

COVAR: Computer Program for Multifactor Relative Risks and Tests of Hypotheses Using a Variance-Covariance Matrix from Linear and Log-Linear Regression

Leif E. Peterson
Center for Cancer Control Research
Baylor College of Medicine
Houston, Texas 77030

November 3, 1997

Abstract

A computer program for multifactor relative risks, confidence limits, and tests of hypotheses using regression coefficients and a variance-covariance matrix obtained from a previous additive or multiplicative regression analysis is described in detail. Data used by the program can be stored and input from an external disk-file or entered via the keyboard. The output contains a list of the input data, point estimates of single or joint effects, confidence intervals and tests of hypotheses based on a minimum modified chi-square statistic. Availability of the program is also discussed.

Key Words: Relative Risks, Confidence Limits, Matrix Manipulation, Hypothesis Testing

1 INTRODUCTION

Additive and multiplicative models are widely used in epidemiology to obtain maximum likelihood (ML) estimates of relative risks (RR), odds-ratios (OR) and hazard ratios (HR). Numerous computer programs have been developed that generate ML estimates of regression coefficients and variance-covariance matrices [1, 2, 3, 4, 5]. Despite this wide diversity of programs, there is a shortage of stand alone algorithms that allow the statistician to generate multifactor relative risks and confidence limits, and conduct tests of hypotheses using results of previous regression analyses.

This report documents a computer program to obtain point estimates and confidence intervals (CI) of RR for single or joint effects using the regression coefficients and a variance-covariance matrix from a previous modeling session.

The program has been developed mainly for Grizzle-Starmer-Koch (GSK) regression models for multinomial categorical data, Cox proportional hazards (PH), logistic and Poisson regression results, but can be used with the results of other models. The critical data required to apply the computer code are:

- Number of parameters
- Variable names and regression coefficients
- Variance-covariance matrix
- Probability $(1-\alpha)$ of the confidence interval of RR
- Number of coefficients used in RR or tests of hypotheses
- Units applied to each coefficient in RR estimation

There is only one output format, listing the regression coefficients, their names, variance-covariance matrix, point estimates, corresponding CIs, minimum modified χ^2 tests of hypotheses with one-sided tail probabilities and units applied to each coefficient. The program and all of its subroutines are written in FORTRAN-77 and are run on a desktop PC with an AMD-K6-166MHz chip. The algorithm may be compiled and run on Unix machines as

well. The PC version of the executable requires about 176,000 bytes of random access memory (RAM) to operate. Average execution time is about 1 second per run.

2 PROGRAM DESCRIPTION

The computer program proceeds as follows. The method of entering parameter names, regression coefficients, and the variance-covariance matrix is specified by the user. The choices are to enter the data from the keyboard or from an external disk-file. If the keyboard is selected as the means of entering data, the user is queried for all of the necessary data. When using the disk-file for data entry, the number of parameters, p , is first read in, followed by the p variable names and p regression coefficients. This is followed by entry of $((p^2 - p)/2) + p$ elements of the variance-covariance matrix. One should note that, since the variance-covariance matrix is symmetric, only the lower triangular is required.

Once the number of parameters, variable names, coefficients and variance-covariance matrix are entered via the keyboard or disk-file, the user must enter the number of coefficients used in RR or tests of hypotheses to be considered for this run. For each coefficient used in RR estimation, i.e., variable, the user must specify the variable index and the units to be considered. Next, the user must specify for each factor one of the following three options: (1) estimate RR and CI, (2) use contrast matrix for test of hypothesis about $\mathbf{b}=\mathbf{0}$, or (3) use a contrast matrix to test for a trend in $\mathbf{b}=\mathbf{0}$. If RR and CI is specified, the program queries the user for the probability to be used for estimating the CI of RR. The choices are 80%, 90%, 95%, 97.5% and 99%, with the last choice left as an option to enter a probability that is different from those listed above.

2.1 NOTATION AND THEORY

Thus far we have discussed the program's necessary input data. In this section we will address the relationship between asymptotic likelihood theory and the ML estimation of the regression coefficients and variance-covariance matrix. According to asymptotic likelihood theory, the score vectors of a regression model are

$$U(\mathbf{b}) = \frac{\partial \log L(\mathbf{b})}{\partial \mathbf{b}}, \quad (1)$$

where $\log L$ is the log of the likelihood function for the particular model under consideration, that is, the GSK, Cox proportional hazards, logit, Poisson or other model. The Fisher information matrix $I(b)$ of the score vectors is

$$\begin{aligned} \mathbf{I}(b) &= E[U(\mathbf{b})U^T(\mathbf{b})] \\ &= -\left[E\left(\frac{\partial^2 \log L(\mathbf{b})}{\partial \mathbf{b}_j \partial \mathbf{b}_k} \right) \right], \end{aligned} \quad (2)$$

the generalized-inverse of which is the variance-covariance matrix, V_b , written as

$$\mathbf{V}_b = \begin{pmatrix} \sigma^2(b_0) & \sigma(b_0, b_1) & \cdots & \sigma(b_0, b_j) \\ \sigma(b_0, b_1) & \sigma^2(b_1) & \cdots & \sigma(b_1, b_j) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2(b_0, b_j) & \sigma(b_1, b_j) & \cdots & \sigma^2(b_j) \end{pmatrix}, \quad (3)$$

which is a $p \times p$ real symmetric matrix whose principal diagonal consists of the parameter variances. The off-diagonals are the covariances between each of the coefficients. Most computer programs print the diagonal and the upper or lower triangular of the variance-covariance matrix, which makes no difference since the matrix is symmetric. When Newton-Raphson iteration is used for convergence, \mathbf{V}_b is the generalized-inverse of the Fisher information matrix. For iteratively reweighted least squares (IRLS) regression, \mathbf{V}_b is the generalized-inverse of the weighted dispersion matrix shown as

$$\begin{aligned} \mathbf{V}_b &= \mathbf{I}(\mathbf{b}) \\ &= (\mathbf{X}^T \mathbf{V} \mathbf{X})^{-1}. \end{aligned} \quad (4)$$

