RECENT PHD THESES

Katholieke Universiteit Leuven (KULeuven)

Roselinde Kessels. *Optimal Designs for the Measurement of Consumer Preferences*. (23/10/2006) – Promotors: Prof. Martina Vandebroek and Prof. Peter Goos (UA).

This thesis focuses on the design of conjoint experiments for measuring the trade-offs people make in choosing between alternative products and service providers. Marketing consultants and researchers frequently use these experiments to predict people's choices for prospective goods. In this way, they assist companies in launching innovative products or services. The entire process from collecting consumer preference data to analyzing them and simulating the marketplace is generally known as conjoint analysis.

Conjoint analysis assumes that a product or service can be decomposed into its component attributes and levels. A good is thereby described by levels for each of the attributes. For example, a car is characterized by the attributes price, transmission, airbags, door lock and audio system. A possible profile or alternative of a car is then a car with a price of 18,000€, manual transmission, front and dual side airbags, an auto lock function and a radio and CD player. By presenting a series of profiles to a number of test persons and finding out which are most preferred, conjoint analysis allows the determination of the relative importance of each attribute and level in the purchasing decision. The relative values or utilities respondents derive from the attribute levels are also called part-worths. Conjoint analysis is based on the fact that the part-worths can better be measured when the attributes are *con*sidered *joint*ly rather than in isolation.

Respondents usually evaluate profiles in one of the following two ways. They either choose their preferred profile from a set of profiles, also called a choice set, and they repeat this task for several other choice sets presented to them. Such a conjoint experiment is a choice-based conjoint experiment, also referred to as a conjoint choice or discrete choice experiment, or more succinctly, a choice experiment. Or, the respondents rate a number of profiles on a scale, for example a 10-point scale. This type of conjoint experiment is a rating-based conjoint experiment. In the thesis we deal with the question of how to properly design choice-based and rating-based conjoint experiments. This means that we search for profiles that, when administered to respondents, yield maximum information on the part-worths. To find the best possible design in each case, we make use of design criteria or optimality criteria resulting in optimal designs. The thesis is split into two parts each involving the design of one type of conjoint experiment.

Katholieke Universiteit Leuven (KULeuven)

Rembert De Blander. *Essays on Endogeneity and Parameter Heterogeneity in Cross-Section and Panel Data*. (28/04/2006) - Promotor: Prof. Marinus Verbeek (Erasmus University Rotterdam), Co-promotor: Prof. Geert Dhaene (K.U.Leuven)

In this thesis, both the Correlated Random Coefficient (*CRC*) model as well as a dynamic panel data model are considered.

The CRC model is a random coefficient model the crucial characteristic of which is the possible correlation between regressors and coefficients. Assuming that the constant term is also associated with a random and correlated coefficient, the CRC model encompasses standard models that allow for endogenous regressors. A first paper presents an estimator for the linear CRC model which is an extension of Garen's (1984) Selectivity Bias Method, but the outcome equation is augmented with more terms compared to Garen. In a second Paper, I discuss semiparametric estimation of a linear correlated random coefficient model (Heckman and Vytlacil (1998)). I include two unknown functions of the residuals in the equation of interest, one of which is multiplied by the treatment variable. I propose to estimate these unknown functions by series regression, resulting in root N-consistent estimation of the parameters of interest. By making use of the unified treatment of models for truncation, sample selection and limited dependent variables (Heckman (1976)), and of generalized residuals (Gouriéroux et al. (1987)), this approach is valid as well for discrete treaments. A third paper, finally, considers estimation methods for the CRC model when panel data are available, which allows for estimation of the average treatment effect by IV as described by Wooldridge (2003), using Hausman and Taylor (1981) style instruments. The part concerning dynamic panel data focusses on the derivation of unit root tests for fixed time dimension, which extends Harris and Tzavalis' (1998) tests to panel data models with AR(1) errors. The limiting distributions of the test statistics (for increasing N and fixed T) are shown to be normal. Closed-form expressions for the first and second moments of the test statistics are derived. Heterogenous initial conditions and drift in the data generating process are taken into account by including fixed effects and individual-specific linear time trends in the regression. This inclusion makes the least squares estimators of the autoregressive parameters inconsistent for fixed T, and appropriate bias-corrections are proposed.

