Welch 检验及 Games-Howell 法在单因素方差分析中的应用

山东第一医科大学第二附属医院(271000) 张维福

【提 要】 单因素方差分析(one-way ANOVA),用于完全随机设计的多个样本均数间的比较,其统计推断是推断各样本所代表的各总体均数是否相等。当多个样本观察变量的整体方差不相等时,且方差并非严重不齐,考虑到方差差异之后的更为稳健的分析结果,可以采用校正检验方法(Welch 法),若要进一步推断具体哪两个总体均数有差别,多重比较采用 Games-Howell 法。

【关键词】 单因素方差分析 方差不齐 Welch 检验 Games-Howell 法
【中图分类号】 R195.1 【文献标识码】 A **DOI** 10.11783/j.issn.1002-3674.2025.02.029

单因素方差分析(one-way ANOVA),用于完全随机设计的多个样本均数间的比较,其统计推断是推断各样本所代表的各总体均数是否相等[1]。进行单因素方差分析时,须满足各样本是随机性、独立性,所在总体均正态性及方差齐性等前提条件[1]。在满足随机性、独立性、正态性条件下,当多个样本观察变量的整体方差相等时,即满足方差齐性时,可采用 F 检验进行分析[2]。当多个样本观察变量的整体方差不相等时,即不满足方差齐性时,可通过数据变换达到方差齐性要求,或者采用校正的单因素方差分析法(Welch 检验)进行比较。

Welch 检验是指采用 Welch 分布的统计量检验多个样本均值是否相等。Welch 分布近似于 F 分布,比 F 检验更稳妥,Welch 检验对方差齐性没有要求,所以当数据的分布不满足方差齐性的要求,且方差并非严重不齐时,采用 Welch 检验。单因素方差分析整体比较,推断结论为拒绝 H_0,接受 H_1 时,只能认为各总体均数之间整体比较有差异,但尚不能说明任意两个总体均数之间都有差别。若要进一步推断具体哪两个总体均数有差别,需要进一步多重比较[3]。

多重比较的方法很多[4-5],方差不齐时,采用多组间两两比较的 Games-Howell 法[6]。目前国内外文献对 Welch 检验和 Games-Howell 法鲜有报道。下面就基本原理理论及经典统计软件 SAS、R、SPSS 的实现进行阐述,通过案例数据进行 SPSS 操作演示。

基本原理理论及软件实现

1. Welch 检验基本原理理论[7-13]

Welch 检验,由美国统计学家 Peter Welch 提出的,主要用于比较两个或多个样本的均值是否存在显著差异,特别适用于各组样本的方差可能不相等的情况,因此其在实际应用中具有更广泛的适用性。需要注意的是,Welch 检验基于数据的独立性、正态性等一定的假设,在使用时,应确保这些假设得到满足,否则结果可能会不准确。

Welch 检验方法具体如下:

(1)建立检验假设,确定检验水准

H_0:$\mu_1=\mu_2,\cdots\cdots,=\mu_k$($k\geq3$),即 k 个处理组观测均数相等

H_1:k 组观测均数不全相等

$\alpha=0.05$

(2)计算检验统计量

$$F=\frac{\frac{1}{k-1}\sum_{j-1}^{k}w_j(\bar{x}_j-\bar{x}')^2}{1+\frac{2(k-2)}{k^2-1}\sum_{j}^{k}(\frac{1}{n_j-1})(1-\frac{w_j}{w})^2}$$

其中,$w_j=\frac{n_j}{s_j^2}$,$w=\sum_{j=1}^{k}w_j$,$\bar{x}'=\frac{\sum_{j=1}^{k}w_j\bar{x}_j}{w}$,$k$ 为处理组数。

$$F\sim F(k-1,df)$$

自由度:$df=\frac{k^2-1}{3\sum_{j=1}^{k}(\frac{1}{n_{j-1}})(1-\frac{w_j}{w})^2}$

(3)确定 P 值,作出推断结论

以 $df1$,$df2$ 查(F 界值表)得 $F_{0.05(df1,df2)}$。按 $\alpha=0.05$ 检验水准,判断检验假设。

统计软件 SAS、R、SPSS 的程序实现:

①在 SAS 中实现 Welch 检验

在 SAS 中,进行校正单因素方差分析,代码如下所示:

```
proc anova data=data1 plots=none;
    class group;
    model Hb=group;
means group/welch;
run;
```

②在 R 中实现 Welch 检验

在 R 中,进行校正单因素方差分析,代码如下所示:

```
library(rstatix)
```

welch_anova_test(data=mydata, Hb~group)　#校正的单因素方差分析

③在 SPSS 中实现 Welch 检验

打开 SPSS 软件,并导入数据集。选择“分析”—“比较平均值”—“单因素 ANOVA 检验”。将两个样本变量选入“变量”栏中。单击“选项”按钮,勾选 Welch 选项。点击“确定”按钮,SPSS 将进行 Welch 检验。

2. Games-Howell 法基本原理理论[14-15]

Games-Howell 检验是一种非参数检验多重比较方法,用于比较两个或多个独立样本的均值,特别是当这些样本量不相等,且总体方差可能不同时。而且该检验不需要数据服从正态分布假设,因此相对于一些参数统计方法具有更大的灵活性。但是与某些参数检验相比,Games-Howell 检验可能具有较低的效能。这意味着在样本量较小或效应较小的情况下,它可能无法检测到真实的差异。

Games-Howell 检验方法具体如下:

(1)建立检验假设,确定检验水准

$H_0:\mu_i=\mu_j$,即任何两处理组观测均数相等

$H_1:\mu_i\neq\mu_j$,即任何两处理组观测均数不等

$\alpha=0.05$

(2)计算检验统计量

$$s.e.=\sqrt{\frac{1}{2}\cdot\left(\frac{s_i^2}{n_i}+\frac{s_j^2}{n_j}\right)}$$

n 为每组样本量。统计量 q 的分布称为 studentized range q 分布。如果 $q>q_{crit}$,则两组均值有统计学差异。

此检验相当于

$$\bar{x}_{max}-\bar{x}_{min}>q_{crit}\sqrt{\frac{1}{2}\cdot\left(\frac{s_i^2}{n_i}+\frac{s_j^2}{n_j}\right)}$$

其中 $[latex]\ df'=\dfrac{\left(\frac{s_i^2}{n_i}+\frac{s_j^2}{n_j}\right)^2}{\frac{\left(\frac{s_i^2}{n_i}\right)^2}{n_i-1}+\frac{\left(\frac{s_j^2}{n_j}\right)^2}{n_j-1}}$

基于 $\alpha=0.05$、k(组数)和 df' 的(studentized range q 表),可得 q_{crit},计算 $q_{crit}\times s.e.$。

同理可计算出其他两组的统计量。

(3)确定 P 值,作出推断结论

$\bar{x}_{max}-\bar{x}_{min}>q_{crit}\times s.e.$,可得 $P<0.05$,$\bar{x}_{max}-\bar{x}_{min}<q_{crit}\times s.e.$,可得 $P>0.05$。按 $\alpha=0.05$ 检验水准,判断检验假设。

统计软件 SAS、R、SPSS 的程序实现:

①在 SAS 中实现 Games-Howell 检验

在 SAS 中,事后检验使用 means 语句指定 LSD、BON、Sidak、SNK、Tukey、Duncan 等事后检验的方法。由于 SAS 并未提供 Games-Howell 法,而 Tukey-Kramer 方法与 Games-Howell 法近似,因此,使用 Tukey-Kramer 方法进行事后检验。代码如下:

```
proc anova data=data1 plots=none;
    class group;
    model Hb=group;
    means group/welch tukey;
run;
```