A best-asymptotic-normal (BAN) estimate of the parameter vector \mathbf{b} is obtained by adding the solution vector

$$\Delta \mathbf{b} = (\mathbf{X}^T \mathbf{V} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V} \mathbf{Y}, \quad (5)$$

to the previous parameter vector at each iteration

$$\mathbf{b}_{(i+1)} = \mathbf{b}_i + \Delta \mathbf{b}_{(i+1)}, \quad (6)$$

which results in \mathbf{b} at convergence

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}. \quad (7)$$

In (5) \mathbf{X} is the design matrix, \mathbf{W} is a diagonal matrix of inverse variance weights and \mathbf{Y} is the column vector of residuals. Using \mathbf{b} and \mathbf{V}_b from, for example, a four-variable model including terms for an intercept b_0 and three risk factors (b_1, b_2 , and b_3), the point estimate of RR for the three factors considered jointly is

$$RR = \exp(b_1 u_1 + b_2 u_2 + b_3 u_3), \quad (8)$$

where b_1, b_2 , and b_3 are the regression coefficients and u_1, u_2 and u_3 are the units applied to the coefficients. The RR in (8) is based on reference cell coding where each p -level risk factor is coded with $p-1$ independent variables. The Taylor series-based $(1-\alpha)$ confidence interval (CI) for the interval estimate can be expressed by

$$(1-\alpha)CI \ RR = \exp(\ln(RR) \pm Z_\alpha \times \sqrt{V}). \quad (9)$$

Abramowitz and Stegun [6] give a useful method for determining percentage points of the standard deviate, Z_α , in (9) for a specified a level in the form

$$Z_\alpha = t - \left[\frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 + d_2 t^2 + d_3 t^3} \right] \quad (10)$$

where

$$t = \sqrt{\ln\left(\frac{1}{\alpha^2}\right)}, \quad (11)$$

and

$$\begin{aligned} c_0 &= 2.515517 & d_1 &= 1.432788 \\ c_1 &= 0.802853 & d_2 &= 0.189269 \\ c_2 &= 0.010328 & d_3 &= 0.001308. \end{aligned}$$

The variance, V , of the logarithm of the RR (9) takes on the form

$$V = \sum_i^p \sigma^2(b_i)(u_i^* - u_i)^2 + \sum_j^p \sum_{i \neq j}^p \sigma_{ij}(b_i, b_j)(u_i^* - u_i)(u_j^* - u_j), \quad (12)$$

where $\sigma^2(b_i)$ is the variance of b_i , $\sigma_{ij}(b_i, b_j)$ is the covariance between b_i and b_j (3) and $(u_i^* - u_i)$ and $(u_j^* - u_j)$ are the change in units applied to the coefficients. A test of hypothesis of no significant effect due to the joint factors of the parameters in (8) is

$$H_o : b_1 = b_2 = b_3 = 0, \quad (13)$$

which can be rewritten as

$$H_o : \mathbf{C} \mathbf{b} = \mathbf{0}, \quad (14)$$

where \mathbf{C} is the contrast matrix forming the desired linear combination of the regression coefficients for the three-parameter test. Thus, for the three joint factors in (8), the \mathbf{C} matrix is

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (15)$$

Grizzle, Starmer, and Koch [7] introduced the test statistic of the linear hypothesis $H_0 : \mathbf{Cb} = \mathbf{0}$, given by

$$\chi^2 = \mathbf{SS}(\mathbf{Cb} = \mathbf{0}) = \mathbf{b}^T \mathbf{C}^T [\mathbf{CV}_b \mathbf{C}^T]^{-1} \mathbf{Cb}, \quad (16)$$

which is χ^2 distributed with degrees of freedom equal to the number of independent rows in \mathbf{C} . Matrix multiplication in (16) is first carried out by obtaining the outer products $\mathbf{b}^T \mathbf{C}^T$ and \mathbf{Cb} . The variance-covariance matrix \mathbf{V}_b is premultiplied by \mathbf{C} to give \mathbf{CV}_b , which is postmultiplied by the transpose of \mathbf{C} (\mathbf{C}^T) to give $\mathbf{CV}_b \mathbf{C}^T$. The inverse of $\mathbf{CV}_b \mathbf{C}^T$, that is, $[\mathbf{CV}_b \mathbf{C}^T]^{-1}$, is obtained by using singular value decomposition, which avoids such problems as nullity (zero elements), singularity and collinearity, which may be inherent in ill-conditioned or “sparse” matrices. Lastly, $[\mathbf{CV}_b \mathbf{C}^T]^{-1}$ is premultiplied by $\mathbf{b}^T \mathbf{C}^T$ then postmultiplied by \mathbf{Cb} to arrive at $\mathbf{b}^T \mathbf{C}^T [\mathbf{CV}_b \mathbf{C}^T]^{-1} \mathbf{Cb}$. If the test statistic exceeds some tabled value of $\chi^2_{(d.f., 1-\alpha)}$, then we conclude that there is a statistically significant joint effect in the presence of the three factors in our example.

