

Invited Speaker and Poster Session Abstracts

Invited Speaker Abstracts

Author: **Burdick, Richard**

Title: Using Variance Component Models to Compare Variability of Manufacturing Processes

Abstract: A common experiment employed in industrial settings is the repeatability and reproducibility (R&R) study. One objective of such an experiment is to determine if a measurement procedure or instrument is adequate for monitoring a process. Such a study may also be described as a gauge capability analysis. If the measurement error is small relative to total process variation, then the measurement procedure is deemed "adequate". R&R studies typically employ a two-factor crossed design with the factors "parts" and "operators". Measures of interest are functions of the variance components of this model. Several methods have been proposed for constructing confidence intervals for these measures of variability. These studies are summarized for cases where operators are either fixed or random. Results from a study to compare the variability of two measurement processes are also presented.

Authors: **Datta, Gauri Sankar, Malay Ghosh, and Yeong-Hwa Kim**

Title: Probability Matching Priors For One-Way Unbalanced Random Effects Models

Abstract: This paper considers development of noninformative priors for unbalanced one-way random effects models. The main focus is on the development of probability matching priors (that is those for which the coverage probabilities of Bayesian credible intervals match asymptotically their frequentist counterparts) when the parameter of interest is the variance ratio. Under a suitable orthogonal reparametrization (Cox and Reid, 1987, JRSS B), it is possible to characterize the class of first order probability matching priors. However, it is shown that there does not exist any second order probability matching prior in this case. Also, it is shown that the one-at-a-time reference prior of Berger and Bernardo works well, and is fairly robust with respect to the choice of prior parameters.

Authors: **Eubank, Lane, Youngjo Lee, Kathleen Purdy and Justus Seely**

Title: Confidence Intervals for Variance Components

Abstract: The random-effect mean squares from an unweighted means ANOVA (UANOVA) have been exploited by a number of writers to form confidence intervals for variance components. Except for a few special cases, these results are limited to models where a UANOVA can be defined. For a two variance component model, we show how this limitation can be eliminated by defining two mean squares that have properties similar to the mean squares from a random one-way UANOVA. These two mean squares can then be used for variance component inference in the same way that mean squares from a UANOVA are utilized. Also, several examples are investigated that illustrate some of the limitations using the UANOVA mean squares. An alternative to the UANOVA mean squares, which seems to have better performance characteristics, is suggested.

Author: **Ghosh, Malay, Ming-Hui Chen, Atalanta Ghosh, and Alan Agresti**

Title: Non-Conjugate Bayesian Analysis Of One Parameter Item Response Models

Abstract: We present a unified Bayesian approach for the analysis of one-parameter item response models. A necessary and sufficient condition is given for the propriety of posteriors under improper priors with nonidentifiable likelihoods. Posterior distributions for item and subject parameters may be improper when the sum of the binary responses for an item or subject takes its minimum or maximum possible value. When the item parameters have a flat prior but the item totals do not fall at a boundary value, we prove the propriety of the Bayesian joint posterior under some sufficient conditions on the joint (proper) distribution of the subject parameters. The methods are implemented using Markov chain Monte Carlo and illustrated with an example from a cross-over study comparing three medical treatments.

The Bayesian methods are compared against the usual generalized mixed linear models. We also provide some results related to the analysis of binary matched pairs data.

Authors: Gianola, Daniel, and R. Rekaya

Title: Bayesian Random Regression Models for Longitudinal Binary Data with Applications to Dairy Cattle Breeding

Abstract: Our concern is genetic analysis of a disease, mastitis, observed longitudinally in dairy cows. Data were 3341 test-day records from 329 first-lactation Holstein cows scored for absence/presence of mastitis at 14, 30 days (d) of lactation, and every 30d thereafter. Susceptibility was related to a Gaussian process. Given the parameters, the probability of the data for each cow was the product of probabilities at each test-day. The conditional probability of infection at time t for each cow was a normal integral, its argument being a function of fixed and random effects, and of time. Models for the latent variable were: 1) year-month of test + a 5-parameter linear regression + breeding value (BV) of cow + environmental effect peculiar to all records of a cow + residual. 2) As before, excluding BV, but with a 5-parameter random regression for each cow. 3) Hierarchical model of a 3-parameter regression for each cow. Computations were by Gibbs sampling, with data augmentation. Model 1 posterior mean of heritability was .05. Model (2) heritabilities were, e.g., .27 (14d), .03 (120d), and .07 (305d). Model 3 heritabilities were .57 (14d), .06 (120d) and .19 (305d). Bayes factors were: .011 (Model 1/Model 2), .017 (1/3) and 1.535 (2/3). Model can be extended: fitting structured covariance matrices for residuals, thick-tailed distributions and oscillatory components.

