Leibniz Universität Hannover

Multiple Contrast Tests in the Presence of Heteroscedasticity

M. Hasler, L.A. Hothorn

Institute of Biostatistics, Leibniz University of Hannover, Germany

Email: *hasler@biostat.uni-hannover.de*

1. Situation

For $j=1,\dots,n_i$ and $i=1,\dots,k$, let X_{ij} denote the j^{th} observation under the i^{th} treatment. Suppose the X_{ij} to be independently normal with means μ_i and possibly unequal variances σ_i^2. We are interested in the vector of ratio contrasts $\gamma = (\gamma_1,\dots,\gamma_q)^T$ where for $1 \le l \le q$

$$\gamma_l = \frac{\sum_{i=1}^{k} c_{li}\mu_i}{\sum_{i=1}^{k} d_{li}\mu_i}.$$

The vectors $c_l = (c_{l1},\dots,c_{lk})^T$ and $d_l = (d_{l1},\dots,d_{lk})^T$ consist of real constants. The hypothesis to test is

$$H_0:\ \gamma_l = \theta_l \quad \forall l = 1,\dots,q$$

against several alternatives due to given testing problems and for given relative thresholds $\theta_1,\dots,\theta_q$.

Problem: Existing approaches are not robust!

2. Procedure

Test statistic for $1 \le l \le q$:

$$T_l = \frac{\sum_{i=1}^{k}(c_{li} - \theta_l d_{li})\overline{X}_i}{\sqrt{\sum_{i=1}^{k}(c_{li} - \theta_l d_{li})^2 \hat{\sigma}_i^2 / n_i}}$$

→ Separate sample variances

Degrees of freedom for $1 \le l \le q$:

$$df_l = \frac{\left(\sum_{i=1}^{k}(c_{li} - \theta_l d_{li})^2 \hat{\sigma}_i^2 / n_i\right)^2}{\sum_{i=1}^{k} \frac{\left((c_{li} - \theta_l d_{li})^2 \hat{\sigma}_i^2 / n_i\right)^2}{n_i - 1}}$$

→ Due to Satterthwaite (1946)

Correlation matrix $R = (\rho_{lm})$ with $(1 \le l \ne m \le q)$

$$\rho_{lm} = \frac{\sum_{i=1}^{k}(c_{li} - \theta_l d_{li})(c_{mi} - \theta_m d_{mi})\hat{\sigma}_i^2 / n_i}{\sqrt{\sum_{i=1}^{k}(c_{li} - \theta_l d_{li})^2 \hat{\sigma}_i^2 / n_i}\sqrt{\sum_{i=1}^{k}(c_{mi} - \theta_m d_{mi})^2 \hat{\sigma}_i^2 / n_i}}$$

→ Plug-in of the sample variances

So, each test statistic T_l $(1 \le l \le q)$ is compared with a **separate** – „its own“ – **quantile** coming from a q-variate t-distribution with adjusted degrees of freedom and a correlation matrix for which a variance plug-in is used. This procedure is referred to as **PI**.

3. Competing approaches

HOM: Multiple contrast test (MCT) for homogeneous variances and originally for difference contrasts; pooled sample variance; correlations without variance estimator; common degree of freedom
→ same quantile for all contrasts

GH: Originally for all-pair comparisons as a difference contrast (Games and Howell, 1976); separate sample variances; correlations without variance estimator; separate degrees of freedom due to Satterthwaite (1946)
→ separate quantiles for the contrasts

HTL: Originally for unbalanced settings (Hochberg and Tamhane, 1987) or for special contrasts only (Tamhane and Logan, 2004); separate sample variances; average of correlations and degrees of freedom as in **PI**
→ same quantile for all contrasts

4. Simulation studies

- Different settings and allocations, numbers of treatments, contrasts
- Global and local α level focused (weak and strong control of FWER)
- 100000 simulation runs in statistic software R (mvtnorm)

Example: Dunnett contrast, one-sided, $\theta_i = 1.25$ $(1 \le i \le q)$, $\alpha = 0.05$

