A Studentized Range Test for the Equivalency of Normal Means under Heteroscedasticity
Miin-Jye Wen* and Hubert J Chen

Department of Statistics, National Cheng Kung University
Email:mjwen@stat.ncku.edu.tw


1. Preface
In analysis of variance (anova), we generally assume that the error terms are independent and normally distributed with a common unknown variance. However, in practice, one often encounters situations where the error variances are not only unknown but also unequal (it is so-called heteroscedasticity). Therefore, there is a need to search for a suitable statistical method to solve this kind of problems in anova. Furthermore, in classical hypothesis testing the interest is often to test the null hypothesis that the population means are all equal. It is well known that, for a large enough sample size, the classical test will almost always reject the null hypothesis. In many real world problems (for example, to test generic drug vs innovative drug in pharmaceutical study), the practical interest is frequently to examine whether the normal population means fall into an indifference zone, not just the equality of means. This idea leads to the consideration of equivalence hypothesis under the situation of heteroscedasticity of error terms stated as $H_0 : \frac{1}{k} \sum_{i=1}^{k} |\mu_i - \bar{\mu}| \leq \delta$ versus the alternative of inequivalence $H_a : \frac{1}{k} \sum_{i=1}^{k} |\mu_i - \bar{\mu}| \geq \delta^* > \delta$, where $\bar{\mu}$ is the grand average of the means $\mu_1, \ldots, \mu_k, \delta \geq 0$ is a predetermined indifference constant and $\delta^* \geq 0$ is a detective amount specified in advance. Thus, the expression on the left side of the inequality in the null hypothesis is often regarded as the average deviation of the means from their grand average, and the null hypothesis claims that the means are falling into a negligible zone.

2. Introduction
When there are only two treatments ($k = 2$), the two one-sided tests by Schuirmann (1987) for bioequivalence of the treatments become a guideline in the field of pharmaceutical industry for drug development and medical studies. In situations where there are three or more treatments ($k > 2$) under study, the equivalence/bioequivalence of these treatments has barely been touched by Giani and Finner (1991), Chen, Xiong and Lam (1993), and Chen and Chen (1999). In this paper, when the population variances are unknown and possibly unequal. We propose an equivalence hypothesis $H_0 : \frac{1}{k} \sum_{i=1}^{k} |\mu_i - \bar{\mu}| \leq \delta$ against an inequivalence alternative $H_a : \frac{1}{k} \sum_{i=1}^{k} |\mu_i - \bar{\mu}| \geq \delta^* > \delta$, where $\delta (>0)$ is an equivalent constant specified by experts in advance. It has been shown that this measure of equivalence is equivalent to the interval hypothesis of equivalence for $k = 2$. When the variances are unknown and unequal, a studentized range test using a two-stage and a one-stage sampling procedure, respectively, is proposed for testing the null hypothesis that the average deviation
of the normal means is falling into a practical indifference zone against the alternative of inequivalence. Since the level and the power of the test are functions of all unknown means and unknown variances, it is necessary to find a least favorable configuration (LFC) of the means to guarantee a maximum level (the probability of rejection region) at a given null hypothesis and a LFC to guarantee a minimum power at a given alternative. By our findings the level under a given null $H_0$ and the power under a given alternative are fully independent of not only the unknown means but also the unknown and unequal variances. Therefore, the critical values and required sample sizes can be simultaneously determined by solving system of integral equations. Statistical tables to implement the procedure are provided for practitioners and an example is given to demonstrate the use of the test.

3. Summary and Conclusion
Testing the null hypothesis of equal treatment means is sometimes impractical in real applications. An alternative measure to detect the difference among means is the average deviation of the means, which extends the idea of equivalence among means. The test of equivalence receives more attention in health sciences, pharmaceutical industry, and other applied areas. When the variances are unknown and unequal, a studentized range test using a two-stage and a one-stage sampling procedure, respectively, is proposed for testing the null hypothesis. The two-stage sampling procedure is a design-oriented procedure that satisfies certain probability requirements and simultaneously determines the required sample sizes (which can be largely increased at the second stage) in an experiment while the one-stage sampling procedure is a data-analysis procedure after the data have been collected, which can supplement the two-stage sampling procedure when the latter has to end its experiment sooner than its required experimental process is completed. At that time the level and power can be approximated, and the one-stage sampling procedure is shown to be quite feasible under heterocedasticity.

Acknowledgement: This research with the second author being principal investigator was supported by National Science Council Grants NSC92-2119-M-006-007, 2003-2004 and NSC93-2118-M-006-008, 2004-2005, Taiwan, ROC.

4. References