda_0$ (即认为 $\lambda > \lambda_0$); 否则不能在置信度 γ 下拒绝条件 $\lambda \leq \lambda_0$ 。

(4) 若 $|2F(\lambda_0) - 1| > \gamma$, 则在置信度 γ 下拒绝条件 $\lambda = \lambda_0$ (即认为 $\lambda \neq \lambda_0$); 否则不能在置信度 γ 下拒绝条件 $\lambda = \lambda_0$ 。

4.1.2 双侧指标弱检验

对于规定的指标值 $\lambda_1, \lambda_2 (\lambda_1 < \lambda_2)$ 和给定的置信度 γ , 双侧指标弱检验可按如下步骤进行:

(1) 由随机抽取的样本确定 λ 的置信分布函数 $F(\lambda)$, 并计算 $F(\lambda_1)$ 和 $F(\lambda_2)$ 。

(2) 若 $1 - F(\lambda_2) + F(\lambda_1) > \gamma$, 则在置信度 γ 下拒绝条件 $\lambda_1 \leq \lambda \leq \lambda_2$; 否则不能在置信度 γ 下拒绝条件 $\lambda_1 \leq \lambda \leq \lambda_2$ 。

(3) 若 $F(\lambda_2) - F(\lambda_1) > \gamma$, 则在置信度 γ 下拒绝条件 $\lambda \leq \lambda_1$ 和 $\lambda \geq \lambda_2$; 否则不能在置信度 γ 下拒

绝条件 $\lambda \leq \lambda_l$ 和 $\lambda \geq \lambda_u$ 。

4.2 置信限表征的弱检验

对于规定的指标值 λ_0 和给定的置信度 γ , 单指标弱检验可按如下步骤进行:

(1) 由随机抽取的样本确定 λ 的置信度为 γ 的单侧置信上限 λ_u 和下限 λ_l 。

(2) 若 $\lambda_u < \lambda_0$, 则在置信度 γ 下拒绝条件 $\lambda \geq \lambda_0$; 否则不能在置信度 γ 下拒绝条件 $\lambda \geq \lambda_0$ 。

(3) 若 $\lambda_l > \lambda_0$, 则在置信度 γ 下拒绝条件 $\lambda \leq \lambda_0$; 否则不能在置信度 γ 下拒绝条件 $\lambda \leq \lambda_0$ 。

(4) 若 $\lambda_l > \lambda_0$ 或 $\lambda_u < \lambda_0$, 则在置信度 $2\gamma - 1$ 下拒绝条件 $\lambda = \lambda_0$; 否则不能在置信度 $2\gamma - 1$ 下拒绝条件 $\lambda = \lambda_0$ 。

5 弱强检验理论

工程中, 有时对产品质量规定两个指标值 λ_0 和 λ_l , λ_0 被称为合格质量, λ_l 为极限质量。若在 λ_0 处采用弱检验(置信度为 γ), 而在 λ_l 处采用强检验(置信度为 γ_l), 则这种检验被称为弱强检验。

5.1 置信概率表征的弱强检验

5.1.1 下限弱强检验

设 $\lambda_0 > \lambda_l$, 当 $\lambda \geq \lambda_0$ 时认为该批产品合格, 且不允许 $\lambda \leq \lambda_l$; γ 和 γ_l 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信分布函数 $F(\lambda)$, 计算 $F(\lambda_0)$ 和 $F(\lambda_l)$, 并由下面两式:

$$F(\lambda_0, n_0, x) = \gamma \quad (9)$$

$$F(\lambda_l, n_0, x) = 1 - \gamma_l \quad (10)$$

联立消去其中的一个参数估计量(通常为 $\bar{x}$) 确定 n_0 。式中 $F(\lambda, n_0, x)$ 是 λ 的置信分布函数。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $1 - F(\lambda_0) > \gamma$ 且 $1 - F(\lambda_l) > \gamma_l$, 则认为该批产品合格; 若 $F(\lambda_0) > \gamma$ 或 $F(\lambda_l) > \gamma_l$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $F(\lambda_0) < \gamma$ 且 $F(\lambda_l) < 1 - \gamma_l$, 则认为该批产品合格; 否则认为该批产品不合格。