Electronic access:
http://www.kuleuven.ac.be/doctoraatsverdediging/cm/3H05/3H050156.htm

Universiteit Gent (UGent)

Beatrijs Moerkerke. *To be or not to be significant: on more powerful methods for multiple testing* (17/11/2006) - Promotor: Prof. dr. E. Goetghebeur.

Recent advances in technology are leading to the production of enormous amounts of data. To turn these data into information, statisticians are equipped with statistical models and techniques for estimation and hypothesis testing. They traditionally draw inference with well understood error margins. The dimension of today's datasets however has led to a multiplicity of analyses and tests that are leading to the present new and fundamental challenges to hypothesis testing. If no adjustments are made to account for this multiplicity, the number of false positive results can grow unacceptably large. This is at the heart of the multiple testing problem, which forms the topic of this thesis.

The multiple testing problem has received particular attention in the field of statistical genetics where a huge number of candidate genes may be tested for an association with phenotype, based on relatively small samples. This has led to a whole range of new (ad hoc) statistical developments. These procedures mainly aim to avoid a flood of false positives, i.e. to protect against false findings that put researchers on the wrong track. They thus focus on the null of no association. Unfortunately, by reducing the number of selected genes, one also loses some ability to detect interesting findings. The strategy that is developed in this thesis also quantifies evidence against an important alternative to gain power.

In chapter 1, we introduce the multiple testing problem and a range of error measures through which repeated tests are being evaluated. We present popular procedures designed to control these different error measures.

In chapter 2, we borrow from lessons learned in the field of statistical genetics to confront multiplicity questions in the study of high dimensional Quality of Life (QOL) outcomes. Specifically, we investigate the impact of treatment on symptoms and functionality in several aspects of daily life of breast cancer patients. Unlike classical analysis of QOL-questionnaires, we extend existing techniques to compare treatments in terms of all individual items as they reflect different symptoms that matter to the patients. More in particular, we develop permutation based procedures to detect items of the QOL-questionnaire that differ significantly over observed treatments even in the face of multiple testing.

While in chapter 1 and 2, existing methods are studied and extended, a new powerful methodology to deal with the multiple testing problem is developed from chapter 3 onwards.

To avoid a flood of false positives and to avoid losing too much power, we strive for a better balance between the magnitude of a genetic effect and its precision in chapter 3 and 4. We complement the traditional p-value (p_0) with an alternative p-value (p_1), a measure of impotence, which summarizes evidence against a target alternative. We build a formal decision criterion by balancing gene-specific type I and type II errors to optimize an expected gain. This criterion turns out to be cast in terms of p_0 and p_1 and leads to an intuitive measure of relative evidence based on which genes are ranked and selected. We call this procedure the balanced test. The technique is applied and developed in the context of statistical selection of genetic markers for plant breeding and for detecting differentially expressed genes in hereditary breast cancer.

In chapter 5, we take this testing procedure one step further and develop two-stage designs for screening genetic markers when the cost of measurements is high. Based on a first sample, a genetic marker is selected if it shows enough evidence against the null and in favor of a specified worthwhile alternative, while a sufficiently convincing lack of effect results in acceptance of the null. Otherwise the genetic marker is situated in a grey zone and more data are gathered at the second stage after which a binary decision is made on all available data. We develop designs and decision rules which optimize an expected gain.

We return to the original multiple testing framework in chapter 6 and study experimentwise properties of our procedure. The balanced test achieves a higher power to detect specified alternatives than methods based on classical p-values. This is a consequence of considering evidence against both the null and alternative hypothesis and handling gene-specific decision criteria. It appears to also provide a more stable solution than standard techniques.

In this thesis, the multiple testing problem is approached in a more specific way to provide a solution for the lack of power of classical tests. The development of new techniques that rethink the ultimate goals of hypothesis testing in a context where many tests are performed is important. This process involves a reflection on existing methods and the recognition of some of their shortcomings when keeping the original (biological) question in mind. Some important steps towards this goal are taken here and smoothen the path for further work on this topic such as the detection of gene-gene interactions.