②在 R 中实现 Games-Howell 检验

在 R 中,当不满足方差齐性时采用“Games-Howell”法进行事后检验,代码如下:

games_howell_test (data = mydata, Hb ~ group, conf. level=0.95, detailed=FALSE)

③在 SPSS 中实现 Games-Howell 检验

选择“分析”—“比较平均值”—“单因素 ANOVA 检验”,单击“事后比较”按钮,选择“Games-Howell”法。

案例数据

本案例采用 SPSS 20.0 软件实现操作。实例数据来自某实验室关于“比较接种不同菌型伤寒杆菌 9D、11C 和 DSC_1 小白鼠存活日数”的研究,观察指标为“存活日数(d)”,菌型为 9D、11C 和 DSC_1,比较接种不同菌型伤寒杆菌 9D、11C 和 DSC_1 小白鼠存活日数是否有差异,案例数据见表 1。具体步骤如下 :

表 1　案例数据

编号	9D(d)	编号	11C(d)	编号	$DSC_1(d)$
1	2	11	5	20	3
2	2	12	5	21	5
3	2	13	6	22	6
4	3	14	6	23	6
5	4	15	6	24	5
6	4	16	7	25	7
7	4	17	8	26	7
8	5	18	10	27	9
9	5	19	12	28	11
10	5			29	12
11	5			30	12

1. 适用条件的判断

(1)正态性检验

①软件操作

a.选择“分析”—“描述统计”—“探索”。

b.在“探索”对话框中将观察变量“存活日数”选入右侧“因变量列表”框、将分组变量“菌型”选入右侧“因子列表”框。

c.点击右侧“图”,在子对话框中勾选“含检验的正态图”,取消勾选“茎叶图”,其他不变,点击“继续”,回到“探索”对话框,点击“确定”。

②结果解读

接种不同菌型伤寒杆菌 9D、11C 和 DSC_t 小白鼠

的存活日数(d)分别为 3.70±1.16、7.22±2.39 和7.55±3.05。图 1 的 Q-Q 图上 3 种菌型散点基本围绕对角线分布,提示 3 种菌型存活日数呈正态分布,表 2 列出了 3 种菌型的正态性检验结果。本案查看 Shapiro-Wilk 检验结果[16],分别显示 3 种菌型的 $P=0.123$、0.080、0.287,均>0.05,也提示 3 种菌型存活日数满足正态性条件。综上,本案例满足正态性条件。

(2)方差齐性检验

①软件操作

a.选择“分析”—“比较平均值”—“单因素 ANOVA 检验”。

b.在“单因素 ANOVA 检验”对话框中将观察变量“存活日数”选入右侧“因变量列表”框、将分组变量“菌型”选入右侧“因子”框。

表 2 正态性检验

	菌型	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		统计量	*df*	显著性	统计量	*df*	显著性
存活日数	9D	0.202	10	0.200*	0.878	10	0.124
	11C	0.251	9	0.106	0.853	9	0.080
	DSC_1	0.207	11	0.200*	0.916	11	0.287