2.2 STRUCTURE

```
CALL VAR(N,A,B,CON,M,U,X,Z,ITRND,BEGTRND,ENDTRND,
ORDTRND,ITRD,ITER,IFLT,PLUS,RR,LL,UL,CHI2,PROB)
```

NP	Parameter	Work:	Maximum number of parameters
NTITER	Parameter	Work:	Maximum number of cycles (runs per session)
INFILE	Character	Input:	Input filename (optional)
OUTFILE	Character	Input:	Output filename
NCOL	Integer	Input:	Number of parameters
B(NCOL)	Real	Input:	Column vector of coefficients, \mathbf{b}
SE(NCOL)	Real	Input:	Standard error of coefficients
X(NCOL,NCOL)	Real	Input:	Variance-covariance matrix, \mathbf{V}_b
Z	Real	Input:	Standard normal deviate
NEFF	Integer	Input:	Number of coefficients used in RR or tests of hypotheses
AUNIT(NCOL)	Real	Input:	Units applied to coefficients
BEGTRND(NCOL)	Integer	Input:	First coefficient in trend test
ENDTRND(NCOL)	Integer	Input:	Last coefficient in trend test
IRTEST(NTITER)	Integer	Input:	Type of test for each row of \mathbf{C} matrix
ITEST(NCOL)	Integer	Input:	Dummy variable for test of trend during each cycle (run per session)
ITER	Integer	Work:	Cycle, or number of tests performed
B1(NCOL,1)	Real	Work:	Column vector of coefficients, \mathbf{b}
BT(1,NCOL)	Real	Work:	Transpose of \mathbf{b} vector
C(NCOL,NCOL)	Real	Work:	Contrast matrix \mathbf{C}
CON(NCOL,NCOL,NTITER)	Real	Work:	Contrast matrix written at each cycle
CB(NCOL,1)	Real	Work:	Cross-product of \mathbf{C} matrix and \mathbf{b} vector, \mathbf{Cb}
CT(NCOL,NCOL)	Real	Work:	Transpose of \mathbf{C} matrix, \mathbf{C}^T
CX(NCOL,NCOL)	Real	Work:	Cross-product of \mathbf{C} and \mathbf{V}_b matrices, \mathbf{CV}_b
CXC(NCOL,NCOL)	Real	Work:	Cross-product of \mathbf{CV}_b and \mathbf{C}^T matrices, $\mathbf{CV}_b \mathbf{C}^T$
CXC1(NCOL,NCOL)	Real	Work:	Generalized inverse of $\mathbf{CV}_b \mathbf{C}^T$ matrix, $[\mathbf{CV}_b \mathbf{C}^T]^{-1}$
M1(1,NCOL)	Real	Work:	Cross-product of $\mathbf{b}^T \mathbf{C}^T$ and $[\mathbf{CV}_b \mathbf{C}^T]^{-1}$ matrices, $\mathbf{b}^T \mathbf{C}^T [\mathbf{CV}_b \mathbf{C}^T]^{-1}$
NUM(1,1)	Real	Work:	Cross-product of $\mathbf{b}^T \mathbf{C}^T [\mathbf{CV}_b \mathbf{C}^T]^{-1}$ and \mathbf{C}_b matrices, $\mathbf{b}^T \mathbf{C}^T [\mathbf{CV}_b \mathbf{C}^T]^{-1} \mathbf{Cb}$
PIND(NCOL)	Integer	Work:	Column vector used to set each row of contrast matrix, \mathbf{C}
V	Real	Work:	Variance of $\ln(RR)$

SD	Real	Work:	Square-root of \mathbf{V}
BU	Real	Work:	Sum-product of coefficients and units $\sum b_i UNIT_i$
PLUS(NCOL)	Character	Output:	Column vector of '+' and '-' symbols, with '+' representing the specified factors
RR	Real	Output:	Relative risk
LL	Real	Output:	Lower bound of relative risk
UL	Real	Output:	Upper bound of relative risk
CHI2	Real	Output:	Chi-square test statistic
PROB	Real	Output:	One-sided tail probability
IFLT	Integer	Output:	Premature exit code
			0=OK
			1=($Z = 0$)
			2=($NEFF = 0$)
			3=($b_i = 0$)
			4=($UNIT_i = 0$)
			5=($IND_i = 0$)
			6=($V = 0$)

2.3 AUXILIARY ALGORITHMS

The algorithms *MATMUL*, *TRNPOS*, *MATINV*, *SVDCMP* are required. *MATMUL* is used for matrix multiplication; *TRNPOS* is used to transpose matrices, both given in [8]. *MATINV* performs a check on the singular values returned from the singular value decomposition algorithm *SVDCMP* based on [9]; the function *CUMNOR* returns Z_α for a given α ; the functions *ALG* and *GAMAIN* return the χ^2 tail probabilities for a given χ^2 value and d.f.

2.4 RESTRICTION

All input parameters are checked for nullity and a fault message is returned if there is an illegal entry.

2.5 TIME

A thorough investigation of absolute timing has not been performed; however, it should be noted that execution time is a function of the number of model parameters and factors to be considered.

2.6 PRECISION

The algorithm may be converted to double precision by making the following changes:

1. Change *REAL* to *DOUBLE PRECISION* in the algorithm.
2. Change the constants to double precision.
3. Change *EXP* to *DEXP* and *ALOG* to *DLOG* in all applicable routines.
4. Make appropriate changes in auxiliary routines *MATMUL*, *TRNPOS*, *MATINV*, *SVDCMP*, and in functions *ALG* and *GAMAIN*.

2.7 APPLICATION

2.7.1 Estimating Relative Risk for Multiple Risk Factors

Let us take, as an example, the results of a unconditional (unmatched) logistic regression analysis of oral contraceptive use (OC), smoking and myocardial infarction (MI) found in [10]. The coefficients are given as

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{pmatrix} = \begin{pmatrix} -9.2760 \\ 1.2756 \\ 0.1517 \\ 1.1977 \\ 2.0828 \\ -1.1709 \\ 0.3392 \end{pmatrix} \begin{matrix} \text{Intercept} \\ \text{OC Use} \\ \text{Age} \\ \text{Smoke}(1 - 24) \\ \text{Smoke} \geq 25 \\ \text{OCuse} * \text{Smoke}(1 - 24) \\ \text{OCuse} * \text{Smoke} \geq 25 \end{matrix} . \quad (17)$$

where the variable names to the right of the coefficients are the intercept, OC use (coded 0-no, 1-yes), age (continuous variable), smoking 1-24 cigarettes per day (coded 0-no, 1-yes), smoking ≥ 25 cigarettes per day (coded 0-no, 1-yes), interaction term for OC use and smoking 1-24 cigarettes per day and an interaction term for OC use and smoking ≥ 25 cigarettes per day. The variance-covariance matrix \mathbf{V}_b for our example is

$$\mathbf{C} = \begin{pmatrix} 0.4045 & . & . & . & . & . & . \\ -0.0676 & 0.3345 & . & . & . & . & . \\ -0.0087 & 0.0008 & 0.0002 & . & . & . & . \\ -0.0387 & 0.0324 & 0.0002 & 0.0476 & . & . & . \\ -0.0470 & 0.0332 & 0.0004 & 0.0321 & 0.0489 & . & . \\ 0.0140 & -0.3294 & 0.0004 & -0.0471 & -0.0310 & 0.7212 & . \\ 0.0072 & -0.3287 & 0.0006 & -0.0313 & -0.0472 & 0.3322 & 0.4424 \end{pmatrix} . \quad (18)$$

