

USING SAS/IML[®] SOFTWARE TO GENERATE ODDS RATIOS AND ASSOCIATED CONFIDENCE INTERVALS BASED ON LOGISTIC REGRESSION OUTPUT

Catherine C. Cowie, Social & Scientific Systems, Inc., Bethesda, MD
Danita Byrd-Holt, Social & Scientific Systems, Inc., Bethesda, MD
Keith F. Rust, Westat, Inc., Rockville, MD



ABSTRACT

Odds ratios and associated confidence intervals are frequently used measures in epidemiologic and biomedical research and are often derived from logistic regression models. Computation of these measures is straightforward when based on simple logistic regression models. Many logistic regression models, however, are complex and involve interactions of independent variables. In this case, computations of these measures may be quite tedious and subject to error if performed by hand. An automated method using the SAS[®] software is not currently available for use in these more complex regressions. We present a simple and quick computer program for calculating odds ratios and confidence intervals from logistic regression output using SAS/IML[®] software. The method may be used both for reporting results from a final reduced logistic regression model and for model fitting in which the magnitude of effect of a variable is a criterion for model inclusion. The method is also applicable for use based on output from other regression models.

INTRODUCTION

Logistic regression analysis is frequently used in epidemiologic and biomedical research to assess the magnitude of effect of a risk factor on an outcome. The specific measure obtained from this analysis, the odds ratio, is directly calculated from the coefficient(s) of the desired independent variable(s). A confidence interval around the odds ratio may also be calculated using coefficient variances and covariances to indicate the degree of confidence that the true population odds ratio is close to the estimated value.

Although the LOGISTIC procedure of SAS/STAT[®] software currently provides odds ratios with associated confidence intervals for the effects of independent variables, these measures only have meaning when the independent variables stand alone as main effect variables, such as if we wished to examine the odds ratio for diabetes in blacks compared to whites. However, because most chronic diseases are multifactorial in etiology, many logistic regression equations are complex. The complexity may be due to the inclusion of both linear and squared independent variables or to the inclusion of two- and three-way interactions of variables. An example of an effect for a two-way interaction would be the odds ratio for diabetes in blacks compared to whites at varying levels

of obesity. The resultant computation of odds ratios, and particularly of associated confidence intervals, is tedious and subject to considerable error when performed by hand. The development of a SAS program to calculate these measures appears beneficial not only for use with final logistic regression models, but also for use in model fitting to determine which variables should be included.

We describe a relatively simple and straightforward computer program for generating odds ratios and associated confidence intervals using SAS/IML software, that relies on outputted results from a logistic regression analysis. We present background on estimation of odds ratios and confidence intervals, matrix algebra development for their estimation, and use of the IML procedure to generate these measures, combined with specific examples.

ESTIMATION OF ODDS RATIOS AND ASSOCIATED CONFIDENCE INTERVALS

The general equation for the logistic regression model is:

$$\text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

where x_1, \dots, x_n denote the values of the independent variables, β_0, \dots, β_n denote the regression coefficients, p denotes the probability that the outcome of interest occurs, and

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

An estimate of the magnitude of effect of an independent variable on the outcome is calculated by the odds ratio (OR):

$$(1) \quad OR = \frac{e^{\beta_1 x_{11} + \beta_2 x_{21} + \dots + \beta_n x_{n1}}}{e^{\beta_1 x_{10} + \beta_2 x_{20} + \dots + \beta_n x_{n0}}}$$

where x_{11} denotes the value of x_1 for the numerator group, and x_{10} denotes the value of x_1 for the denominator group.

