

The Editor has agreed to publish these abstracts of papers presented during the 1990 academy meeting held at George Mason University. In the future, late abstracts will not be accepted.

Environmental Sciences

SUBSURFACE MICRO-IRRIGATION OF ROW CROP IN VIRGINIA. Norris L. Powell, Dept. of Crop & Soil Environ. Sci., Tidewater Agricultural Experiment Station, VPI & SU, Suffolk, VA 23437, and F. Scott Wright*, Tidewater Agricultural Experiment Station, USDA-ARS, Suffolk, VA 23437. By burying micro-irrigation tubing fourteen to sixteen inches below the soil surface a row crop such as corn and peanut can be easily irrigated. When compared with overhead sprinkler irrigations this offers the advantages of more efficient use of water, lower labor requirement, lower energy requirements, more efficient use of water, ease of automation, reduced water runoff (from irrigation) reduced soil erosion (caused by irrigation) and the system can be phased in overtime. Initial investment cost varies with the distance between the micro-irrigation tubing lines buried below the crop and is equal to or higher than center pivots utilizing the full circle or the hose tow traveling guns. Operating cost is one-half (medium pressure center pivot) to one-fifth (hose tow traveling gun) of the cost of the overhead sprinkler systems. Between 1986 and 1989 corn yields were increased by five (1989) to 384 percent (1987) with the use of subsurface micro-irrigation when compared with no irrigation. For peanut the yield increases were nine (1986) to 18 percent greater with the use of subsurface micro-irrigation when compared with no irrigation. A permanently installed subsurface micro-irrigation system should last longer than 10 years with proper management.

Statistics

RELATIONSHIPS BETWEEN A SURVIVAL MODEL AND LOGISTIC REGRESSION: EXAMPLES IN MODELING OUTCOMES OF MEDICARE DATA. R. Clifton Bailey, Statistical Advisor, Health Standards and Quality Bureau, ME 2-D-2, Health Care Financing Administration, 6325 Security Blvd., Baltimore, MD 21207. In the HCFA publication, Hospital Mortality Information, a logistic regression was used to risk adjust the mortality status of medicare beneficiaries 30 days post admission. In an effort to develop a more informative evaluation of mortality status, I have investigated the use of survival models to characterize the mortality status of patients over time. This presentation will review the background of the problem and demonstrate some useful relationships between survival models and logistic regression. Methods will be demonstrated for using logistic regression to evaluate a survival model for a specific time post admission. Furthermore, as a byproduct of the evaluation one obtains polishing factors which can be used to fine tune a survival model for a specific time. Furthermore, it is shown that the survival model with concomitant variables can be used to estimate the equivalent coefficients for a logistic regression. Also, results relating predictions for average values of concomitant variables with the results of averaging individual predictions are presented.

EARTHQUAKES AND THE EXPONENTIAL PROBABILITY DISTRIBUTION.

Charlotte Blair, W. Michael Gentry, and Susan Zabel. Mary Baldwin College, Staunton, VA. On October 17, 1989, an earthquake measuring 6.9 on the Richter scale devastated sections of San Francisco. It is difficult to prepare for earthquakes, since they are basically random phenomena, but valuable information is gained by studying their seismic history and the time between earthquakes. Parkfield is located on the San Andreas Fault approximately midway between Los Angeles and San Francisco. Earthquakes occurred in Parkfield in 1857, 1881, 1901, 1922, 1934, and 1966. From these data, it is possible to compute the mean time between earthquakes, and then use the Exponential distribution to compute the probability of an earthquake occurring within a specified period of time. Although the use of the Exponential distribution represents an idealization of the problem at hand, the knowledge gained is valuable.

SAMPLING PROBLEMS IN PRE-ELECTION POLLS. A. Richard Bolstein, Center for Computational Statistics, George Mason University, Fairfax, Virginia 22030. A 1988 validated presidential pre-election poll of registered voters in a small city was used to compare the likelihood to vote among the respondent and various non-respondent groups, to predict the likelihood to vote of individual respondents, to compare different bases of respondents as predictors of the election outcome, and to estimate the effect of non-respondents and undecided respondents on the prediction. A 1989 validated pre-election poll was also conducted using the respondents from the 1988 poll to predict the outcome of the Virginia gubernatorial election. The impact of the abortion and racial issues on the outcome are discussed.

