

文章编号: 1000-8055(2005)03-0345-06

多因子异方差分析

傅惠民, 邬文娟

(北京航空航天大学 固体力学研究所, 北京 100083)

摘要: 提出一种多因子异方差分析方法, 在变异系数相同的条件下, 详细讨论了单因子、双因子以及多因子异方差分析。该方法可以对多个正态母体是否具有某种不同的均值进行检验和估计。并且通过少量试验, 对仿真和数字化设计结果进行检验, 分析各个因素(因子)对仿真和数字化设计结果的影响, 以确定哪些因素已被正确仿真, 而哪些因素尚未被正确模拟, 从而指导仿真软件的编制。此外, 该方法还可以根据试验数据对仿真和数字化设计结果进行修正。

关键词: 航空、航天推进系统; 方差分析; 因子分析; 小样本; 检验; 仿真; 数字化设计
中图分类号: TB301; O211 **文献标识码:** A

Multi-Way Heteroscedastic Analysis

FU Huimin, WU Wenjuan

(Institute of Solid Mechanics,

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

Abstract: A method for multiway heteroscedastic analysis is presented, which detailedly discusses the heteroscedastic analysis of one, two and multiple factors, in the condition of the same coefficient of variation in all states. It can test whether the multiple populations have a certain kind of different means, and estimate their parameters. With small sample, it can test the results of simulation or digital design and analyze the influence of each factor in the results. And then, it can point out which factors have been simulated correctly and which haven't. Furthermore, this method can correct the results of simulation or digital design according to the experimental data.

Key words: aerospace propulsion system; variance analysis; factor analysis; small sample; test; simulation; digital design

文献[1]提出一种能够对仿真和数字化设计结果进行小样本检验和修正的方法, 解决了仿真和数字化设计结果的可信性问题。鉴于核仿真、结构强度仿真、振动仿真、可靠性仿真、电磁仿真、经济仿真等需要对众多复杂的影响因素(如温度、应力等)进行模拟, 如何判断这些影响因素哪些已被正确仿真, 而哪些尚未被正确模拟, 这是当前亟待

解决的难题。为此, 本文进一步建立多因子异方差分析方法, 该方法可以用于仿真和数字化设计结果的因子分析, 能够对各个影响因素逐一进行检验, 以指导仿真软件的编制。

传统的多因子方差分析是在方差相同的条件下, 对多个母体的均值是否相同(即各因子及其交互作用是否对均值有显著影响)进行检验^[2~4]; 而

收稿日期: 2005-04-11; 修订日期: 2005-05-12

基金项目: 北京市重点学科工程力学项目(XK100060418)

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

多因子异方差分析方法则是在变异系数相同的条件下, 检验多个母体是否具有某种不同的均值(即在仿真和数字化设计中是否已正确考虑各因子及其交互作用对均值的影响), 所以它们在功能上完全不同。

1 单因子异方差分析方法

1.1 基本思想

设 $A_1, A_2, \dots, A_m$ 是因子 A 的 m 个水平, 在第 i 个水平 A_i 下进行了 n_i 次试验, 试验值 X_{ij} 满足:

$$X_{ij} \sim N(\mu_i, \sigma_i^2) \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n_i \quad (1)$$

式中:

$$\mu_i = k_i \mu \quad i = 1, 2, \dots, m \quad (2)$$

并假定各个不同因子水平下的变异系数 $\sigma_i/\mu_i (i=1, 2, \dots, m)$ 相同(在误差分析中相当于相对误差相同), 则有:

$$\sigma_i = k_i \sigma \quad i = 1, 2, \dots, m \quad (3)$$

式中 μ 和 σ 均为未知参数。若令:

$$Y_{ij} = \frac{X_{ij}}{k_i} \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n_i \quad (4)$$

则当仿真、数字化设计等方法求得的 $k_i (i=1, 2, \dots, m)$ 满足式(2)和式(3)时, 有:

$$Y_{ij} \sim N(\mu, \sigma^2) \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n_i \quad (5)$$