Expression (1) is algebraically equivalent to:

$$e^{\beta_1 (x_{11} - x_{10}) + \beta_2 (x_{21} - x_{20}) + \dots + \beta_n (x_{n1} - x_{n0})}$$

When a logistic regression model is fitted to a set of data, maximum likelihood estimates $\hat{\beta}_j$ of the

coefficients $\hat{\beta}_j$ are obtained, together with estimates of the variances of the $\hat{\beta}_j$, $\text{var}(\hat{\beta}_j)$, and all pairwise covariances, $\text{cov}(\hat{\beta}_j, \hat{\beta}_k)$. An estimate of the odds ratio is given by (Kleinbaum et al., 1982):

$$\widehat{OR} = e^{\left(\sum_{j=1}^n \hat{\beta}_j(x_{1j} - x_{0j})\right)}$$

A large sample confidence interval around the odds ratio estimate is obtained by (Kleinbaum et al., 1982):

$$(2) \quad e^{\log \widehat{OR} \pm z_{\alpha/2} \sqrt{\text{var}(\log \widehat{OR})}}$$

where $z_{\alpha/2}$ = standard normal deviate corresponding to a probability of $\alpha/2$ in each tail,

$$\log \widehat{OR} = \sum_{j=1}^n \hat{\beta}_j(x_{1j} - x_{0j})$$

and

$$\text{var}(\log \widehat{OR}) = \sum_{j=1}^n \text{var}(\hat{\beta}_j) \cdot (x_{1j} - x_{0j})^2 + \sum_{j \neq k}^n \text{cov}(\hat{\beta}_j, \hat{\beta}_k) \cdot (x_{1j} - x_{0j})(x_{1k} - x_{0k})$$

We illustrate both the calculation and interpretation of the OR and its associated confidence interval in the following example. Table 1 shows a final logistic regression model after fitting a series of models to assess the effect of race and other risk factors on risk of diabetes in the U.S. using data from a national health survey. The relationship of the independent variables with the outcome variable of diabetes is complex. In this model, only the variables of education and family history of diabetes are related to diabetes solely as main effects. All other variables are related to diabetes as two-way interactions, through the modifying effects of other variables. As will be seen below, the calculation of odds ratios and associated confidence intervals is more complicated in this case.

Suppose that we wish to examine the effect of having a family history of diabetes on the odds of having diabetes. The OR would be estimated as:

$$\widehat{OR} = e^{\hat{\beta}_1(x_{11} - x_{01})}$$

where: x_{1j} = family history of diabetes in group j : 1 if yes, 0 if no.

Substituting in actual values, the expression becomes:

$$\widehat{OR} = e^{(0.9795)(1-0)} = 2.7$$

According to this estimate, the odds of diabetes is 2.7-fold greater in those with a family history of diabetes compared to those without a family history.

Table 1. Logistic Regression Model Assessing Race and Other Risk Factors for Diabetes in the U.S.

| | Coefficient | p-value |
|-------------------------------------|-------------|---------|
| Race (black/white) | 0.4015 | * |
| Sex (women/men) | 0.4706 | * |
| Age† | 1.0264 | * |
| Family history of diabetes (yes/no) | 0.9795 | <0.001 |
| PDW‡ | 0.3334 | * |
| PDW, squared | 0.2025 | * |
| Education level§ | -0.1933 | 0.04 |
| Subscapular/triceps ratio | 0.1689 | * |
| Race x PDW | 0.2627 | * |
| Race x PDW, squared | -0.1215 | 0.05 |
| Sex x PDW | 0.2731 | * |
| Sex x PDW, squared | -0.2310 | 0.03 |
| Age x PDW | 0.1157 | * |
| Age x PDW, squared | -0.0580 | 0.03 |
| Sex x subscapular/triceps ratio | 0.3997 | 0.04 |

* p value not specified because effect of variable depends on other covariates

† Age (in yrs) standardized by subtracting mean (42.71) and dividing resultant quantity by standard deviation (15.27)

‡ PDW (percent desirable weight) standardized by subtracting mean (124.07) and dividing resultant quantity by standard deviation (23.42)

§ Education levels: 12th/9-12th/≤ 8th grades

|| Subscapular/triceps ratio standardized by subtracting mean (1.15) and dividing resultant quantity by standard deviation (0.56)

Reprinted with permission of *The American Journal of Epidemiology*. Cowie, CC, Harris, MI, Silverman, RE et al. (1993), Effect of multiple risk factors on differences between blacks and whites in the prevalence of non-insulin-dependent diabetes mellitus in the United States, 137(7). Copyright © 1993 by The Johns Hopkins University School of Hygiene and Public Health.