Authors: Khuri, Andre, and Juneyoung Lee

Title: Modeling the Power Function For a Variance Component Under Heterogeneous Error Variances

Abstract: Traditional analysis of variance (ANOVA) tests are based on the assumption that the error variances are homogeneous. Departure from such an assumption combined with a certain degree of imbalance, in the data set under consideration, can have a considerable effect on the performance of a given ANOVA test. In this paper, we consider the effect of heterogeneity of the error variances on the power of the usual F -test concerning the among-group variance component for an unbalanced random one-way model. The power is modeled empirically using generalized linear models techniques. The purpose of the proposed modeling approach is to provide added insight into the combined effects of heterogeneity of variances and data imbalance on the test. This includes, in particular, the ability to monitor changes in the power values, and to detect influential error variances, which can be helpful in increasing robustness of the F -test.

Author: Lahiri, Partha

Title: Jackknife Estimation of Variance Components

Abstract: A general theory of jackknifing M-estimators is considered for a general mixed model which includes mixed linear normal model and mixed logistic model. Our M-estimators cover the modified profile MLE (e.g., REML estimators), penalized MLE, or M-estimators not associated with a maximization process (e.g., the method of moment estimators). We also present the related mean square error estimation of empirical best predictor of a general mixed effect. Simulation results show that our method is quite robust against departure from normality. Our method is illustrated using a real life data.

Author: **Lohr, Sharon**

Title: Optimal Design for Variance Components and Design Implementation

Abstract: Since variance components are nonlinear parameters in a linear model, classical optimal designs for estimating variance components in the m-stage nested model depend on the values of the parameters. We derive optimal local and Bayesian designs for estimating variance components in the nested model. These designs are then compared to designs in the quality control literature. Tree structures are used to allow computation and visualization of the resulting designs.

Author: **Mathew, Thomas**

Title: The One-Way Random Model: Some New Problems and Applications

Abstract: In the talk, I will present some new applications of the one-way random effects model. The applications include the assessment of occupational exposure to contaminants, and testing bioequivalence. These applications give rise to some new hypotheses testing problems, and confidence and tolerance interval problems, involving the one-way random model. I will give a review of these problems and the available solutions.

Authors: **Rao, Jon N.K., and K. Das**

Title: Estimation of MSE of EBLUP Estimator under General Mixed ANOVA Models

Abstract: Estimation of a linear combination of fixed effects and realized values of random effects in general mixed ANOVA models is of considerable interest in practice; for example, in animal breeding experiments and small area estimation. Empirical BLUP estimators are obtained by replacing the variance components by consistent estimators; in particular ML, REML or invariant quadratic estimators. We obtain a second order approximation to MSE of EBLUP and study its accuracy. An estimator of MSE, correct to second order, is also obtained. We show that the MSE estimator when ML is used involves an extra term that is absent in the case of REML and invariant quadratic estimators of variance components. The special case of a balanced two-way crossed random effects model with interaction is used to illustrate the results.

Author: **Rao, Poduri**

Title: Nonnegative Estimators for Variance Components Utilizing Prior Information

Abstract: For the variance components of the one-way Analysis of variance model, nonnegative estimators that can utilize prior information are developed from the Analysis of Variance, Unweighted Sums of Squares, Minimum Variance Quadratic Unbiased Estimation and related procedures. Biases and Mean Square Errors of these estimators are evaluated and their sensitivity to the prior values are examined. Analogous procedures for the One-way Analysis of Covariance Model are also examined.

Authors: **Qaqish, Bahjat, and Pranab Kumar Sen**

Title: Variance Components for Dependent Qualitative Responses

Abstract: The concept of variance components for qualitative and semi-quantitative traits will be explored. Measures of heterogeneity for categorical responses will be discussed with reference to genetic analyses. Issues of decomposability will be discussed and connections with generalized mixed models explored with a view toward interpretation and applicability.

Author: **Sinha, Bimal K.**

Title: Nonnegative estimation of variance components in multivariate unbalanced mixed models with two variance components

Abstract: In this talk we address the problem of nonnegative estimation of variance components in multivariate unbalanced mixed models with two variance components. This generalizes the work in Mathew/Niyogi/Sinha (JMA, 1994). We will illustrate our results with an example.