TESTS OF HYPOTHESES BASED ON RANKS IN THE STANDARD MULTIVARIATE LINEAR MODEL. Enoch B. Bortey, Dept. of Biostat., Va. Commonwealth Univ., Richmond, Va. 23298-0032. A unified approach is developed for testing hypotheses in the standard multivariate linear model based on the ranks of the residuals. Hypotheses concerning a subset of specified parameters can be tested, while the remaining parameters are treated as nuisance parameters. Asymptotically, the test statistic is shown to have a chi-square distribution under the null hypothesis. This result is then extended to cover a sequence of contiguous alternatives from which the Pitman efficacy is derived. The general application of the test requires the consistent estimation of a functional of the underlying distribution and one such estimate is furnished.

MEASURES OF PROBABILITY JUDGMENT PERFORMANCE: THE PROBLEM OF ACCURACY IN SOCIAL PERCEPTION AND PREDICTION. Susan E. Brodt, Darden Graduate School of Business, Univ. of Va., Charlottesville, Va. 22906-6550. Traditional measures of probability judgment performance used to verify forecasts about events in the physical domain (e.g., weather forecasts) offer both opportunities and pitfalls for scientists studying social judgment and prediction. Measures such as the Brier score and its various decompositions (e.g., calibration, resolution, reliability-in-the-small), provide statistical rigor and insight into the study of social cognition. Researchers' attention shifts away from problems of criterion verification and reliance on outcome accuracy, toward examination of the external validity of one's beliefs or the appropriateness of one's confidence in his/her knowledge. Ironically, if social scientists overlook methodological and statistical requirements and de facto definitions of accuracy these measures employ, one potential liability may be inevitable—judgmental overconfidence. Sources of benefit and potential cost are discussed.

STRONG MODERATE DEVIATION THEOREMS FOR m -DEPENDENT RANDOM VARIABLES. Narasinga Rao Chaganty, Dept. of Mathematics and Statistics, Old Dominion University, Norfolk, Va. 23529. Consider a stationary sequence $\{X_n, n \geq 1\}$ of m -dependent random variables. Let $S_n = \sum_{i=1}^n X_i$ be the partial sum. Under some moment conditions, we asymptotic expressions for the probability of moderate deviations, $P(S_n > x_n)$, where $x_n = O(\sqrt{\log(n)})$. These extend some well known results for independent and identically distributed sequences of random variables.

A PARTIALLY-WEIGHTED GENERALIZED MULTIVARIATE ANALYSIS OF VARIANCE MODEL. Vernon M. Chinchilli, Dept. of Biostatistics, VA Commonwealth Univ., Richmond, VA 23298-0032, & Mary Hall Gregg, Merck Sharp & Dohme, Chaussee De Waterloo 1135, 1180 Brussels, Belgium. We examine the generalized multivariate analysis of variance model in terms of a partially-weighted analysis, in which we consider only a subset of the available covariates. We develop estimation and hypothesis testing for the partially-weighted model under multivariate normality, and we establish conditions under which the partially-weighted estimator is more efficient than the weighted and unweighted estimators. Also, we show that for an important subclass of general linear hypotheses, the test statistics are invariant to the choice of transformation matrix in the partially-weighted model. We propose a procedure for selecting an appropriate subset of covariates, which is based on an examination of likelihood ratio statistics from a best subsets regression and where each likelihood ratio statistic is a test of the independence of two multi-normal vectors. We provide a modification to multivariate data sets with missing values via estimated generalized least squares. Also, we indicate how a partially-weighted analysis is applicable to longitudinal data analysis or repeated measurements regression.

DISTRIBUTION-FREE TESTS FOR INTERACTION IN A TWO-WAY DESIGN. Tsui-Hsien (Joanna) Chien*, Dept. of Biostatistics, Va. Commonwealth Univ., Richmond, Va. 23298-0032. The usual test for interaction effect in a two-way layout is based on the assumptions of independence of the observations, constancy of variance and normality. When the assumptions do not hold, it is useful to have available other inferential procedures with less restrictive assumptions. In this paper, two distribution-free test statistics based on U-statistics are discussed. One measures the magnitude of the interactions while the other measures 'directional' interactions. There is an equivalent relationship between these two test statistics and both of them have asymptotic chi-square distributions. These two tests are shown to be consistent and the asymptotic relative efficiencies of them are also studied. At the end, a simulation study is conducted to investigate the relative performance of the tests with smaller sample sizes, and some numerical examples are presented.

METHODS FOR COUNTING RARE AND ELUSIVE POPULATIONS. Charles D. Cowan, Chief Statistician, Opinion Research Corporation, 500 E. Street, SW, Suite 940, Washington, DC, 20024. Statisticians often want to draw inferences regarding populations that are hard to find, ill-defined, or just elusive. Examples of such populations are the homeless, missing children, users of particular social services (such as agricultural giveaway programs), and migrant workers. At times, the populations are well defined but not readily observable because they are very rare and there is no register or list with which to locate them. An example would be persons with relatively rare genetic disorders. This presentation looks at four procedures for finding and enumerating rare populations: sampling, multiplicity techniques, multiple frame techniques, and multiple capture techniques. The models underlying each technique will be compared, and the costs and benefits of use of each technique will be explored. Some examples of studies will be given based on the author's experiences.

