LogXact® PROCs

The Most Powerful Software for Exact Logistic Regression, and the Only One for Exact Poisson Regression.

Why You Need Exact Logistic and Poisson Regression

The conventional method of estimating logistic regression parameters is via large sample approximation. However, the conventional method can be inaccurate and misleading if you are modeling small, sparse or imbalanced data.

These inaccuracies are not just a matter of the third or fourth significant digit. Conventionally derived p-values can differ from their exact counterparts by orders of magnitude and by large absolute amounts as well. King and Ryan published a logistic regression example in the *American Statistician* (2002, v. 56) reviewing an earlier study of the rate at which red blood cells settle out of suspension in blood. The sedimentation rate (the independent variable) was coded as a binary variable (i.e. above or below a cutoff level), and the predictor variables were the blood levels of Fibronogen and gamma-globulin. Exact logistic regression yielded a p-value of 0.001, indicating a highly significant relationship. Conventional logistic regression, in stark contrast, yielded a p-value of 0.439, and a conclusion completely opposite to the correct conclusion.

Choosing an inappropriate analysis method can transform significant effects into non-significant ones, and vice-versa, and how would you know? In some cases, the conventional method produces no answer at all. LogXact PROCs' exact methods will never produce an incorrect answer regardless of how irregular your data are.

SAS® Already has PROC LOGISTIC Regression – Why do I Need LogXact PROCs?

LogXact PROCs is the most powerful exact logistic and Poisson regression software available. SAS's logistic regression tool is sufficient for only a limited range of problems. Cytel's LogXact solves more problems, and solves more problems faster than any other package. The exact inferences in PROC LOGISTIC relies upon algorithms published a decade ago. Research has moved forward considerably since that time.

LogXact PROCs Features

- Computes exact p-values and confidence intervals for small or unbalanced data sets, and for large data sets with very low response rates
- Uses new algorithms (modified version of Mehta, Patel and Senchaudhuri, JASA, March 2000) to speed up computations by a factor of 50 to 1000 relative to LogXact 4
- Powerful Monte Carlo procedures enable fast exact inference for much larger data sets
- Exploration Mode tells you how long a problem will take before you commit computing resources to it, and also lets you customize the exact algorithms to optimize the use of time and memory for difficult problems
- The only package with exact inference for Poisson regression
- Handles matched case-control data under general M:N matching, by conditional likelihood inference
- Runs on Windows NT®/2000/XP, SAS v8.0, 8.1, 8.2 and SAS v9.

$$P_1(\bar{R}_j^{opt}) \geq P_1(\bar{R}_j) \text{ for all } j$$

Cytel
STATISTICAL SOFTWARE

LogXact PROCs now solves problems that are far too difficult for the earlier algorithms.

Six problems brought to us by practicing statisticians are shown in the following table. Each problem tested LogXact against the PROC LOGISTIC procedures in SAS Version 9. You can try these tests yourself. All of the data sets are available for download at http://www.cytel.com/LogXact/examples.asp.

The Benefits of Having LogXact's Capabilities Available as a SAS PROC

LogXact PROCs works seamlessly in your SAS environment. Simply call up LogXact PROCs as you would call up any SAS procedure from within SAS itself. No new software to learn, no clumsy transferring of data from one program to another. Since LogXact PROCs lets you perform the entire data analysis within the SAS system, the batch file of SAS commands used in the analysis serves to document each step of the report.

Execution Times for Cytel's LogXact and SAS's PROC-LOGISTIC

Data Set	SAS's PROC LOGISTIC	Cytel's LogXact	
		LogXact 4	*LogXact 5*
Crash	Can't solve	32 min.	2 min.
UTI	Can't solve	Can't solve	11 min.
Death Penalty	Can't solve	Can't solve	15 min.
Gene	Can't solve	Can't solve	6 min.
Titanic	Can't solve	Can't solve	4 min.
Diarrhea	Can't solve	Can't solve	12 min.

CRASH: Data from 58 simulated car crashes were analyzed; the relationship between the crash outcome (fatal, non-fatal) and 3 covariates was modeled.

UTI: Sexually active college women were studied to determine how the development of first-time urinary tract infection (UTI) is related to contraceptive use. (437 observations, 9 covariates)

DEATH PENALTY: This study was conducted to determine whether race was a factor in applying the death penalty in NJ between 1982 and 1998. (160 observations, 10 covariates.)

GENE: A case-control study of colon cancer investigated the relationship between the presence/absence of a genetic mutation and the presence/absence of colon cancer. (771 observations, 4 covariates.)

TITANIC: The records of the sinking of the Titanic were studied to establish the relationship between survival and status on the ship (1st class, 2nd class, crew, etc.). (2201 observations, 3 covariates, 1 covariate factored.)

DIARRHEA: Hospital patients were studied to determine the relationship between antibiotic use and an acute form of diarrhea. (2493 observations, 10 covariates.)

Anyone who routinely performs logistic regression and/or Poisson regression needs to have a LogXact. Those who regularly use these techniques will almost certainly encounter situations which will not be correctly handled by any other statistical software package.

Robert A. Oster, Ph.D.,
***The American Statistician*,**
August 2002 [Vol. 56, No. 3, Page 246]

Cytel Software Corporation
675 Massachusetts Avenue
Cambridge, MA 02139-3309

www.cytel.com
sales@cytel.com
T: +1.617.661.2011
F: +1.617.661.4405

$$P_1(\bar{R}_j^{opt}) \geq P_1(\bar{R}_j) \text{ for all } j$$