5.1.2 上限弱强检验

设 $\lambda_0 < \lambda_l$, 当 $\lambda \leq \lambda_0$ 时认为该批产品合格, 且不允许 $\lambda \geq \lambda_l$; γ 和 γ_l 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信分布函数 $F(\lambda)$, 计算 $F(\lambda_0)$ 和 $F(\lambda_l)$, 并由下面两式:

$$F(\lambda_0, n_0, x) = 1 - \gamma \quad (11)$$

$$F(\lambda_l, n_0, x) = \gamma_l \quad (12)$$

确定 n_0 。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $F(\lambda_0) > \gamma$ 且 $F(\lambda_l) > \gamma_l$, 则认为该批产品合格; 若 $1 - F(\lambda_0) > \gamma$ 或 $1 - F(\lambda_l) > \gamma_l$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $F(\lambda_0) > 1 - \gamma$ 且 $F(\lambda_l) > \gamma_l$, 则认为该批产品合格; 否则认为该批产品不合格。

5.1.3 双侧区间弱强检验

设 λ_l, λ_0 和 λ_u, λ_e 为规定的指标值, 其中 $\lambda_l' < \lambda_l < \lambda_0 < \lambda_e < \lambda_u'$, 当 $\lambda_l \leq \lambda \leq \lambda_e$ 时认为该批产品合格, 且不允许 $\lambda \leq \lambda_l$ 或 $\lambda \geq \lambda_u$; λ_l 和 λ_e 分别为合格质量的下限和上限, λ_l' 和 λ_u' 分别为极限质量的下限和上限; γ 和 γ_l 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信分布函数 $F(\lambda)$, 计算 $F(\lambda_l)$, $F(\lambda_e)$, $F(\lambda_l')$ 和 $F(\lambda_u')$, 并由下面两式:

$$F(\lambda_e, n_0, x) - F(\lambda_l, n_0, x) = 1 - \gamma \quad (13)$$

$$F(\lambda_u', n_0, x) - F(\lambda_l', n_0, x) = \gamma_l \quad (14)$$

确定 n_0 。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $F(\lambda_e) - F(\lambda_l) > \gamma$ 且 $F(\lambda_u') - F(\lambda_l') > \gamma_l$, 则认为该批产品合格; 若 $1 - F(\lambda_e) + F(\lambda_l) > \gamma$ 或 $1 - F(\lambda_u') + F(\lambda_l') > \gamma_l$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $F(\lambda_e) - F(\lambda_l) > 1 - \gamma$ 且 $F(\lambda_u') - F(\lambda_l') > \gamma_l$, 则认为该批产品合格; 否则认为该批产品不合格。

5.1.4 外侧区间弱强检验

设 λ_l, λ_0 和 λ_l', λ_e 为规定的指标值, 其中 $\lambda_l < \lambda_l' < \lambda_0 < \lambda_e$, 当 $\lambda \leq \lambda_l$ 或 $\lambda \geq \lambda_e$ 时认为该批产品合格, 且不允许 $\lambda_l' \leq \lambda \leq \lambda_e$; λ_l 和 λ_e 为合格质量, λ_l' 和 λ_e 为极限质量; γ 和 γ_l 为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信分布函数所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信分布函数 $F(\lambda)$, 计算 $F(\lambda_0)$, $F(\lambda_1)$, $F(\lambda_0)$ 和 $F(\lambda_1)$, 并由下面两式:

$$F(\lambda_0, n_0, x) - F(\lambda_1, n_0, x) = \gamma \quad (15)$$

$$F(\lambda_0', n_0, x) - F(\lambda_1', n_0, x) = 1 - \gamma_1 \quad (16)$$

确定 n_0 。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $1 - F(\lambda_0) + F(\lambda_1) > \gamma$ 且 $1 - F(\lambda_0') + F(\lambda_1') > \gamma_1$, 则认为该批产品合格; 若 $F(\lambda_0) - F(\lambda_1) > \gamma$ 或 $F(\lambda_0') - F(\lambda_1') > \gamma_1$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $F(\lambda_0) - F(\lambda_1) < \gamma$ 且 $F(\lambda_0') - F(\lambda_1') < 1 - \gamma_1$, 则认为该批产品合格; 否则认为该批产品不合格。