The variance of the estimated coefficient for family history of diabetes, $\hat{\beta}_1$, is 0.015172. Since family history is related to diabetes as a main effect only and no interactions are involved, a 95 percent confidence interval for the odds ratio is calculated by:

$$e^{(0.9795) \pm 1.96 \sqrt{0.015172}} = (2.1, 3.4)$$

since: $\text{var}(\log \widehat{OR}) = (1 - 0)^2 (0.015172) = 0.015172$

The calculation of odds ratios and associated confidence intervals based on interactions of variables is more complicated. Suppose that we wish to examine the odds of diabetes in blacks compared to whites. Since the effect of race varies with level of obesity (measured as percent desirable weight or PDW), the risk in blacks relative to whites can only be expressed at specific levels of PDW. All terms involving race and

race by PDW interactions are needed in the estimation of the OR. The estimation of the OR at a given PDW level would be:

$$\widehat{OR} = e^{\hat{\beta}_1(x_{11} - x_{10}) + \hat{\beta}_2(x_{21} - x_{20}) + \hat{\beta}_3(x_{31} - x_{30})}$$

where x_{ij} = race: 1 if black, 0 if white; x_{2j} = race x PDW = $x_{1j}x_{4j}$: race (x_{1j}) is coded as above and PDW (x_{4j}) is value of PDW standardized according to Table 1; and $x_{3j} = x_{1j}x_{4j}^2$.

For a PDW level of 150, substituting in actual variable values and parameter estimates, the expression becomes:

$$\begin{aligned} \widehat{OR} &= e^{(0.4015)(1-0) + (0.2627)\left((1)\left(\frac{150-124.07}{23.42}\right) - (0)\left(\frac{150-124.07}{23.42}\right)\right)} \\ &\quad + (-0.1215)\left((1)\left(\frac{150-124.07}{23.42}\right)^2 - (0)\left(\frac{150-124.07}{23.42}\right)^2\right)} \\ &= e^{0.54336} = 1.7 \end{aligned}$$

According to this result, the odds of diabetes is seventy percent greater in blacks than whites at a PDW of 150.

Referring to Table 2 for coefficient variances and covariances, a 95 percent confidence interval would be calculated by the following:

$$e^{0.54336 \pm 1.96\sqrt{0.06126}} = (1.1, 2.8)$$

since var (log OR) is obtained as:

$$\begin{aligned} &(0.078558)(1-0)^2 + (0.064489)\left((1)\left(\frac{150-124.07}{23.42}\right) - (0)\left(\frac{150-124.07}{23.42}\right)\right)^2 \\ &+ (0.003644)\left((1)\left(\frac{150-124.07}{23.42}\right)^2 - (0)\left(\frac{150-124.07}{23.42}\right)^2\right)^2 \\ &+ 2\left[(-0.032814)(1-0)\left((1)\left(\frac{150-124.07}{23.42}\right) - (0)\left(\frac{150-124.07}{23.42}\right)\right)\right. \\ &+ (0.002523)(1-0)\left((1)\left(\frac{150-124.07}{23.42}\right)^2 - (0)\left(\frac{150-124.07}{23.42}\right)^2\right) \\ &+ (-0.012285)\left((1)\left(\frac{150-124.07}{23.42}\right) - (0)\left(\frac{150-124.07}{23.42}\right)\right) \\ &\left.\left((1)\left(\frac{150-124.07}{23.42}\right)^2 - (0)\left(\frac{150-124.07}{23.42}\right)^2\right)\right] = 0.06126 \end{aligned}$$