因此, 通过检验 Y_{ij} 是否来自于同一母体可以判断仿真、数字化设计等方法求得的 k_i 正确与否。

1.2 单因子分析方法

下面给出检验仿真、数字化设计得到的 $k_i (i=1, 2, \dots, m)$ 是否满足式(2)的方法, 并对未知参数 μ 和 σ^2 进行估计。令:

$$\mu = \frac{1}{n} \sum_{i=1}^m \frac{n_i \mu_i}{k_i} \quad (6)$$

式中 $n = \sum_{i=1}^m n_i$ 。设:

$$\delta_i = \mu_i/k_i - \mu \quad i = 1, 2, \dots, m \quad (7)$$

则检验 $k_i (i=1, 2, \dots, m)$ 是否满足式(2), 相当于检验假设:

$$\left. \begin{aligned} H_0: & \delta_1 = \delta_2 = \dots = \delta_m = 0 \\ H_1: & \delta_1, \delta_2, \dots, \delta_m \text{ 不全为零} \end{aligned} \right\} \quad (8)$$

单因子异方差分析的统计模型为:

$$\left. \begin{aligned} X_{ij} &= k_i \mu + k_i \delta_i + \epsilon_j \\ \epsilon_j &\sim N(0, \sigma^2), \text{ 各 } \epsilon_j \text{ 独立} \\ i &= 1, 2, \dots, m; \quad j = 1, 2, \dots, n_i \\ \sum_{i=1}^m n_i \delta_i &= 0 \end{aligned} \right\} \quad (9)$$

设总离差平方和为:

$$S_T = \sum_{i=1}^m \sum_{j=1}^{n_i} (X_{ij}/k_i - \bar{X})^2 \quad (10)$$

组内平方和(误差平方和)为:

$$S_E = \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{1}{k_i^2} (X_{ij} - \bar{X}_{i\cdot})^2 \quad (11)$$

组间平方和为:

$$S_A = \sum_{i=1}^m n_i (\bar{X}_{i\cdot}/k_i - \bar{X})^2 \quad (12)$$

并且, 有:

$$S_T = S_E + S_A \quad (13)$$

式中 $\bar{X}$ 为样本总平均:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{X_{ij}}{k_i} \quad (14)$$

$\bar{X}_{i\cdot}$ 为第 i 个水平 A_i 下的样本平均值:

$$\bar{X}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \quad (15)$$

可以证明:

$$\frac{S_E}{\sigma^2} \sim \chi^2(n-m) \quad (16)$$

并且 $E(S_A) = (m-1)\sigma^2 + \sum_{i=1}^m n_i \delta_i^2$, 所以, 当 H_0 为真时, 有:

$$\frac{S_A}{\sigma^2} \sim \chi^2(m-1) \quad (17)$$

由式(16)和(17)可知:

$$\frac{S_A/(m-1)}{S_E/(n-m)} \sim F(m-1, n-m) \quad (18)$$

因此, 对于给定的显著性水平 α , 假设 H_0 的拒绝域为:

$$F = \frac{S_A/(m-1)}{S_E/(n-m)} \geq F_\alpha(m-1, n-m) \quad (19)$$

未知参数 σ^2 的估计量为:

$$\hat{\sigma}^2 = \frac{S_E}{n-m} \quad (20)$$

当不拒绝假设 H_0 时, μ 的估计量为:

$$\hat{\mu} = \bar{X} \quad (21)$$

2 双因子异方差分析方法

2.1 双因子异方差分析

设 A 和 B 是影响试验结果的两个因子, $A_1,$

$A_2, \dots, A_r$ 是因子 A 的 r 个水平, $B_1, B_2, \dots, B_s$ 是因子 B 的 s 个水平。在每对因子水平组合 (A_i, B_j) 下进行了 t ($t \geq 2$) 次试验 (称为等重复试验), 试验值记为 X_{ijk} , $i = 1, 2, \dots, r$; $j = 1, 2, \dots, s$; $k = 1, 2, \dots, t$ 并设 $X_{ijk} \sim N(\mu_{ij}, \sigma_{ij}^2)$, 其中:

$$\mu_{ij} = k_{ij} \mu$$

$$i = 1, 2, \dots, r; \quad j = 1, 2, \dots, s \quad (22)$$

在各个不同因子水平下的变异系数相同时, 有:

$$\sigma_{ij} = k_{ij} \sigma$$

$$i = 1, 2, \dots, r; \quad j = 1, 2, \dots, s \quad (23)$$

式中 μ 和 σ^2 均为未知参数。下面将对仿真、数字化设计等方法求得的 k_{ij} ($i = 1, 2, \dots, r$; $j = 1, 2, \dots, s$) 是否满足式(22)进行检验。令:

$$\mu = \frac{1}{rs} \sum_{i=1}^r \sum_{j=1}^s \frac{\mu_{ij}}{k_{ij}} \quad (24)$$

$$\mu_{i\cdot} = \frac{1}{s} \sum_{j=1}^s \frac{\mu_{ij}}{k_{ij}} \quad i = 1, 2, \dots, r \quad (25)$$

$$\mu_{\cdot j} = \frac{1}{r} \sum_{i=1}^r \frac{\mu_{ij}}{k_{ij}} \quad j = 1, 2, \dots, s \quad (26)$$

$$\alpha_i = \mu_{i\cdot} - \mu \quad i = 1, 2, \dots, r \quad (27)$$

$$\beta_j = \mu_{\cdot j} - \mu \quad j = 1, 2, \dots, s \quad (28)$$

$$\gamma_{ij} = \mu_{ij}/k_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu \quad i = 1, 2, \dots, r; \quad j = 1, 2, \dots, s \quad (29)$$

则检验由仿真、数字化设计等方法求得的 k_{ij} ($i = 1, 2, \dots, r$; $j = 1, 2, \dots, s$) 是否满足式(22), 相当于检验假设:

$$H_{01}: \alpha_1 = \alpha_2 = \dots = \alpha_r = 0 \quad (30)$$

$$H_{11}: \alpha_1, \alpha_2, \dots, \alpha_r \text{ 不全为零}$$

$$H_{02}: \beta_1 = \beta_2 = \dots = \beta_s = 0 \quad (31)$$

$$H_{12}: \beta_1, \beta_2, \dots, \beta_s \text{ 不全为零}$$

$$H_{03}: \gamma_{11} = \gamma_{12} = \dots = \gamma_{rs} = 0 \quad (32)$$

$$H_{13}: \gamma_{11}, \gamma_{12}, \dots, \gamma_{rs} \text{ 不全为零}$$

其双因子异方差分析的统计模型为:

$$\left. \begin{aligned} X_{ijk} &= k_{ij} \mu + k_{ij} \alpha_i + k_{ij} \beta_j + k_{ij} \gamma_{ij} + \epsilon_{ijk} \\ \epsilon_{ijk} &\sim N(0, \sigma_{ij}^2), \text{ 各 } \epsilon_{ijk} \text{ 独立} \\ i &= 1, 2, \dots, r; \quad j = 1, 2, \dots, s; \quad k = 1, 2, \dots, t \\ \sum_{i=1}^r \alpha_i &= 0; \quad \sum_{j=1}^s \beta_j = 0; \quad \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} = 0; \quad \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t \gamma_{ij} = 0 \end{aligned} \right\} \quad (33)$$

设总离差平方和为:

$$S_T = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t (X_{ijk}/k_{ij} - \bar{X})^2 \quad (34)$$

误差平方和为:

$$S_E = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t \frac{1}{k_{ij}} (X_{ijk} - \bar{X}_{ij\cdot})^2 \quad (35)$$

因子 A 效应平方和为:

$$S_A = st \sum_{i=1}^r (\bar{X}_{i\cdot\cdot} - \bar{X})^2 \quad (36)$$

因子 B 效应平方和为:

$$S_B = rt \sum_{j=1}^s (\bar{X}_{\cdot j\cdot} - \bar{X})^2 \quad (37)$$

A 和 B 交互效应平方和为:

$$S_{A \times B} = t \sum_{i=1}^r \sum_{j=1}^s (\bar{X}_{ij\cdot}/k_{ij} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot} + \bar{X})^2 \quad (38)$$