5.2 置信限表征的弱强检验

5.2.1 下限弱强检验

设 $\lambda_0 > \lambda_1$, 当 $\lambda \geq \lambda_0$ 时认为该批产品合格, 且不允许 $\lambda \leq \lambda_1$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信限所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信度为 γ 的单侧置信上限 λ_U 和下限 λ_L , 置信度为 γ_1 的单侧置信上限 λ_{U1} 和下限 λ_{L1} , 并由下面两式:

$$\lambda_U(n_0, x) = \lambda_0 \quad (17)$$

$$\lambda_{L1}(n_0, x) = \lambda_1 \quad (18)$$

联立消去其中的一个参数估计量(通常为 $\bar{x}$) 确定 n_0 。式中 $\lambda_U(n_0, x)$ 和 $\lambda_{L1}(n_0, x)$ 分别是 λ 的置信度为 γ 的单侧置信上限公式和置信度为 γ_1 的单侧置信下限公式。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $\lambda > \lambda_0$ 且 $\lambda_{L1} > \lambda_1$, 则认为该批产品合格; 若 $\lambda_U < \lambda_0$ 或 $\lambda_{U1} < \lambda_1$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $\lambda_U > \lambda_0$ 且 $\lambda_{L1} > \lambda_1$, 则认为该批产品合格; 否则认为该批产品不合格。

5.2.2 上限弱强检验

设 $\lambda_0 < \lambda_1$, 当 $\lambda \leq \lambda_0$ 时认为该批产品合格, 且不允许 $\lambda \geq \lambda_1$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 样本大小为确定 λ 的置信限所需的最小样本容量。

(2) 由所获得的样本 x 确定 λ 的置信度为 γ 的单侧置信上限 λ_U 和下限 λ_L , 置信度为 γ_1 的单侧置信上限 λ_{U1} 和下限 λ_{L1} , 并由下面两式:

$$\lambda_U(n_0, x) = \lambda_0 \quad (19)$$

$$\lambda_{U1}(n_0, x) = \lambda_1 \quad (20)$$

确定 n_0 。式中 $\lambda_U(n_0, x)$ 和 $\lambda_{U1}(n_0, x)$ 分别是 λ 的置信度为 γ 的单侧置信下限公式和置信度为 γ_1 的单侧置信上限公式。若实际抽取的试样总数 $n \geq n_0$, 则转至步骤(4)。

(3) 若 $\lambda_U < \lambda_0$ 且 $\lambda_{U1} < \lambda_1$, 则认为该批产品合格; 若 $\lambda > \lambda_0$ 或 $\lambda_{L1} > \lambda_1$, 则认为该批产品不合格; 否则, 继续从母体中随机抽取一个试样, 并返回步骤(2)。

(4) 若 $\lambda < \lambda_0$ 且 $\lambda_{L1} < \lambda_1$, 则认为该批产品合格; 否则认为该批产品不合格。

6 正态母体均值的置信检验

为了便于理解和掌握置信检验理论, 下面将以方差已知情况下正态母体均值的置信检验为例进行详细讨论。

设 $X \sim N(\mu, \sigma^2)$, 其中方差 σ^2 已知, 从母体 X 中随机抽取样本 $x = (x_1, x_2, \dots, x_n)$, 则均值 μ 的置信度为 γ 的单侧置信上限 μ_U 和下限 μ_L 为:

$$\mu_U = \bar{x} + u_\gamma \sigma / \sqrt{n} \quad (21)$$

$$\mu_L = \bar{x} - u_\gamma \sigma / \sqrt{n} \quad (22)$$

根据定理 1 可知, μ 的置信分布函数 $F(\mu)$ 由下式给出:

$$F(\mu) = \Phi[\sqrt{n}(\mu - \bar{x})/\sigma] \quad (23)$$

式中 u_γ 为标准正态偏量; $\bar{x}$ 为子样均值, 即:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (24)$$