The RR for a woman who is an OC user that smokes ≥ 25 cigarettes per day would be determined by substituting in the values of the coefficients and applying a unit of 1.0 to each in (8), we get

$$\begin{aligned} RR &= \exp(1.2756 \times (1 + 2.0828) \times (1 + 0.3392) \times 1) \\ &= \exp(3.6976) \\ &= 40.35. \end{aligned} \quad (19)$$

If we now substitute into (12) the variances and covariances for the three variables in (18), we obtain V as

$$\begin{aligned} V &= 0.3345(1 - 0)^2 + 0.0489(1 - 0)^2 + 0.4424(1 - 0)^2 + 2(0.0332) + 2(-0.3287) + 2(-0.0472) \\ &= 0.1404. \end{aligned} \quad (20)$$

The 95% CI for the RR is then calculated from (9) as

$$\begin{aligned} 95\% \text{ CI } RR &= \exp(\ln(RR) \pm Z_\alpha \sqrt{V}) \\ &= \exp(3.6976 \pm 1.96\sqrt{0.1404}) \\ &= (19.4, 84.1). \end{aligned} \quad (21)$$

Now that the RR and Taylor series-based CI have been estimated, we can conduct a simultaneous test of hypothesis for no significant effect due to OC use, smoking ≥ 25 cigarettes per day with respect to the RR :

$$H_o : b_1 = b_4 = b_6 = 0. \quad (22)$$

For the three factors, the \mathbf{C} matrix is

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (23)$$

By substituting the appropriate values into (16) and performing the matrix manipulation to obtain χ^2 we get

$$\begin{aligned} \mathbf{Cb} &= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -9.2760 \\ 1.2756 \\ 0.1517 \\ 1.1977 \\ 2.0828 \\ -1.1709 \\ 0.3392 \end{pmatrix}, \\ &= \begin{pmatrix} 1.2756 \\ 2.0828 \\ 0.3392 \end{pmatrix}, \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbf{b}^T \mathbf{C}^T &= \begin{pmatrix} -9.2760 & 1.2756 & 0.1517 & 1.1977 & 2.0828 & -1.1709 & 0.3392 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ &= \begin{pmatrix} 1.2756 & 2.0828 & 0.3392 \end{pmatrix}, \end{aligned} \quad (25)$$

$$\begin{aligned} \mathbf{CV}_b &= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.4045 & . & . & . & . & . & . \\ -0.0676 & 0.3345 & . & . & . & . & . \\ -0.0087 & 0.0008 & 0.0002 & . & . & . & . \\ -0.0387 & 0.0324 & 0.0002 & 0.0476 & . & . & . \\ -0.0470 & 0.0332 & 0.0004 & 0.0321 & 0.0489 & . & . \\ 0.0140 & -0.3294 & 0.0004 & -0.0471 & -0.0310 & 0.7212 & . \\ 0.0072 & -0.3287 & 0.0006 & -0.0313 & -0.0472 & 0.3322 & 0.4424 \end{pmatrix}, \\ &= \begin{pmatrix} -0.0676 & 0.3345 & 0.0008 & 0.0324 & 0.0332 & -0.3294 & -0.3287 \\ -0.0470 & 0.0332 & 0.0004 & 0.0321 & 0.0489 & -0.0310 & -0.0472 \\ 0.0072 & -0.3287 & 0.0006 & -0.0313 & -0.0472 & 0.3322 & 0.4424 \end{pmatrix}, \end{aligned} \quad (26)$$

$$\begin{aligned} \mathbf{CV}_b \mathbf{C}^T &= \begin{pmatrix} -0.0676 & 0.3345 & 0.0008 & 0.0324 & 0.0332 & -0.3294 & -0.3287 \\ -0.0470 & 0.0332 & 0.0004 & 0.0321 & 0.0489 & -0.0310 & -0.0472 \\ 0.0072 & -0.3287 & 0.0006 & -0.0313 & -0.0472 & 0.3322 & 0.4424 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ &= \begin{pmatrix} 0.3345 & 0.0332 & -0.3287 \\ 0.0332 & 0.0489 & -0.0472 \\ -0.3287 & -0.0472 & 0.4424 \end{pmatrix}, \end{aligned} \quad (27)$$

$$[\mathbf{CV}_b \mathbf{C}^T]^{-1} = \begin{pmatrix} 11.0866 & 0.4724 & 8.2877 \\ 0.4724 & 22.8178 & 2.7855 \\ 8.2877 & 2.7855 & 8.7153 \end{pmatrix}, \quad (28)$$

$$\begin{aligned} \mathbf{b}^T \mathbf{C}^T [\mathbf{C} \mathbf{V}_b \mathbf{C}^T]^{-1} &= \begin{pmatrix} 1.2756 & 2.0828 & 0.3392 \end{pmatrix} \begin{pmatrix} 11.0866 & 0.4724 & 8.2877 \\ 0.4724 & 22.8178 & 2.7855 \\ 8.2877 & 2.7855 & 8.7153 \end{pmatrix}, \\ &= \begin{pmatrix} 17.9373 & 49.0723 & 19.3296 \end{pmatrix}, \end{aligned} \quad (29)$$

$$\mathbf{b}^T \mathbf{C}^T [\mathbf{C} \mathbf{V}_b \mathbf{C}^T]^{-1} \mathbf{C} \mathbf{b} = \begin{pmatrix} 17.9373 & 49.0723 & 19.3296 \end{pmatrix} \begin{pmatrix} 1.2756 \\ 2.0828 \\ 0.3392 \end{pmatrix} = 131.6453. \quad (30)$$

The resulting χ^2 test statistic of 131.65 is significantly greater than the tabled $\chi^2_{(3,0.95)}$ value of 7.82, indicating that there is a significant joint effect. Thus we reject the hypothesis (22) that there is no joint effect due to OC use, smoking ≥ 25 cigarettes per day and the interaction between OC use and smoking ≥ 25 cigarettes per day.