OUTLIERS AND INTERACTION-IS THERE A CONNECTION? Barbara R. Kuzmak and Eric P. Smith', Dept. of Statistics, Va Polytechnic Inst., Blacksburg, VA 24061. Outliers and interaction are generated by different mechanisms. An outlier may be produced by a recording error, equipment failure, violation of model assumptions, etc., whereas an interaction is the result of a synergistic effect among several factors. Interaction reveals important knowledge about the system under investigation, but an outlier usually does not convey useful information. However, in a nonreplicated two-way analysis of variance, outlier effects appear as interaction. The additive plus multiplicative model, $Y_{ij} = \mu + \alpha_i + \beta_j + \Sigma \lambda_{ij} \tau_{ij}$, has been used to describe multiplicative interaction in an unreplicated experiment. We use this model in the same setting to study outliers. In data sets with significant interaction, one may be interested in determining whether the cause of the interaction is due to a true interaction, outliers or both. We developed a new technique which can show how outliers can be distinguished from interaction when there are simple outliers in a two-way table. Several examples illustrating the use of this model to describe outliers and interaction are presented.

ON THE OPTIMALITY PROPERTIES OF SETS OF YODEN DESIGNS. J. P. Morgan, Dept. of Math. & Stat., Old Dominion Univ., Norfolk, Va. 23529. A generalized Youden design $GYD(v, p, q)$ is a $p \times q$ row/column design for v treatments for which each of rows and columns forms a balanced block design. A GYD is regular if either p or q is a multiple of v , and otherwise is non-regular. That regular GYD's are universally optimum, and non-regular GYD's are A-, E-, and D-optimum, was shown by Kiefer (1975). Here some optimality properties of sets of GYD's are established when $b > 1$ $p \times q$ designs are to be used. In particular, in the regular case b GYD's are universally optimum, while in the non-regular case counterexamples are given to show that A- and D-optimality are lost. Some results on the E-behavior of $b = 2$ GYD's are also given.

ANALYSIS OF MISSING DATA UNDER NON-RANDOM MECHANISMS: BAYESIAN INFERENCE.

Patricia A. Pepple, Department of Mathematical Sciences, Virginia Commonwealth University, Richmond, Va. 23284, & Sung C. Choi, Department of Biostatistics, Virginia Commonwealth University, Richmond, Va. 23298. Inferences are made concerning population proportions when data are not missing at random. Both one-sample and two-sample situations are considered with examples in clinical trials. The one-sample situation involves the existence of response related missing data in a study conducted to make inferences involving the proportion. The two-sample problem involves comparing two treatments in clinical trials when there exists nonrespondents due to both the treatment and the response to the treatment. Bayes procedures are used in estimating parameters of interest and testing hypotheses of interest in these two situations. An ad-hoc approach to the classical inference is presented for each of the two situations and compared with the Bayesian approach discussed. To illustrate the theory developed, data from clinical trials of severe head trauma patients at the Medical College of Virginia Head Injury Center from 1984 to 1987 is considered.

DIFFERENCES BETWEEN ATMOSPHERIC OZONE PROFILES DERIVED FROM OZONE SONDES AND THE SAGE I AND II SATELLITE INSTRUMENTS. Robert E. Veiga,

S. T. Systems Corp., 28 Research Drive, Hampton, VA 23666. Ozone sonde profiles from 14 stations have been spatially and temporally collocated with satellite ozone profiles measured by SAGE I and SAGE II over the period 1979-1989. Estimates of difference profiles measured with respect to the satellite observations indicate three altitude dependent types of behavior. From 10-14 km the differences increase from values as small as -50% to large positive values ranging from 10-40%. A discontinuity exists in the difference profile at 14 km. From 15 km to the ozone density peak the satellite measurements are larger than the in situ measurements, peaking at 16 km with values ranging from -5% to -35%. From 16 km to the peak of the ozone density the relative differences decrease to zero. From the ozone maximum to 27 km there are no statistically significant differences. Above 27 km the ozone sonde values tend to be lower than the satellite measurements, a behavior consistent with sonde pump efficiency losses at low ambient pressures. Regression analyses using monthly means over an 11 year period from satellite data and observations from two specific ground stations were performed in order to contrast ozone trends. The results show that an ozone decrease occurred from 15-25 km.