6.1 均值的强检验

6.1.1 置信概率表征的强检验

6.1.1.1 单指标检验

对于规定的指标值 μ_0 和给定的置信度 γ , 可按如下步骤对均值 μ 进行强检验, 从第一个样本观测值 x_1 开始:

(1) 根据所获样本由式(23)计算 $F(\mu_0)$ 。

(2) 若 $F(\mu_0) \geq \gamma$, 则可在置信度 γ 下断定 $\mu \leq \mu_0$ 。

(3) 若 $1 - F(\mu_0) \geq \gamma$, 则可在置信度 γ 下断定

$$\mu \geq \mu_0$$

(4) 若 $1 - \gamma < F(\mu_0) < \gamma$, 则样本信息量太小, 无法在置信度 γ 下作出判断, 需继续从母体中随机抽取一个试样, 并重复步骤(1)~(4)。

6.1.1.2 双侧指标检验

对于规定的指标值 $\mu_1, \mu_2 (\mu_1 < \mu_2)$ 和给定的置信度 γ , 可按如下步骤对均值 μ 进行强检验, 从第一个样本观测值 x_1 开始:

(1) 根据所获得的样本由式(23)计算 $F(\mu_1)$ 和 $F(\mu_2)$ 。

(2) 若 $F(\mu_2) - F(\mu_1) \geq \gamma$, 则可在置信度 γ 下断定 $\mu_1 \leq \mu \leq \mu_2$ 。

(3) 若 $1 - F(\mu_2) + F(\mu_1) \geq \gamma$, 则可在置信度 γ 下断定 μ 位于区间 $[\mu_1, \mu_2]$ 之外。

(4) 若 $1 - \gamma < F(\mu_2) - F(\mu_1) < \gamma$, 则无法在置信度 γ 下作出判断, 需继续从母体中随机抽取一个试样, 并重复步骤(1)~(4)。

6.1.2 置信限表征的强检验

对于规定的指标值 μ_0 和给定的置信度 γ , 可按如下步骤对均值 μ 进行强检验, 从第一个样本观测值 x_1 开始:

(1) 根据所获样本由式(21)和(22)计算 μ 的置信度为 γ 的单侧置信上限 μ_U 和下限 μ_L 。

(2) 若 $\mu_U \leq \mu_0$, 则可在置信度 γ 下断定 $\mu \leq \mu_0$ 。

(3) 若 $\mu_L \geq \mu_0$, 则可在置信度 γ 下断定 $\mu \geq \mu_0$ 。

(4) 若 $\mu_L < \mu_0 < \mu_U$, 则样本信息量太小, 无法在置信度 γ 下作出判断, 需继续从母体中随机抽取一个试样, 并重复步骤(1)~(4)。

6.2 均值的弱检验

6.2.1 置信概率表征的弱检验

6.2.1.1 单指标弱检验

对于规定的指标值 μ_0 和给定的置信度 γ , 母体均值 μ 的单指标弱检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 并根据式(23)计算 $F(\mu_0)$ 。

(2) 若 $F(\mu_0) > \gamma$, 则在置信度 γ 下拒绝条件 $\mu \geq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \geq \mu_0$ 。

(3) 若 $1 - F(\mu_0) > \gamma$, 则在置信度 γ 下拒绝条件 $\mu \leq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \leq \mu_0$ 。

(4) 若 $|2F(\mu_0) - 1| > \gamma$, 则在置信度 γ 下拒

绝条件 $\mu = \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu = \mu_0$ 。

6.2.1.2 双侧指标弱检验

对于规定的指标值 $\mu_1, \mu_2 (\mu_1 < \mu_2)$ 和给定的置信度 γ , 母体均值 μ 的双侧指标弱检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 并根据式(23)计算 $F(\mu_1)$ 和 $F(\mu_2)$ 。