则有:

$$S_T = S_E + S_A + S_B + S_{A \times B} \quad (39)$$

式中:

$$\bar{X} = \frac{1}{rst} \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t \frac{X_{ijk}}{k_{ij}} \quad (40)$$

$$\bar{X}_{i\cdot\cdot} = \frac{1}{st} \sum_{j=1}^s \sum_{k=1}^t \frac{X_{ijk}}{k_{ij}} \quad (41)$$

$$\bar{X}_{\cdot j\cdot} = \frac{1}{rt} \sum_{i=1}^r \sum_{k=1}^t \frac{X_{ijk}}{k_{ij}} \quad (42)$$

$$\bar{X}_{ij\cdot} = \frac{1}{t} \sum_{k=1}^t X_{ijk} \quad (43)$$

可以证明:

$$\frac{S_E}{\sigma^2} \sim \chi^2(rs(t-1)) \quad (44)$$

由于 $E(S_A) = (r-1)\sigma^2 + st \sum_{i=1}^r \alpha_i^2$, 故当 H_{01} 为真时, 有:

$$\frac{S_A}{\sigma^2} \sim \chi^2(r-1) \quad (45)$$

因 $E(S_B) = (s-1)\sigma^2 + rt \sum_{j=1}^s \beta_j^2$, 故当 H_{02} 为真时, 有:

$$\frac{S_B}{\sigma^2} \sim \chi^2(s-1) \quad (46)$$

同样有 $E(S_{A \times B}) = (r-1)(s-1)\sigma^2 + t \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij}^2$, 当 H_{03} 为真时, 有:

$$\frac{S_{A \times B}}{\sigma^2} \sim \chi^2((r-1)(s-1)) \quad (47)$$

由此可知, 在显著性水平 α 下, 假设 H_{01} 的拒绝域为:

$$F_A = \frac{S_A/(r-1)}{S_E/[rs(t-1)]} \geq F_{\alpha}(r-1, rs(t-1)) \quad (48)$$

假设 H_{02} 的拒绝域为:

$$F_B = \frac{S_B/(s-1)}{S_E/[rs(t-1)]} \geq F_{\alpha}(s-1, rs(t-1)) \quad (49)$$

假设 H_{03} 的拒绝域为:

$$F_{A \times B} = \frac{S_{A \times B} / [(r-1)(s-1)]}{S_E / [rs(t-1)]} \geq F_{\alpha} \left[(r-1)(s-1), rs(t-1) \right] \quad (50)$$

若 H_{01} (或 H_{02})被拒绝,则说明由仿真、数字化设计等方法求 k_{ij} ($i=1,2,\dots,r; j=1,2,\dots,s$)时,未正确考虑因子A(或B)的影响。若 H_{03} 被拒绝,则说明未正确考虑因子A和B的交互效应。

未知参数 σ^2 的估计量为:

$$\hat{\sigma}^2 = \frac{S_E}{rs(t-1)} \quad (51)$$

当不拒绝假设 H_{01} 、 H_{02} 和 H_{03} 时, μ 的估计量为 $\hat{\mu} = \bar{X}$ 。

2.2 双因子无重复试验的异方差分析

当因子A和B无交互作用或交互作用对试验结果的影响很小时,在每对水平组合(A_i, B_j)下可以只进行一次试验(称为无重复试验),试验值记为 X_{ij} ,设 $X_{ij} \sim N(\mu_{ij}, \sigma_{ij}^2)$, $i=1,2,\dots,r; j=1,2,\dots,s$ 。其异方差分析模型为:

$$\left. \begin{aligned} X_{ij} &= k_{ij}\mu + k_{ij}\alpha + k_{ij}\beta_j + \epsilon_{ij} \\ \epsilon_{ij} &\sim N(0, \sigma_{ij}^2), \text{各 } \epsilon_{ij} \text{ 独立} \\ i &= 1, 2, \dots, r; j = 1, 2, \dots, s \\ \sum_{j=1}^s \alpha &= 0; \sum_{i=1}^r \beta_j = 0 \end{aligned} \right\} \quad (52)$$

式中 $\mu, \mu_i, \mu_j, \alpha$ 和 β_j 仍由式(24)~(28)给出。因此,检验由仿真、数字化设计等方法求得的 k_{ij} ($i=1,2,\dots,r; j=1,2,\dots,s$)是否满足式(22),相当于检验式(30)和(31)给出的假设 H_{01} 和 H_{02} 。

设总离差平方和为:

$$S_T = \sum_{i=1}^r \sum_{j=1}^s (X_{ij}/k_{ij} - \bar{X})^2 \quad (53)$$

误差平方和为:

$$S_E = \sum_{i=1}^r \sum_{j=1}^s (X_{ij}/k_{ij} - \bar{X}_{i\cdot} - \bar{X}_{\cdot j} + \bar{X})^2 \quad (54)$$