*:真实显著水平的下限;a:Lilliefors 显著水平修正

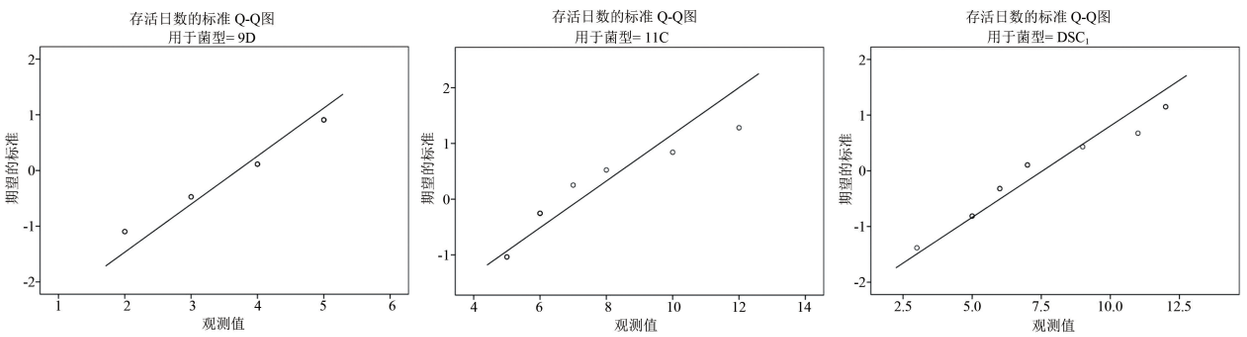

图 1 3 种菌型 Q-Q 图

c.单击“选项”按钮,在弹出的对话框中选中“方差齐性检验”框,其他不变,点击“继续”,回到“单因素 ANOVA 检验”对话框,点击“确定”按钮。

②结果解读

表 3 方差齐性检验结果,可见 $F=4.183$,$P=0.026<0.1$,提示三组数据方差不齐,不满足方差齐性的要求。

本案例数据满足正态性要求,不满足方差齐性要求,但方差并非严重不齐,可采用 Welch 检验进行数据分析。

表 3 方差齐性检验

	Levene 统计量	*df*1	*df*2	显著性
存活日数	4.183	2	27	0.026

2. Welch 检验

(1) 软件操作

①选择“分析”—“比较平均值”—“单因素 ANOVA 检验”。

②单击“选项”按钮,勾选 Welch 选项。

(2)结果解读

表 4 校正单因素方差分析(Welch)法的结果,$F=12.979$,$P=0.001$;提示各组均数不全相等(至少有两组均数不相同)。

表 4 均值相等性的键壮性检验

		统计量[a]	*df*1	*df*2	显著性
存活日数	Welch	12.979	2	15.190	0.001

a:渐近 *F* 分布。

3. 多重比较(Games-Howell 法)

通过上述分析得出了“各组均数不全相等”的结论,但具体是哪些组别之间存在差异尚不清楚,因此需要进行多重比较。

(1) 软件操作

操作步骤同上:

①选择“分析”—“比较平均值”—“单因素 ANOVA 检验”;

②单击“事后比较”按钮,选择“Games-Howell”法。

(2)结果解读

由表 5 可知,菌型 9D 组和 11C 组之间的差异为 3.522d,差异有统计学意义($P=0.005$);菌型 9D 组和 DSC_1 组之间的差异为 3.845d,差异有统计学意义($P=0.005$);菌型 11C 和 DSC_1 之间的差异为 0.323d,差异无统计学意义($P=0.962$)。

表 5 多重比较

(I)菌型	(J)菌型	均值差(I-J)	标准误	显著性	95% 置信区间	
					下限	上限
9D	11C	-3.522*	0.876	0.005	-5.88	-1.17
	DSC_1	-3.845*	0.989	0.005	-6.45	-1.24
11C	9D	3.522*	0.876	0.005	1.17	5.88
	DSC_1	-0.323	1.215	0.962	-3.42	2.78
DSC_1	9D	3.845*	0.989	0.005	1.24	6.45
	11C	0.323	1.215	0.962	-2.78	3.42

*:$P<0.05$。

4. Kruskal-Wallis *H* 检验

（1）软件操作

①选择“分析”—“非参数检验”—“独立样本检验”。

②单击“字段”按钮，将字段“存活日数”选入右侧“检验字段”框、将分组变量“菌型”选入右侧“分组”框。

③单击“设置”按钮，选择“自定义检验”，对话框中选中“Kruskal-Wallis 单因素 ANOVA”框，其他不变，点击“运行”按钮。

（2）结果解读

表 6 描述 Kruskal-Wallis *H* 检验的结果，*F* = 14.960，*P* = 0.001；提示各组均数不全相等（至少有两组均数不相同）。两两比较显示 9D 组和 11C 组之间的差异为 3.270d，差异有统计学意义（*P* = 0.003）；9D 组和 DSC_1 组之间的差异为 3.420d，差异有统计学意义（*P* = 0.002）；11C 和 DSC_1 之间的差异为 0.018d，差异无统计学意义（*P* = 1.000）。