(2) 若 $1 - F(\mu_2) + F(\mu_1) > \gamma$, 则在置信度 γ 下拒绝条件 $\mu_1 \leq \mu \leq \mu_2$; 否则不能在置信度 γ 下拒绝条件 $\mu_1 \leq \mu \leq \mu_2$ 。

(3) 若 $F(\mu_2) - F(\mu_1) > \gamma$, 则在置信度 γ 下拒绝条件 $\mu \leq \mu_1$ 和 $\mu \geq \mu_2$; 否则不能在置信度 γ 下拒绝条件 $\mu \leq \mu_1$ 和 $\mu \geq \mu_2$ 。

6.2.2 置信限表征的弱检验

对于规定的指标值 μ_0 和给定的置信度 γ , 母体均值 μ 的单指标弱检验可按如下步骤进行:

(1) 从母体中随机抽取一个样本, 并根据式(21)和(22)计算 μ 的置信度为 γ 的单侧置信上限 μ_U 和下限 μ_L 。

(2) 若 $\mu_U < \mu_0$, 则在置信度 γ 下拒绝条件 $\mu \geq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \geq \mu_0$ 。

(3) 若 $\mu_L > \mu_0$, 则在置信度 γ 下拒绝条件 $\mu \leq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \leq \mu_0$ 。

(4) 若 $\mu_U < \mu_0$ 或 $\mu_L > \mu_0$, 则在置信度 $2\gamma - 1$ 下拒绝条件 $\mu = \mu_0$; 否则不能在置信度 $2\gamma - 1$ 下拒绝条件 $\mu = \mu_0$ 。

6.3 均值的弱强检验

6.3.1 置信概率表征的弱强检验

6.3.1.1 下限弱强检验

设 $\mu_0 > \mu_1$, 当 $\mu \geq \mu_0$ 时认为该批产品合格, 且不允许 $\mu \leq \mu_1$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 由下面两式:

$$\Phi[\sqrt{n_0}(\mu_0 - \bar{x})/\sigma] = \gamma \quad (25)$$

$$\Phi[\sqrt{n_0}(\mu_1 - \bar{x})/\sigma] = 1 - \gamma_1 \quad (26)$$

联立消去 $\bar{x}$ 求得:

$$n_0 = [(u\gamma + u\gamma_1)/(\mu_1 - \mu_0)]^2 \sigma^2 \quad (27)$$

(2) 从母体中随机抽取 $n (n \geq n_0)$ 个试样, 根据式(23)计算 $F(\mu_0)$ 和 $F(\mu_1)$ 。

(3) 若 $F(\mu_0) < \gamma$ 且 $F(\mu_1) < 1 - \gamma_1$, 则认为该批产品合格; 否则认为该批产品不合格。

μ 的上限弱强检验可类似进行讨论(略)。

6.3.1.2 双侧区间弱强检验

设 μ_1, μ_2 和 μ_1', μ_2' 为规定的指标值, 其中 $\mu_1' < \mu_1 < \mu_2 < \mu_2'$, 当 $\mu_1 \leq \mu \leq \mu_2$ 时认为该批产品合格, 且不允许 $\mu \leq \mu_1$ 或 $\mu \geq \mu_2$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 由下面两式:

$$\Phi\left(\frac{\mu_2 - \bar{X}}{\sigma/\sqrt{n_0}}\right) - \Phi\left(\frac{\mu_1 - \bar{X}}{\sigma/\sqrt{n_0}}\right) = 1 - \gamma \quad (28)$$

$$\Phi\left(\frac{\mu_2' - \bar{X}}{\sigma/\sqrt{n_0}}\right) - \Phi\left(\frac{\mu_1' - \bar{X}}{\sigma/\sqrt{n_0}}\right) = \gamma_1 \quad (29)$$