After performing estimating relative risk or testing hypotheses, the user is asked to continue with processing. If the response is “no,” the program is terminated, otherwise the algorithm recycles and queries the user for input data for the next cycle (run).

2.7.2 Testing Hypotheses for Trends

This section describes how to use COVAR for conducting tests of hypothesis for trends in coefficients. The example is based on multinomial categorical data related to pathological outcome among 2525 thyroid surgical patients. These particular patients resided near the Semipalatinsk Test Site, Kazakstan, where the former Soviet Union conducted nuclear weapons testing from 1949-89. These data are shown in Table 1, which contains 7 columns for outcome (Goiter, Adenoma, Cancer, Hashimoto’s thyroiditis, Riedel’s thyroiditis, Dequervain’s thyroiditis, and Other).

Our goal is to determine whether there has been an increase in the proportion of thyroid cancers identified at surgery throughout the time periods. In the “Cancer” column of Table 1, one can notice the proportion of cancer out of all other outcomes increases over time. Although inferences can be made by visualizing non-zero effects and increasing trends of proportions, we nevertheless must perform hypothesis tests to discredit the null hypothesis of no effect, no interaction, and no trend.

Table 1: Frequency distribution of 2525 thyroid abnormalities among surgical patients in the Semipalatinsk Region of Kazakstan, number (%).

Calendar period	Goiter (all forms)	Adenoma	Cancer	Hashimoto’s thyroiditis	Riedel’s thyroiditis	de Quervain’s thyroiditis	Other	Total
I (1966-1971)	324 (83.9)	12 (3.1)	5 (1.3)	37 (9.6)	3 (0.8)	3 (0.8)	2 (0.5)	386 (100)
II (1972-1976)	212 (59.9)	95 (26.8)	16 (4.5)	25 (7.1)	1 (0.8)	-	5 (1.4)	354 (100)
III (1977-1981)	229 (60.4)	85 (22.4)	17 (4.5)	36 (9.5)	5 (1.3)	1 (0.3)	6 (1.6)	379 (100)
IV (1982-1986)	276 (60.1)	59 (12.8)	48 (10.5)	69 (15.0)	1 (0.2)	1 (0.2)	5 (1.1)	459 (100)
V (1987-1991)	353 (54.6)	68 (10.5)	105 (16.2)	111 (17.2)	4 (0.6)	2 (0.3)	3 (0.5)	647 (100)
VI (1992-1994)	201 (67.0)	29 (9.7)	29 (9.7)	34 (11.3)	2 (0.6)	2 (0.7)	3 (1.0)	300 (100)

First, we made a GSK regression by regressing the ratio of the count for each outcome (column) in each time period by the total row count for each time period. As an example, the first record in the data file for the GSK regression has 324/386 as the dependent variable, followed by 12/386 in the second record, 5/386 in the third record, 37/386 if the fourth, 3/386 in the fifth record, 3/386 in the sixth, and 2/386 in the seventh. The eighth record begins with 212/354 as the dependent variable, and so on. Dummy coding with 0 and 1 was used

to specify the level of the outcome, time period, and interaction between cancer and time period. Arranged in matrix notation, the column vector of coefficients is

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \\ b_8 \\ b_9 \\ b_{10} \\ b_{11} \\ b_{12} \\ b_{13} \\ b_{14} \\ b_{15} \\ b_{16} \\ b_{17} \\ b_{18} \\ b_{19} \end{pmatrix} = \begin{pmatrix} 0.5557 \\ 0.0618 \\ 0.0013 \\ 0.0478 \\ -0.0700 \\ -0.0723 \\ -0.0663 \\ 0.0770 \\ 0.0742 \\ 0.0802 \\ 0.0744 \\ 0.0747 \\ 0.0770 \\ -0.0653 \\ -0.0303 \\ -0.0367 \\ 0.0289 \\ 0.0863 \\ 0.0183 \end{pmatrix} \begin{matrix} \text{Goiter} \\ \text{Adenoma} \\ \text{Cancer} \\ \text{Hashimoto's} \\ \text{Riedel's} \\ \text{Dequervain's} \\ \text{Other} \\ 1966 - 71 \\ 1972 - 76 \\ 1977 - 81 \\ 1982 - 86 \\ 1987 - 91 \\ 1992 - 94 \\ \text{Cancer} * 1966 - 71 \\ \text{Cancer} * 1972 - 76 \\ \text{Cancer} * 1977 - 81 \\ \text{Cancer} * 1982 - 86 \\ \text{Cancer} * 1987 - 91 \\ \text{Cancer} * 1992 - 94 \end{matrix}. \quad (31)$$

In the parameter vector, we can see coefficients for 7 outcomes (Goiter, Adenoma, Cancer, Hashimoto's thyroiditis, Riedel's thyroiditis, Dequervain's thyroiditis, and Other), 6 time periods (1966-71, 1972-76, 1977-81, 1982-86, 1987-91, and 1992-94), and 6 interaction terms for Cancer and Period (Cancer * 1966-71, Cancer * 1972-76, Cancer * 1977-81, Cancer * 1982-86, Cancer * 1987-91, Cancer * 1992-94). Using sum-to-zero constraints, we have a total of 19 coefficients. (Sum-to-zero constraints do not use of a constant term nor $p - 1$ levels to describe each factor. Rather, p terms are used for each factor without a constant). We notice that coefficients for the outcomes (coefficients 1-7), periods (coefficients 8-13), and cancer*period interaction (coefficients 14-19) are non-zero. The increasing proportion of cancer over time is also noticeable by the increasing trend in values of coefficients 14 through 18, but the trend did not continue for the last time period (coefficient 19).