表 6　Kruskal-Wallis *H* 检验多重比较

样本 1-样本 2	检验统计量	标准误	标准检验统计量	显著性	调整显著性
9D-DSC_1	-13.018	3.806	-3.420	0.001	0.002
9D-11C	-13.089	4.003	-3.270	0.001	0.003
DSC_1-11C	0.071	3.916	0.018	0.986	1.000

注：每行检验原假设：样本 1 和样本 2 分布相同。显示渐进显著性（2-sided 检验）。显著性水平是 0.05。

Welch 检验及 Games-Howell 法与 Kruskal-Wallis *H* 检验结论一致。

讨　论

在单因素方差分析时，首先进行适用条件的判断，如果数据不是严重偏态或者只有部分组别数据不满足正态性要求，鉴于其对数据分布具有一定的耐受力，且参数检验的统计学效能优于非参数检验，还是可以使用单因素方差分析方法。如果数据满足整体满足正态分布，也可使用该种分析方法。但严格来讲，应尽可能分组别检验数据的正态性。当组间方差差异较大时，可以进行变量转换或采用非参数检验。如果数据正态性和方差齐性都不满足，最好使用 Kruskal-Wallis *H* 检验。对于不是特别严重的方差不齐，考虑到方差差异之后的更为稳健的分析结果，可以采用校正检验方法（Welch 法），当单因素方差分析有显著意义时，多重比较采用 Games-Howell 法。

参　考　文　献

[1] 刘浩，史雨梅，余晓美．基于 SPSS 单因素方差分析在专业认同研究中的应用[J]．经济研究导刊，2020(5)：71-73.

[2] 邹祎．SPSS 软件单因素方差分析的应用[J]．价值工程，2016，35(34)：219-222.

[3] 胡良平，高辉．如何正确运用单因素多水平设计定量资料方差分析[J]．中西医结合学报，2008(3)：316-319.

[4] 胡纯严，胡良平．如何正确运用方差分析——多个均值之间的多重比较[J]．四川精神卫生，2022，35(1)：21-25.

[5] SAS Institute Inc. SAS/STAT 15.1 user's guide[M]. CaryNC：SAS Institute Inc，2018：3957-4142.

[6] 王鹏伟．血清 NFL、SBDP145 对早产儿脑损伤早期的预测价值[D]．大连：大连医科大学，2021.

[7] 肖枝洪，朱强．统计模拟及其 R 实现[M]．武汉：武汉大学出版社，2010：82-83.

[8] 董玉恒．当方差不齐时两组数据比较的近似 F′检验法与 Welch—Aspin 检验法的等价性证明[J]．数理医药学杂志，1994(2)：111-112.

[9] 曹胜炎，张金发．Welch 检验中有关自由度的几点讨论[J]．华中农业大学学报，1989(4)：369-372.

[10] 叶鸿烈．方差分析中 F 统计量计算教学研究[J]．大众科技，2014，16(5)：165-167.

[11] 李丛．方差分析的理论与应用[J]．科技视界，2014(12)：146-147.

[12] 高卓．单因素方差分析应用的实证研究[J]．赤峰学院学报（自然科学版），2014，30(6)：4-6.

[13] 田兵．单因素方差分析的数学模型及其应用[J]．阴山学刊（自然科学），2013，27(2)：24-27.

[14] 蓝石．交互性多维变量的量化研究方法[M]．上海：格致出版社，2018：14-20.

[15] 李俊，丁建华，金显文，等．方差分析多重比较中 *q* 值表和 SSR 表的构建[J]．廊坊师范学院学报（自然科学版），2019，19(2)：64-67.

[16] 如何正确运用单因素方差分析——药物研究中的统计学（二）[J].药学与临床研究，2018，26(2)：159-160.

（责任编辑：郭海强）