确定 n_0 。

(2) 从母体中随机抽取 $n (n \geq n_0)$ 个试样, 由式

(23) 计算 $F(\mu_1), F(\mu_2)$ 和 $F(\mu_1'), F(\mu_2')$ 。

(3) 若 $F(\mu_2) - F(\mu_1) > 1 - \gamma$ 且 $F(\mu_2') - F(\mu_1') > \gamma_1$, 则认为该批产品合格; 否则认为该批产品不合格。

μ 的外侧区间弱强检验可类似进行讨论。

6.3.2 置信限表征的弱强检验

6.3.2.1 下限弱强检验

设 $\mu_0 > \mu_1$, 当 $\mu \geq \mu_0$ 时认为该批产品合格, 且不允许 $\mu \leq \mu_1$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 由下面两式:

$$\mu_0 = \bar{X} + u_\gamma \sigma / \sqrt{n_0} \quad (30)$$

$$\mu_1 = \bar{X} - u_{\gamma_1} \sigma / \sqrt{n_0} \quad (31)$$

联立消去 $\bar{X}$ 求得:

$$n_0 = [(u_\gamma + u_{\gamma_1}) / (\mu_1 - \mu_0)]^2 \sigma^2 \quad (32)$$

(2) 从母体中随机抽取 $n (n \geq n_0)$ 个试样, 由式

(21) 和 (22) 计算均值 μ 的置信度为 γ 的单侧置信上限 μ_U 和置信度为 γ_1 的单侧置信下限 μ_{L1} 。

(3) 若 $\mu_U > \mu_0$ 且 $\mu_{L1} > \mu_1$, 则认为该批产品合格; 否则认为该批产品不合格。

6.3.2.2 上限弱强检验

设 $\mu_0 < \mu_1$, 当 $\mu \leq \mu_0$ 时认为该批产品合格, 且不允许 $\mu \geq \mu_1$; γ 和 γ_1 分别为给定的置信度。该问题的弱强检验可按如下步骤进行:

(1) 由下面两式

$$\mu_0 = \bar{X} - u_\gamma \sigma / \sqrt{n_0} \quad (33)$$

$$\mu_1 = \bar{X} + u_{\gamma_1} \sigma / \sqrt{n_0} \quad (34)$$

联立求得

$$n_0 = [(u_\gamma + u_{\gamma_1}) / (\mu_1 - \mu_0)]^2 \sigma^2 \quad (35)$$

(2) 从母体中随机抽取 $n (n \geq n_0)$ 个试样, 由式 (21) 和 (22) 计算均值 μ 的置信度为 γ 的单侧置信下限 μ_L 和置信度为 γ_1 的单侧置信上限 μ_{U1} 。

(3) 若 $\mu_L < \mu_0$ 且 $\mu_{U1} < \mu_1$, 则认为该批产品合格; 否则认为该批产品不合格。

7 弱检验与假设检验的关系

弱检验等价于显著性假设检验, 表 1 以正态母体均值(方差已知)的检验为例进行了对比说明。从中可以看到弱检验完全具有显著性假设检验的功能, 而且在实际计算和使用中更加简单方便, 许多显著性假设检验难以进行的问题, 采用弱检验能够容易地得到解决, 例如百分率和百分位值的检验, 详见文献[4]中例 1。

8 弱强检验与抽样检验的关系

弱强检验等价于抽样检验, 表 2 对正态母体均值(方差已知)的抽样检验和弱强检验进行了对比, 从中可以看到弱强检验完全具有抽样检验的功能, 而且弱强检验在合格质量处是弱检验, 在极限质量处是强检验, 检验的性质非常明确, 虽然抽样检验也是如此, 但何处“弱”, 何处“强”, 混淆不清。此外, 在实际计算和使用中弱强检验也要比抽样检验更加方便, 在抽样检验中许多复杂和困难的问题, 在弱强检验中将变得简单易行。

9 结 论

(1) 提出了母体特征值(如母体均值、方差、百分位值和百分率等)置信分布的概念, 为置信检验奠定了理论基础。置信分布实质上是一种特殊的 Fiducial 分布, 其求得的置信限与 Neyman 意义下的置信限是相同的, 从而使置信度既能用频率解释, 又在一次抽样中有明确的意义。