For the purpose of this example, we shall focus on three hypothesis tests. First, we would like to know if there was a "period effect" present in the data, that is, if the apportionment of various thyroid outcomes changed over time. In (31) the period effect is expressed by the 6 coefficients 8-13. Thus, we are conducting a simultaneous test of the 6 coefficients with the hypothesis

$$H_o : b_8 = b_9 = b_{10} = b_{11} = b_{12} = b_{13} = 0. \quad (32)$$

Since there are 6 coefficients to test and 19 total coefficients, this results in a 6×19 \mathbf{C} matrix shown as

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (33)$$

The second test we will perform identifies whether the cancer*period interaction is significant. The interaction terms are coefficients 14-19, so the test notation is

$$H_o : b_{14} = b_{15} = b_{16} = b_{17} = b_{18} = b_{19} = 0. \quad (34)$$

The \mathbf{C} matrix now becomes

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (35)$$

Table 2: Format specification for input disk-file (filename specified by user at run-time).

Description	Form	Columns	Number of Records
Number of parameters	Integer	1-2	1
Variable names	Character	1-20	p
Regression coefficients	Real	1-16	p
Variance-covariance matrix	Real	$p(1-16)$	p

Our last test determines the presence of a statistically significant trend in the cancer*period interaction term. Trend tests using contrast matrices are performed by using equal “spacing” throughout the coefficients to be tested. Thus, we have the following notation

$$H_o : -3b_{14} - 2b_{15} - b_{16} + b_{17} + 2b_{18} + 3b_{19} = 0, \quad (36)$$

and the \mathbf{C} matrix

$$\mathbf{C} = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ -3 \ -2 \ -1 \ 1 \ 2 \ 3). \quad (37)$$

Results of the χ^2 tests using (16) and \mathbf{C} matrices above are 4575.67 (6 d.f.) for the period effect, 164.99 (6 d.f.) for the cancer*period interaction and 64.18 (1 d.f.) for the test of trend for the cancer*period interaction. These results indicate that there is a significant period effect, a significant cancer*period interaction, and a significant trend for the proportion of cancer to increase over time.

2.8 INPUT

There are seven forms of input into the computer program: the number of parameters, the variable names of the regression coefficients, the regression coefficients, the variance-covariance matrix of the regression coefficients, the number of coefficients used in RR or tests of hypotheses, the probability of the CIs, and the units applied to each coefficient. The first four items, i.e., the number of parameters, variable names, regression coefficients, and variance-covariance matrix can be entered from either the keyboard or a disk-file. The last three items, namely, the number of coefficients used in RR or tests of hypotheses, probability of the CIs, and units applied to each coefficient can only be entered by using the keyboard.

3 USER INPUT DATA FILE

This section discusses the data format when the user has specified that the regression coefficients and variance-covariance matrix are to be entered using the input disk-file. Table 2 lists the data in each record (card) of the input disk-file. When entering the regression coefficients and variance-covariance matrix with the keyboard, however, the data must follow the form described below.

The examples in the Appendix give listings of input files according to the format specified in Table 2.

4 SAMPLE INPUT/OUTPUT CASES

The output for all computer runs is in tabular form, giving the variable names, the coefficients, variance-covariance matrix, single or joint effects, contrast matrices, units, point estimates and lower and upper bounds of the $1-\alpha$ CIs for each run, the chi-square test statistic with degrees of freedom and p-values. Two examples using an input disk-file are given in the Appendix.

Table 3: Names and descriptions of each COVAR file.

Filename	Description	Format
INFILE (optional)*	Input file	ASCII text
OUTFILE*	Output file	ASCII text
COVAR.FOR	FORTRAN source code	ASCII text
COVAR.EXE	Executable file	FORTRAN

* Input (optional) and output filenames are specified by the user at run-time.

5 FILENAMES

Table 3 lists the names of the files that were used to program, link and execute COVAR. As one notices, the source code is an ASCII text file and the object and executable files have been compiled with Microsoft FORTRAN Powerstation Version 4.10. There is only one optional input data file, which is read when the user specifies data input from disk. In this case, the filename to be read is specified by the user at run-time.

6 AVAILABILITY

The program and all of its subroutines are available from the Journal of Statistical Software free of charge at <http://www.stat.ucla.edu/journals/jss/>

7 ACKNOWLEDGEMENTS

The author gratefully acknowledges Drs. Barry Davis, Ronald Firthofer and Wenyaw Chan of the University of Texas for their review and helpful commentary. Work on this algorithm was partially performed under NASA contract NAS-9 18236.

References

- [1] Firthofer, R.N. and Koch, G.G. (1974). A program for the analysis of compounded functions of categorical data. *Comp. Prog. Biomed.* 3:237-248.
- [2] Cox, D.R. (1972) Regression Models and Life Tables. *J. Royal Stat. Soc., Series B.* 34:187-220.
- [3] Engelman, L. (1991) Stepwise Logistic Regression, PLR Module. BMDP Statistical Software Manual. Berkeley: U. of California.
- [4] Frome, E. L. (1981) PREG: A computer program for Poisson regression analysis. Oak Ridge Associated Universities (ORAU). ORAU Publication 178. Oak Ridge: ORAU.
- [5] Preston, D.L. and Pierce, D.A. (1989). AMFIT: A program for parameter estimation in additive and multiplicative rate models with grouped survival data - methods, models, and examples. AMFIT User's Guide. Hirosoft International Corporation. Seattle: Hirosoft.
- [6] Abramowitz, M. and Stegun, A. (1965) Handbook of Mathematical Functions. New York: Dover.
- [7] Grizzle, J.E., Starmer, C.F., Koch, G.G. (1969) Analysis of categorical data for linear models. *Biometrics.* 25:489-504.
- [8] Heiberger, R. (1989) Computation for the analysis of designed experiments. New York: Wiley.
- [9] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1989) Numerical Recipes: The Art of Scientific Computing. New York: Cambridge University Press.

- [10] Schlesselman, J.J. (1982) Case-Control Studies: Design, Conduct, Analysis. Oxford: Oxford University.
- [11] Kalbfleisch, J.D., Prentice, R.L. (1980) The Statistical Analysis of Failure Time Data. New York: Wiley.
- [12] Hosmer, D.W., Lemeshow, S. (1989) Applied Logistic Regression. New York: Wiley.