因子A效应平方和为:

$$S_A = s \sum_{i=1}^r (\bar{X}_{i\cdot} - \bar{X})^2 \quad (55)$$

因子B效应平方和为:

$$S_B = r \sum_{j=1}^s (\bar{X}_{\cdot j} - \bar{X})^2 \quad (56)$$

则有:

$$S_T = S_E + S_A + S_B \quad (57)$$

式中:

$$\bar{X} = \frac{1}{rs} \sum_{i=1}^r \sum_{j=1}^s \frac{X_{ij}}{k_{ij}} \quad (58)$$

$$\bar{X}_{i\cdot} = \frac{1}{s} \sum_{j=1}^s \frac{X_{ij}}{k_{ij}} \quad (59)$$

$$\bar{X}_{\cdot j} = \frac{1}{r} \sum_{i=1}^r \frac{X_{ij}}{k_{ij}} \quad (60)$$

可以证明:

$$\frac{S_E}{\sigma^2} \sim \chi^2 \left[(r-1)(s-1) \right] \quad (61)$$

S_A 和 S_B 仍分别满足式(45)和(46)。

因此,当因子A和B无交互作用或交互作用对试验结果的影响很小时,在显著性水平 α 下,假设 H_{01} 的拒绝域为:

$$F_A = \frac{S_A / (r-1)}{S_E / [(r-1)(s-1)]} \geq F_{\alpha} \left[r-1, (r-1)(s-1) \right] \quad (62)$$

假设 H_{02} 的拒绝域为:

$$F_B = \frac{S_B / (s-1)}{S_E / [(r-1)(s-1)]} \geq F_{\alpha} \left[s-1, (r-1)(s-1) \right] \quad (63)$$

未知参数 σ^2 的估计量为:

$$\hat{\sigma}^2 = \frac{S_E}{(r-1)(s-1)} \quad (64)$$

当不拒绝假设 H_{01} 和 H_{02} 时, μ 的估计量为 $\hat{\mu} = \bar{X}$ 。

3 多因子异方差分析方法

对于多因子试验,设试验总次数为 n ,试验值 X_i 满足:

$$X_i \sim N(\mu_i, \sigma_i^2) \quad i=1,2,\dots,n \quad (65)$$

式中:

$$\mu_i = k_i \mu \quad i=1,2,\dots,n \quad (66)$$

在变异系数 σ_i/μ_i ($i=1,2,\dots,n$)相同的条件下,有:

$$\sigma_i = k_i \sigma \quad i=1,2,\dots,n \quad (67)$$

式中 μ 和 σ^2 均为未知参数。下面给出检验仿真、数字化设计等方法求得的 k_i ($i=1,2,\dots,n$)是否满足式(66)的方法。

3.1 总离差平方和

设总离差平方和为:

$$S_T = \sum_{i=1}^n (X_i/k_i - \bar{X})^2 \quad (68)$$

即:

$$S_T = Q_T - T^2/n \quad (69)$$

式中 Q_T 为所有试验数据的加权平方和, T 为所有试验数据的加权总和, $\bar{X}$ 为其加权总平均值:

$$Q_T = \sum_{i=1}^n X_i^2/k_i^2 \quad (70)$$

$$T = \sum_{i=1}^n X_i / k_i \tag{71}$$

$$\bar{X} = T / n \tag{72}$$

S_T 的自由度为:

$$f_T = n - 1 \tag{73}$$

可以证明, 当 $k_i (i=1, 2, \cdots, n)$ 满足式(66)和(67)时, 有:

$$\frac{S_T}{\sigma^2} \sim \chi^2(f_T) \tag{74}$$

3.2 因子效应平方和

设 n_{Ai} 为因子 A 的第 $i (i=1, 2, \cdots, m_A)$ 个水平下的试验次数, X_{Aik} 为其第 k 个试验结果, 则因子 A 的效应平方和由下式计算:

$$S_A = \sum_{i=1}^{m_A} n_{Ai} (\bar{X}_{Ai} - \bar{X})^2 \tag{75}$$

即:

$$S_A = \sum_{i=1}^{m_A} T_{Ai}^2 / n_{Ai} - T^2 / n \tag{76}$$

式中 T_{Ai} 为因子 A 第 i 个水平下的所有试验数据的加权总和, $\bar{X}_{Ai}$ 为其加权平均值:

$$T_{Ai} = \sum_{k=1}^{n_{Ai}} X_{Aik} / k_{Aik} \tag{77}$$

$$\bar{X}_{Ai} = T_{Ai} / n_{Ai} \tag{78}$$

其中 k_{Aik} 为因子 A 第 i 个水平下第 k 个试验结果的均值系数, 即 $E(X_{Aik}) = \mu_{Aik} = k_{Aik} \mu$ 。 S_A 的自由度为:

$$f_A = m_A - 1 \tag{79}$$

可以证明, 当式(66)和(67)成立时, 有:

$$\frac{S_A}{\sigma^2} \sim \chi^2(f_A) \tag{80}$$

3.3 两因子交互效应平方和

因子 A 和因子 B 的交互作用效应平方和 $S_{A \times B}$ 由下式计算:

$$S_{A \times B} = S_{AB} - S_A - S_B \tag{81}$$

式中:

$$S_{AB} = \sum_{i=1}^{m_A} \sum_{j=1}^{m_B} n_{ABij} (\bar{X}_{ABij} - \bar{X})^2 \tag{82}$$

即:

$$S_{AB} = \sum_{i=1}^{m_A} \sum_{j=1}^{m_B} T_{ABij}^2 / n_{ABij} - T^2 / n \tag{83}$$

式中:

$$T_{ABij} = \sum_{k=1}^{n_{ABij}} X_{ABijk} / k_{ABijk} \tag{84}$$

$$\bar{X}_{ABij} = T_{ABij} / n_{ABij} \tag{85}$$

其中 T_{ABij} 为因子 A 第 i 个水平和因子 B 第 j 个水平

组合 (A_i, B_j) 下的所有试验数据的加权总和, $\bar{X}_{ABij}$ 为其加权平均值, k_{ABijk} 为 (A_i, B_j) 下第 k 个试验结果的均值系数, 即 $E(X_{ABijk}) = \mu_{ABijk} = k_{ABijk} \mu$; n_{ABij} 为 (A_i, B_j) 下的试验次数, X_{ABijk} 为其第 k 个试验结果; $i=1, 2, \cdots, m_A$; $j=1, 2, \cdots, m_B$; $k=1, 2, \cdots, n_{ABij}$ 。

$S_{A \times B}$ 的自由度等于因子 A 和因子 B 的自由度 f_A 和 f_B 乘积:

$$f_{A \times B} = f_A \times f_B \tag{86}$$

可以证明, 当式(66)和(67)成立时, 有:

$$\frac{S_{A \times B}}{\sigma^2} \sim \chi^2(f_{A \times B}) \tag{87}$$

3.4 采用正交表时的两因子交互效应平方和

若采用正交表^[2]安排多因子有交互作用的试验, 其中因子 A 和因子 B 的交互作用在正交表上占了若干列, 那么该交互作用效应平方和 $S_{A \times B}$ 等于正交表上这若干列的效应平方和之和, 而每列的效应平方和可将每列看成单一因子按式(75)或(76)计算。也可按式(81)计算 $S_{A \times B}$ 。

3.5 误差平方和

误差平方和为:

$$S_E = S_T - S_F - S_I \tag{88}$$

式中 S_F 为所有因子的效应平方和之和, S_I 为所有交互作用的效应平方和之和。 S_E 的自由度为:

$$f_E = f_T - f_F - f_I \tag{89}$$

式中 f_F 为所有因子平方和的自由度之和, f_I 为所有交互作用效应平方和的自由度之和。可以证明:

$$\frac{S_E}{\sigma^2} \sim \chi^2(f_E) \tag{90}$$

3.6 显著性检验

在显著性水平 α 下, 当

$$F_A = \frac{S_A / f_A}{S_E / f_E} \geq F_{\alpha}(f_A, f_E) \tag{91}$$

或

$$F_{A \times B} = \frac{S_{A \times B} / f_{A \times B}}{S_E / f_E} \geq F_{\alpha}(f_{A \times B}, f_E) \tag{92}$$

时, 则拒绝式(65), 即可以断定由仿真、数字化设计等方法求 $k_i (i=1, 2, \cdots, n)$ 时, 未正确考虑因子 A 或交互作用 $A \times B$ 的影响。