(2) 建立了依据“大概率原则”的置信检验理论, 包括强检验、弱检验和弱强检验三种方法, 分别满足了不同工程检验的需求。

(3) 强检验能够以高概率确认工程中所关心的“条件”是否成立, 而且当试样较少时也能作出准确判断, 并具有序贯抽样的优点。与假设检验和抽样检验相比, 强检验能够大大提高工程检验精度。

(4) 弱检验属于拒绝检验, 等价于显著性假设检验, 但是其检验的性质更加清楚, 明确地指出

了当某一“条件”不被拒绝时就接受它的理由是不充分的。在使用中弱检验也更加简单方便,许多显著性假设检验难以进行的问题,采用弱检验能容易地得到解决,例如百分率和百分位值的检验。

(5) 弱强检验等价于抽样检验,弱强检验在合格质量处是弱检验,在极限质量处是强检验,检验的性质非常明确,虽然抽样检验也是如此,但何处“弱”,何处“强”,混淆不清。与抽样检验相比,弱强检验还具有序贯抽样的优点,常常在试样较少时就能作出高置信度的判断,可以节省大量试样。

对于工程上要求预先确定样本量的情况,如果 n_0 不含其他参数估计量,则 n_0 即为需抽取的最小样本量;如果 n_0 含有其他参数估计量,则可先根据历史数据计算这些参数估计量,也可直接估算其均值、上限或下限,然后再确定样本量 n_{00} 。此外,弱强检验中的 $1-\gamma$ 和 $1-\gamma_0$ 与抽样检验中的生产方风险 $\alpha=1-\gamma$ 和使用方风险 $\beta=1-\gamma_0$ 相同。

(6) 置信检验不但对正态分布,而且对任意分布母体特征值的检验都简单易行。

表 1 正态母体均值(方差已知)的弱检验和显著性假设检验对比表

Table 1 Comparison between weak test and sign if cance hypothesis test for the mean of normal population

条件	弱检验	显著性假设检验
$\mu \geq \mu_0$	若 $\mu_0 = \bar{x} - u\gamma\sigma/\sqrt{n} < \mu_0$, 则在置信度 γ 下拒绝条件 $\mu \geq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \geq \mu_0$ 。	若 $\bar{x} < \mu_0 - u\gamma\sigma/\sqrt{n}$, 则 显著性水平 $\alpha=1-\gamma$ 下拒绝条件 $\mu \geq \mu_0$; 否则不能在显著性水平 $\alpha=1-\gamma$ 下拒绝条件 $\mu \geq \mu_0$ 。
$\mu \leq \mu_0$	若 $\mu_0 = \bar{x} - u\gamma\sigma/\sqrt{n} > \mu_0$, 则在置信度 γ 下拒绝条件 $\mu \leq \mu_0$; 否则不能在置信度 γ 下拒绝条件 $\mu \leq \mu_0$ 。	若 $\bar{x} > \mu_0 + u\gamma\sigma/\sqrt{n}$, 则在显著性水平 $\alpha=1-\gamma$ 下拒绝条件 $\mu \leq \mu_0$; 否则不能在显著性水平 $\alpha=1-\gamma$ 下拒绝条件 $\mu \leq \mu_0$ 。
$\mu = \mu_0$	若 $\mu_0 = \bar{x} + u\gamma\sigma/\sqrt{n} < \mu_0$ 或 $\mu_0 = \bar{x} - u\gamma\sigma/\sqrt{n} > \mu_0$, 则在置信度 $2\gamma-1$ 下拒绝条件 $\mu = \mu_0$; 否则不能在置信度 $2\gamma-1$ 下拒绝条件 $\mu = \mu_0$ 。	若 $\bar{x} < \mu_0 - u\gamma\sigma/\sqrt{n}$ 或 $\bar{x} > \mu_0 + u\gamma\sigma/\sqrt{n}$, 则在显著性水平 $\alpha=2(1-\gamma)=1-(2\gamma-1)$ 下拒绝条件 $\mu = \mu_0$; 否则不能在显著性水平 $\alpha=1-(2\gamma-1)$ 下拒绝条件 $\mu = \mu_0$ 。