Author: **Speed, Terry**

Title: Variance component models for mapping human quantitative trait genes: a brief survey

Abstract: With the ready availability of high-resolution genetic maps and efficient high-throughput genotyping, considerable attention is now being given to mapping genes contributing to quantitative traits in humans. These include traits related to heart disease, alcoholism, and many biochemical and physiological traits. Because the genetic basis of such traits is usually completely unknown, so-called allele-sharing methods are popular in these contexts, as they make very few specific assumptions concerning the genotype-phenotype relationship. One of the simplest such involves sib-pairs, and I will focus on them in my talk. My aim is to give a brief partial survey of the use of variance components for mapping quantitative traits using allele-sharing in sib-pairs.

Author: **Sutradahr, Brajendra C.**

Title: On recent estimation approaches for variance components in generalized linear mixed models

Abstract: In view of the cumbersome and often intractable numerical integrations required for a full likelihood analysis, several suggestions have been made recently for approximate inference in generalized linear mixed models. For example, we refer to the penalized quasi-likelihood (PQL) approach of Breslow and Clayton (1993, JASA), bias corrected PQL approach of Breslow and Lin (1995, Biometrika), hierarchical likelihood approach of Lee and Nelder (1996, JRSSB). But as shown by Sutradhar and Qu (1998, CJS) (see also Jiang (1998, JASA), these approaches do not yield consistent estimators for the variance components of the random effects of the mixed model. In this talk I will discuss a small variance based likelihood approximation technique which yields consistent estimators of the variance components both for binary and Poisson mixed models, which are important special cases of the generalized linear mixed model. I will also outline a simulation based general likelihood procedure as an extension of Jiang's (1998) simulation based method of moments for the estimation of the variance components.

Author: **Wolfinger, Russell**

Title: Non-Conjugate Bayesian Analysis of Variance Component Models

Abstract: We consider the usual normal linear mixed model for variance components from a Bayesian viewpoint. With conjugate priors and balanced data, Gibbs sampling is easy to implement. However, simulating from full conditionals can become difficult for the analysis of unbalanced data with possibly non-conjugate priors, thus leading one to consider alternative Markov chain Monte Carlo schemes. We propose and investigate a method for posterior simulation based on an independence chain. The method is customized to exploit the structure of the variance component model, and it works with arbitrary prior distributions. As a default reference prior, we consider a version of Jeffreys prior based on the integrated ("restricted") likelihood. We demonstrate the ease of application and flexibility of this approach in several familiar settings, even in the presence of unbalanced data. This work is joint with Rob Kass, Carnegie Mellon University.

Poster Session Abstracts

Author: **Aviles, Ana I.**

Title: Optimal Assembled Designs with Variance Components

Abstract: An assembled design is a hybrid design which places a nested design at each design point in a crossed design. This allows for the estimation of both main and interaction effects as well as variance components that arise from nested random factors such as lot number and batch number. Assembled designs are also discussed for the estimation of dispersion effects on variance components. Optimal assembled designs are constructed and discussed for many practical situations.

Authors: **Balzarini, Monica, and Raul E. Macchiavelli**

Title: On Mixed AMMI Models for Investigating Genotype-Environment Interaction

Abstract: Additive main effects multiplicative interaction (AMMI) models provide a tool for analyzing genotype-environment interaction in plant breeding. Typically, they have been used in a fixed effects model framework for the analysis of complete genotype by environment data sets. Assuming environments as random effects, multiplicative interaction models can be estimated in a mixed model framework for incomplete data. Under normality, parameter estimates are obtained by maximum likelihood based procedures. A factor analytic covariance matrix has been used to model the structure of the multiplicative interaction terms within an environment. Both fixed and mixed AMMI models are discussed. Biplots under both approaches are presented for a complete set of data from a plant breeding experiment. Although the differing procedures to obtain the biplots under both approaches, they showed same interaction pattern.

Author: **Chiang, Andy K. L.**

Title: A Simple General Method for Constructing Confidence Intervals for Functions of Variance Components

Abstract: A single simple general method is proposed for constructing confidence intervals for arbitrary functions of variance components in balanced normal theory models. The method produces the commonly known exact (Chi-Squares and F distribution based) confidence intervals for expected mean squares and ratios of them. The concept of "surrogate variables" is introduced as part of the description of the method. "Equal-tail" and "shortest-length" confidence intervals from this method can be easily computed using Monte Carlo simulations. The two-way random effects model without interaction is considered. Using computer simulations, it is shown that the proposed method produces intervals that maintain the nominal confidence level and have comparable or smaller average lengths than those produced by the best existing methods.