3.7 未知参数估计

未知参数 σ^2 的估计量为:

$$\hat{\sigma}^2 = \frac{S_E}{f_E} \tag{93}$$

当不拒绝式(66)时, μ 的估计量为 $\hat{\mu} = \bar{X}$ 。

4 多母体变异系数齐性检验方法

设试验值 $X_{ij} \sim N(\mu_i, \sigma_i^2)$, 其中 $\mu_i = k_i \mu$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n_i$ 。若:

$$\sigma_i = k_i \sigma \quad i = 1, 2, \dots, m \quad (94)$$

则变异系数 σ_i/μ_i ($i = 1, 2, \dots, m$) 相同。因此, 可在多母体方差齐性检验的 Bartlett 方法^[4]基础上, 给出多母体变异系数齐性的检验方法。令:

$$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \frac{1}{k_i^2} (X_{ij} - \bar{X}_i)^2$$

$$i = 1, 2, \dots, m \quad (95)$$

$$S^2 = \frac{1}{n - m} \sum_{i=1}^m (n_i - 1) S_i^2 \quad (96)$$

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \quad (97)$$

式中 $n = \sum_{i=1}^m n_i$ 。

当变异系数 σ_i/μ_i ($i = 1, 2, \dots, m$) 相同, 即 k_i ($i = 1, 2, \dots, m$) 满足式(94)时, Bartlett 统计量 χ^2 近似服从 $\chi^2(m-1)$ 分布, 其中:

$$\chi^2 = \frac{1}{C} \left[(n - m) \ln S^2 - \sum_{i=1}^m (n_i - 1) \ln S_i^2 \right] \quad (98)$$

式中修正系数 C 由下式给出^[4]:

$$C = 1 + \frac{1}{3(m-1)} \left(\sum_{i=1}^m \frac{1}{n_i - 1} - \frac{1}{n - m} \right) \quad (99)$$

在显著性水平 α 下, 式(94)的拒绝域为 $\chi^2 \geq \chi_{\alpha}^2(m-1)$ 。

5 结 论

(1) 多因子异方差分析方法可以对仿真和数字化设计结果中的各个影响因素逐一进行分析和检验, 能够确定哪些因素已被正确仿真, 哪些因素尚未被正确模拟。

(2) 多因子异方差分析方法也能够对仿真和

数字化设计结果中各个因素的交互作用逐一进行小样本检验。

(3) 建立了双因子等重复和无重复试验两种情况的异方差分析方法, 前者可用于两因素有交互作用时仿真和数字化设计结果的因子分析; 后者则用于无交互作用时仿真和数字化设计结果的因子分析。

(4) 对于一般的多因子试验, 给出了异方差分析的总离差平方和、各因子效应平方和、交互作用效应平方和、误差平方和以及它们的自由度计算公式, 并建立了各因子及其交互作用的分析和检验方法。

(5) 建立了多因子正交试验的异方差分析方法, 多因子正交试验具有信息量大、样本小的特点。

(6) 多因子异方差分析方法要求各状态下的变异系数(或相对误差)相同, 因此文中给出一种多母体变异系数齐性检验方法。

参考文献

- [1] 傅惠民 异方差分析方法[J]. 机械强度, 2005, 27(2): 196~201.
FU Huimin Method for Heteroscedastic Analysis [J]. *Journal of Mechanical Strength*, 2005, 27(2): 196~201.
- [2] 《现代应用数学手册》编委会 现代应用数学手册: 概率统计与随机过程卷[M]. 北京: 清华大学出版社, 2004.
Council of Modern Application Mathematics Handbook Modern Applied Mathematics Handbook (Volume: Probability Statistics and Random Process) [M]. Beijing: Tsinghua University Press, 2004.
- [3] 盛骤, 谢式千, 潘承毅 概率论与数理统计[M]. 北京: 高等教育出版社, 2004.
SHENG Zhou, XIE Shi-qian, PAN Cheng-yi Probability Theory and Symbolic Statistics [M]. Beijing: Higher Education Press, 2004.
- [4] 袁执法, 周静芋 试验设计与分析[M]. 北京: 高等教育出版社, 2003.
YUAN Zhi-fa, ZHOU Jing-yu Experimental Design and Analysis [M]. Beijing: Higher Education Press, 2003.