表 2 正态母体均值(方差已知)的弱强检验和抽样检验对比表

Table 2 Comparison between weak-strong test and sampling inspection for the mean of normal population

条件	弱强检验	抽样检验
当 $\mu \geq \mu_0$ 时认为该批产品合格,且不允许 $\mu \leq \mu_0$ ($\mu_0 > \mu_1$)	(1) 样本量 $n_0 = [(u\gamma + u\gamma_1)/(\mu_1 - \mu_0)]^2 \sigma^2$ 。 (2) 若 $\mu_0 = \bar{x} + u\gamma\sigma/\sqrt{n_0} > \mu_0$ 和 $\mu_1 = \bar{x} - u\gamma_1\sigma/\sqrt{n_0} > \mu_1$ 同时成立,即 $\bar{x} > (\mu_1 u\gamma + \mu_0 u\gamma_1)/(u\gamma + u\gamma_1)$, 则认为该批产品合格; 否则认为该批产品不合格。 (3) 生产方风险和使用方风险分别为 $1-\gamma$ 和 $1-\gamma_0$ 。	(1) 样本量 $n_0 = [(u\gamma + u\gamma_1)/(\mu_1 - \mu_0)]^2 \sigma^2$ 。 (2) 若 $\bar{x} > (\mu_1 u\gamma + \mu_0 u\gamma_1)/(u\gamma + u\gamma_1)$, 则认为该批产品合格; 否则认为该批产品不合格。 (3) 生产方风险和使用方风险分别为 $\alpha=1-\gamma$ 和 $\beta=1-\gamma_0$ 。
当 $\mu \leq \mu_0$ 时认为该批产品合格,且不允许 $\mu \geq \mu_0$ ($\mu_0 < \mu_1$)	(1) 样本量 $n_0 = [(u\gamma + u\gamma_1)/(\mu_1 - \mu_0)]^2 \sigma^2$ 。 (2) 若 $\mu_0 = \bar{x} - u\gamma\sigma/\sqrt{n_0} < \mu_0$ 和 $\mu_1 = \bar{x} + u\gamma_1\sigma/\sqrt{n_0} < \mu_1$ 同时成立,即 $\bar{x} < (\mu_1 u\gamma + \mu_0 u\gamma_1)/(u\gamma + u\gamma_1)$, 则认为该批产品合格; 否则认为该批产品不合格。 (3) 生产方风险和使用方风险分别为 $1-\gamma$ 和 $1-\gamma_0$ 。	(1) 样本量 $n_0 = [(u\gamma + u\gamma_1)/(\mu_1 - \mu_0)]^2 \sigma^2$ 。 (2) 若 $\bar{x} < (\mu_1 u\gamma + \mu_0 u\gamma_1)/(u\gamma + u\gamma_1)$, 则认为该批产品合格; 否则认为该批产品不合格。 (3) 生产方风险和使用方风险分别为 $\alpha=1-\gamma$ 和 $\beta=1-\gamma_0$ 。

参考文献

[1] 茆诗松,王静龙,濮晓龙 高等数理统计[M] 北京:高等教育出版社;海德堡:施普林格出版社,2000

[2] Acheson J Duncan Quality Control and Industrial Statistics [M] Fifth Edition, Homewood, Illinois: Richard D. Irwin, NC, 1986

[3] 于善奇 抽样检验与质量控制[M] 北京:北京大学出版社,1991

[4] 傅惠民 可靠性假设检验[J] 航空学报,1992,13(7):A 365 - A 370

[5] 傅惠民 连续母体百分位值和百分率的区间估计和假设检验[J] 航空学报,1993,14(11):A 569- A 577.

[6] 傅惠民 百分回归分析[J] 航空学报,1994,15(2): 141- 148