8 APPENDIX: SAMPLE INPUT/OUTPUT CASES

Example 1: Data from logistic regression analysis. Odds-ratios and 95

Contents of the input file, EXAMPLE1.DAT, provided in the distribution are listed below:

```
5
LWD
SMOKE
AGE
LWD x SMOKE
LWD x AGE
1.730
1.153
0.084
1.407
0.147
3.490000
0.006260 0.210000
0.044700 0.008630 0.002080
0.120000 0.165000 0.000776 0.670000
0.147000 0.002840 0.002030 0.010000 0.006870
```

Keyboard specifications at run-time for Example 1.

INPUT FROM KEYBOARD OR FILE (K/F)? F

```
1 - 80.0
2 - 90.0
3 - 95.0
4 - 97.5
5 - 99.0
6 - OTHER
```

```
SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 2
SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 1
ENTER UNIT FOR EFFECT #1:
1
SPECIFY TEST FOR ROW #2: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 5
ENTER UNIT FOR EFFECT #2:
30
CONTINUE(Y/N)? Y
```

```
1 - 80.0
2 - 90.0
3 - 95.0
4 - 97.5
5 - 99.0
6 - OTHER
```

```
SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 3
SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 1
```

ENTER THE COEFFICIENT INDEX FOR ROW # 1: 1
 ENTER UNIT FOR EFFECT #1:
 1
 SPECIFY TEST FOR ROW #2: 1-RR 2-EFFECT 3-TREND: 1
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 4
 ENTER UNIT FOR EFFECT #2:
 1
 SPECIFY TEST FOR ROW #3: 1-RR 2-EFFECT 3-TREND: 1
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 5
 ENTER UNIT FOR EFFECT #3:
 30
 CONTINUE(Y/N)? Y

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
 ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 1
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 1
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 2
 ENTER UNIT FOR EFFECT #1:
 1
 CONTINUE(Y/N)? Y

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
 ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 2
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 1
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 2
 ENTER UNIT FOR EFFECT #1:
 1
 SPECIFY TEST FOR ROW #2: 1-RR 2-EFFECT 3-TREND: 1
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 4
 ENTER UNIT FOR EFFECT #2:
 1
 CONTINUE(Y/N)? Y

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

```

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 4
SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 1
ENTER UNIT FOR EFFECT #1:
1
SPECIFY TEST FOR ROW #2: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 2
ENTER UNIT FOR EFFECT #2:
1
SPECIFY TEST FOR ROW #2: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 4
ENTER UNIT FOR EFFECT #3:
1
SPECIFY TEST FOR ROW #3: 1-RR 2-EFFECT 3-TREND: 1
ENTER THE COEFFICIENT INDEX FOR ROW # 1: 5
ENTER UNIT FOR EFFECT #4:
30
CONTINUE(Y/N)? N

```

Output file for Example 1. Results are in agreement with those in Table 4.13 of [12].

RUN DATE: 4/10/1997

VARIANCE-COVARIANCE MATRIX OF COEFFICIENTS

	B1	B2	B3	B4	B5
B1	3.4900000				
B2	.62600000E-02	.21000000			
B3	.44700000E-01	-.86300000E-02	.20800000E-02		
B4	-.12000000	-.16500000	.77600000E-03	.67000000	
B5	-.14700000	.28400000E-02	-.20300000E-02	-.10000000E-01	.68700000E-02

INDEX	NAME	COEFFICIENT	STD. ERROR
1	- LWD	-1.73000000	1.86815400
2	- SMOKE	1.15300000	.45825760
3	- AGE	-.08400000	.04560702
4	- LWD x SMOKE	-1.40700000	.81853530
5	- LWD x AGE	.14700000	.08288547

EFFECTS, CIs AND TEST STATISTICS AT EACH CYCLE

CYCLE	B1	B2	B3	B4	B5	CI	RR	LOWER	UPPER	CHI2	DF	PROB
1	+	-	-	-	+	95.00	14.59	2.39	89.17	8.9594	2	.0113
2	+	-	-	+	+	95.00	3.57	.71	18.05	9.0557	3	.0286
3	-	+	-	-	-	95.00	3.17	1.29	7.78	6.3305	1	.0119
4	-	+	-	+	-	95.00	.78	.18	3.32	6.7952	2	.0335
5	+	+	-	+	+	95.00	11.31	2.08	61.51	13.5448	4	.0089

UNITS APPLIED TO COEFFICIENTS AT EACH CYCLE

CYCLE	U1	U2	U3	U4	U5
1	1.0	.0	.0	.0	30.0
2	1.0	.0	.0	1.0	30.0
3	.0	1.0	.0	.0	.0
4	.0	1.0	.0	1.0	.0
5	1.0	1.0	.0	1.0	30.0

Example 2: Data from multinomial categorical regression of data for thyroid pathologies among 2525 surgical patients who lived near the Semipalatinsk Test Site, Kazakstan, where nuclear weapons were tested from 1949-89. Test of hypotheses for effects of time period, cancer and time period interaction, and trend for increasing cancer over time periods. Regression coefficients obtained from use of a Grizzle-Starmer-Koch additive model for multinomially distributed data. All data stored in input data file EXAMPLE2.DAT.

Contents of input file EXAMPLE2.DAT provided in distribution are shown below:

```

19
Goiter
Adenoma
Cancer
Hashimoto's
Riedel's
Dequervain's
Other
1966-71
1972-76
1977-81
1982-86
1987-91
1992-94
Cancer * 1966-71
Cancer * 1972-76
Cancer * 1977-81
Cancer * 1982-86
Cancer * 1987-91
Cancer * 1992-94
.5557
.0618
.0013
.0478
-.0700
-.0723
-.0663
.0770
.0742
.0802
.0744
.0747
.0770
-.0653
-.0303
-.0367
.0289
.0863
.0183
.79073110E-04
-.95527380E-05 .40990960E-04