Author: **Hartless, Glen, James G. Booth, and Ramon C. Littell**

Title: Local influence diagnostics for the prediction of random effects in a linear mixed model

Abstract: The local influence approach to diagnostics developed by Cook (1986) provides a general method for assessing how slight perturbations to problem formulation can influence maximum likelihood (ML) inference. In this poster, Cook's likelihood displacement measure is extended to the problem of prediction through the use of the profile predictive likelihood. It is shown that application to the linear mixed model with known variance is straightforward. In particular, we will demonstrate the use of three perturbation schemes for assessing influence on the estimation of fixed effects, random effects and linear combinations thereof. The three schemes perturb the error variance, random effects variance and the response, respectively. The techniques are illustrated with an example.

Author: **Parmet, Yisrael**

Title: Variance Components in disassembly and reassembly experiments

Abstract: One of the main demands of a production process is stability. The typical modern industrial product is a complicated combination of different types of components. This research is stimulated by the problem of detecting a component that systematically causes degradation and variance of the product. The basic problem is to find components that have major influence on the variance of assembled products in order to stabilize the process, reduce variance and improve quality. Taguchi (1987) proposed so called disassembly and reassembly experiments. In the simple case when we examine only two components, A and B , the two components together with the time factor yield a three factor experiment (where each factor has m levels) with only m^2 observations. Such a design is known as the Latin Square or m^{3-1} design. Since component units used in the experiment design were sampled from a large population of units, it is natural to consider the factors A and B as random effects. The influence of each factor is naturally measured by its variance component. The larger the variance component of the factor the stronger its effect on the process variance. Thus, we may test the significance of each factor by estimating the variance components from the data. We present the model resulting from this experiment as a special case of mixed linear model and compared several estimators for the variance components. We also suggested and compared several ways to extend the basic Latin square design for disassembly and reassembly experiments.

Author: **Smith, Stephen P.**

Title: Some Facts that Illustrate the Connection Between REML and the Cholesky Decomposition

Abstract: A well known fact is that the Cholesky decomposition leads to computation of the likelihood function for REML. The Cholesky decomposition is an interesting data mining tool for treating large, sparse and

symmetric matrices, and when used to compute the REML likelihood function the availability of derivatives by backward differentiation permits maximum likelihood estimation. It is less well known, but when singular variance-covariance matrices are encountered (such as with Kalman filtering) and the mixed model equations can not be formed, the Cholesky decomposition continues to lead to likelihood evaluation. However, the Cholesky decomposition must be applied to a matrix that is symmetric and unfortunately indefinite. Better decomposition algorithms are available to factorize indefinite matrices (such as the Bunch-Parlett method), but these approaches pre-assume existence of the Cholesky decomposition as an intrinsic part of likelihood evaluation. Alternatively, even when the Cholesky algorithm fails because of the property of indefiniteness, the aborted calculations continue to permit likelihood evaluation for REML. This poster describes some of the matrix facts that have lead to the above revelations.

Author: **Unver, Yakut**

Title: Comparison of Variance Component Estimation Methods for Clutch Pattern in Laying Hens

Abstract: A study was carried out to investigate the efficiency of different variance component estimation methods to determine the genetic parameters for clutch traits in laying hens (sequence length and delay). Most of variance component estimation methods have been developed to solve the problems in animal breeding application. Recently a wide array of methods have been developed to estimate variance components. For instance, Henderson Method's, ML (Maximum Likelihood), REML (Restricted Maximum Likelihood), DFREML (Derivative Free Restricted Maximum Likelihood), MIVQUE (Minimum Variance Quadratic Unbiased Estimation), Bayesian Method, Gibbs Sampling. In this study we compared a mostly used methods in animal breeding studies such as TYPE I and TYPE III of Henderson's Method, ML, REML, DFREML, MIVQUE, MINQUE0. Data were obtained from a commercial sire line in Turkey. A total 1980 animal with 43 sire and approximately 8 dam per sire was used. The hens were housed in individual cages and selected based on records made between 22 and 40 week of age. Eggs were collected daily and individually. At the end of the period, hens have more than ten pause day were excluded from the data set. Each sequence length consists of consecutively laid egg number and delay was calculated pause day between sequence lengths. Mean of sequence length and delay were obtained 7.41 and 1.13 respectively. The heritability estimates of sequence length were 0.11 (TYPE I), 0.12 (TYPE III), 0.13 (ML), 0.13 (REML), 0.40 (DFREML), 0.14 (MIVQUE) and 0.10 (MINQUE0). The heritability estimates of delay was 0.03 for all estimation methods except DFREML was 0.05. Because of including pedigree information DFREML can explain the variability and relationship better.

Last modified: Wed Dec 29 08:52:36 EST 1999