Setting	HOM	GH	PI	HTL
n=(10,10,10), σ=(10,10,50)	0.0471	0.0556	0.0496	0.0506
n=(4,13,13), σ=(10,10,50)	0.0055	0.0654	0.0527	0.0585
n=(13,13,4), σ=(10,10,50)	0.1970	0.0545	0.0514	0.0770
n=(10,10,10), σ=(30,30,30)	0.0492	0.0487	0.0485	0.0487
n=(10,10,10,10,10), σ=(10,10,10,10,50)	0.0634	0.0562	0.0498	0.0496
n=(6,11,11,11,11), σ=(10,10,10,10,50)	0.0338	0.0619	0.0500	0.0485
n=(11,11,11,11,6), σ=(10,10,10,10,50)	0.1462	0.0573	0.0520	0.0593
n=(10,10,10,10,10), σ=(30,30,30,30,30)	0.0497	0.0500	0.0493	0.0493

Results:

- **HOM** does not control the α level
- **PI** almost exact, tightest α ranges
- **GH** and **HTL** often too both liberal and conservative, respectively
- **GH** differs more dependent on the special contrasts
- **HTL** differs more dependent on settings, widest α ranges

5. Simultaneous confidence intervals

Lower bounds of approximate (1−α)100% SCI of **PI** for $\gamma = (\gamma_1,\dots,\gamma_q)^T$:

$$\hat{\theta}_l^{(l)} = \frac{-B_l - \sqrt{B_l^2 - 4A_lC_l}}{2A_l}$$

$$A_l = \left(\sum_{i=1}^{k} d_{li}\overline{X}_i\right)^2 - t_{q,1-\alpha}^2(df_l, R)\sum_{i=1}^{k} d_{li}^2\hat{\sigma}_i^2 / n_i, \quad C_l = \left(\sum_{i=1}^{k} c_{li}\overline{X}_i\right)^2 - t_{q,1-\alpha}^2(df_l, R)\sum_{i=1}^{k} c_{li}^2\hat{\sigma}_i^2 / n_i,$$

$$B_l = -2\left(\left(\sum_{i=1}^{k} c_{li}\overline{X}_i\right)\left(\sum_{i=1}^{k} d_{li}\overline{X}_i\right) - t_{q,1-\alpha}^2(df_l, R)\sum_{i=1}^{k} c_{li}d_{li}\hat{\sigma}_i^2 / n_i\right), \quad A_l > 0, \quad 1 \le l \le q$$

6. Conclusions

- Adjusted degrees of freedom not sufficient to handle heteroscedasticity, plug-in **variance estimators necessary**
- New approach (**PI**) keeps FWER best for all contrasts and settings
- Taking **averages** of correlations and degrees of freedom **too rough**
- Also other approaches studied which take the minimum (conservative) or maximum (liberal) of the Welch-adjusted degrees of freedom, respectively
- Same theory also considered for MCT of DIFFERENCES in means with similar results
- R code available from first author

References:

Dilba, G., Bretz, F., Guiard, V., Aug. 2006. Simultaneous confidence sets and confidence intervals for multiple ratios. Journal of Statistical Planning and Inference 136 (8), 2640–2658.

Dilba, G., Bretz, F., Guiard, V., Hothorn, L. A., 2004. Simultaneous confidence intervals for ratios with applications to the comparison of several treatments with a control. Methods of Information in Medicine 43 (5), 465–469.

Games, P. A., Howell, J. F., 1976. Pairwise multiple comparison procedures with unequal n's and/or variances: a Monte Carlo study. Journal of Educational Statistics 1 (2), 113–125.

Hochberg, Y., Tamhane, A. C., 1987. Multiple comparisons procedures. John Wiley and Sons, Inc.

R Development Core Team, 2006. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL http://www.R-project.org

Satterthwaite, F. E., 1946. An approximate distribution of estimates of variance components. Biometrics 2, 110–114.

Tamhane, A. C., Logan, B. R., 2004. Finding the maximum safe dose level for heteroscedastic data. Journal of Biopharmaceutical Statistics 14 (4), 843–856.

Welch, B. L., 1938. The significance of the difference between two means when the population variances are unequal. Biometrika 29, 350–362.