```

-.66085770E-05	-.36034480E-05	.18156760E-04		
-.92355610E-05	-.57279850E-05	-.33316720E-05	.37546310E-04	
-.60641810E-05	-.25531110E-05	-.63490720E-06	-.22354460E-05	.31471900E-05
-.59866910E-05	-.24739670E-05	-.57919900E-06	-.21561200E-05	.11089230E-05
-.61796430E-05	-.26697410E-05	-.72608150E-06	-.23522160E-05	.86166600E-06
.58893280E-05	.23836660E-05	.42398200E-06	.20666520E-05	-.10018460E-05
.59764610E-05	.24589970E-05	.49477690E-06	.21406180E-05	-.12317640E-05
.58217260E-05	.23477210E-05	.31618880E-06	.20342310E-05	-.66307480E-06
.59077310E-05	.23905490E-05	.53694550E-06	.20721830E-05	-.12833040E-05
.58233450E-05	.23068350E-05	.58408650E-06	.19885520E-05	-.13350450E-05
.60272670E-05	.25221220E-05	.31697150E-06	.22051640E-05	-.85483440E-06
.35120810E-06	.85173070E-06	-.14164310E-04	.89697650E-06	.12687060E-05
-.72239960E-06	-.21007690E-06	-.23973420E-05	-.16347970E-06	.51214890E-06
-.46918090E-06	-.32303120E-09	-.34006680E-05	.41403640E-07	.41954150E-07
-.15658960E-05	-.10538430E-05	.85072400E-05	-.10072680E-05	-.34852620E-06
-.15494840E-05	-.10381020E-05	.92756790E-05	-.99160810E-06	-.36476370E-06
-.26528570E-05	-.21528250E-05	.20336160E-04	-.21076640E-05	-.17444170E-05

Keyboard specifications at run-time for Example 2.

INPUT FROM KEYBOARD OR FILE (K/F)? F

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
 ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 6
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 1: 8
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 2: 9
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 3: 10
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 4: 11
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 5: 12
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 6: 13
 CONTINUE(Y/N)? Y

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
 ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 6
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2

ENTER THE COEFFICIENT INDEX FOR ROW # 1: 14
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 2: 15
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 3: 16
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 4: 17
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 5: 18
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 2
 ENTER THE COEFFICIENT INDEX FOR ROW # 6: 19
 CONTINUE(Y/N)? Y

1 - 80.0
 2 - 90.0
 3 - 95.0
 4 - 97.5
 5 - 99.0
 6 - OTHER

SELECT PERCENTAGE FOR CONFIDENCE INTERVALS: 2
 ENTER NUMBER OF ROWS IN CONTRAST MATRIX: 1
 SPECIFY TEST FOR ROW #1: 1-RR 2-EFFECT 3-TREND: 3
 ENTER BEGINNING COEFFICIENT INDEX: 14
 ENTER ENDING COEFFICIENT INDEX: 19
 CONTINUE(Y/N)? N

Output file for Example 2.

VARIANCE-COVARIANCE MATRIX OF COEFFICIENTS

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19
B1	.79073110E-04																		
B2	-.95527380E-05	.40990960E-04																	
B3	-.66085770E-05	-.36034480E-05	.18156760E-04																
B4	-.92355610E-05	-.57279850E-05	-.33316720E-05	.37546310E-04															
B5	-.60641810E-05	-.25531110E-05	-.63490720E-06	-.22354460E-05	.31471900E-05														
B6	-.59866910E-05	-.24739670E-05	-.57919900E-06	-.21561200E-05	.11089230E-05	.23669360E-05													
B7	-.61796430E-05	-.26697410E-05	-.72608150E-06	-.23522160E-05	.86166600E-06	.96027090E-06	.43740												
B8	.57360130E-05																		
B9	.61598620E-06	.39929060E-05																	
B10	-.20915850E-07	.25510090E-06	.79012830E-05																
B11	.65550000E-06	.99558190E-06	.29706330E-06	.35407700E-05															
B12	.69092170E-06	.10233200E-05	.33625410E-06	.10592090E-05	.30728940E-05														
B13	.16885510E-06	.46789670E-06	-.16786120E-06	.50748810E-06	.54302610E-06	.70760190E-05													
B14	-.57919500E-05	-.74271730E-06	.72774950E-07	-.82439780E-06	-.90696110E-06	-.11778330E-06	.48295												
B15	.10982830E-03																		
B16	-.14714190E-04	.10707580E-03																	
B17	-.26892360E-04	-.26489030E-04	.16784310E-03																

B18	-.27701040E-04	-.27286250E-04	-.39481890E-04	.17181930E-03	
B19	-.40216420E-04	-.39750260E-04	-.51993600E-04	-.52794410E-04	.23285270E-03

INDEX	NAME	COEFFICIENT	STD. ERROR
1	- Goiter	.55570000	.00889231
2	- Adenoma	.06180000	.00640242
3	- Cancer	.00130000	.00426108
4	- Hashimoto's	.04780000	.00612750
5	- Riedel's	-.07000000	.00177403
6	- Dequervain's	-.07230000	.00153848
7	- Other	-.06630000	.00209143
8	- 1966-71	.07700000	.00239500
9	- 1972-76	.07420000	.00199823
10	- 1977-81	.08020000	.00281092
11	- 1982-86	.07440000	.00188169
12	- 1987-91	.07470000	.00175297
13	- 1992-94	.07700000	.00266008
14	- Cancer * 1966-71	-.06530000	.00694946
15	- Cancer * 1972-76	-.03030000	.01047990
16	- Cancer * 1977-81	-.03670000	.01034774
17	- Cancer * 1982-86	.02890000	.01295543
18	- Cancer * 1987-91	.08630000	.01310799
19	- Cancer * 1992-94	.01830000	.01525951

EFFECTS, CIs AND TEST STATISTICS AT EACH CYCLE

CYCLE	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	CI	RR	LOWER	UPPER	CHI2
1	- EFFECT TEST, CONTRAST MATRIX BELOW																			N/A	N/A	N/A	N/A	4575.676
	CONTRAST MATRIX:																							
	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.	0.					
2	- EFFECT TEST, CONTRAST MATRIX BELOW																			N/A	N/A	N/A	N/A	164.997
	CONTRAST MATRIX:																							
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.	0.					
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	1.					
3	- TREND TEST, SEE CONTRAST MATRIX BELOW																			N/A	N/A	N/A	N/A	64.185
	CONTRAST MATRIX:																							
	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	-3.	-2.	-1.	1.	2.					3.

UNITS APPLIED TO COEFFICENTS AT EACH CYCLE

CYCLE	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15	U16	U17
-------	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----