Table 2. Variance-covariance matrix for selected coefficients of final logistic regression model.

| | $\hat{\beta}_1$ (Race) | $\hat{\beta}_2$ (Race x PDW) | $\hat{\beta}_3$ (Race x PDW ²) |
|--|------------------------|------------------------------|--|
| $\hat{\beta}_1$ (Race) | 0.076558 | | |
| $\hat{\beta}_2$ (Race x PDW) | -0.032814 | 0.064489 | |
| $\hat{\beta}_3$ (Race x PDW ²) | 0.002523 | -0.012285 | 0.003644 |

The tediousness of the above calculation is apparent. For this model, calculations of odds ratios involving comparison of the odds of diabetes at different levels of PDW would require more terms in the expressions (1) and (2). All terms involving PDW would be needed, and a level of race, sex, and age would have to be specified. For example, if we were to compare the effect of PDW

at a level of 150 vs. 100 for black women at age 50, eight terms would be needed for the calculation of the odds ratio and 29 variances and covariances of coefficients would be required for estimating the variance of log \widehat{OR} .

An alternative representation of these formulae to calculate these measures relies on matrix algebra. It uses the coefficients, variances, and covariances from a logistic regression model, and the differences between the independent (X) variables that the analyst wishes to compare, as input to PROC IML. A large number of odds ratios and associated confidence intervals can be calculated easily, quickly, and without error.

MATRIX ALGEBRA DEVELOPMENT

We previously saw that

$$\log \widehat{OR} = \sum_{i=1}^n \hat{\beta}_i(x_{i1} - x_{i0})$$

where x_{i1} denotes the value of x_i of the numerator group and x_{i0} denotes the value of x_i of the denominator group.

If we let \underline{X} be the vector of X differences:

$$\underline{X} = \begin{bmatrix} x_{11} - x_{10} \\ x_{21} - x_{20} \\ \vdots \\ x_{n1} - x_{n0} \end{bmatrix}$$

And $\underline{\beta}$ be the vector of estimated coefficients:

$$\underline{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{bmatrix}$$

Then:

$$(3) \log \widehat{OR} = \underline{\beta}' \underline{X}$$

$$\text{and } (4) \widehat{OR} = e^{\log \widehat{OR}}$$

We also previously saw that:

$$\begin{aligned} \text{var}(\log \widehat{OR}) &= \sum_{i=1}^n \text{var}(\hat{\beta}_i) \cdot (x_{i1} - x_{i0})^2 + \\ &\quad \sum_{i \neq j}^n \sum_{j=1}^n \text{cov}(\hat{\beta}_i, \hat{\beta}_j) \cdot (x_{i1} - x_{i0})(x_{j1} - x_{j0}) \end{aligned}$$

If we let V be the variance-covariance matrix of the coefficients:

$$V = \begin{bmatrix} \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_1, \hat{\beta}_n) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{var}(\hat{\beta}_2) & & \\ \vdots & & \ddots & \\ \text{cov}(\hat{\beta}_n, \hat{\beta}_1) & & & \text{var}(\hat{\beta}_n) \end{bmatrix}$$

Then:

$$(5) \text{var}(\log \overline{OR}) = X' V X$$

This result may be incorporated in the formula we saw earlier for the confidence interval around the odds ratio:

$$e^{\log \overline{OR} \pm z_{\alpha/2} \sqrt{\text{var}(\log \overline{OR})}}$$

USE OF THE IML PROCEDURE

To calculate these measures using the IML procedure, the vector of coefficients, $\hat{\beta}$, and the variance-covariance matrix of the coefficients, V , are obtained from output generated from a logistic regression program. The analyst specifies the vector of X differences for the groups he wishes to compare for risk of the outcome. For a term not involved in the calculation of the particular odds ratio of interest, the value of the difference is set to zero. The program may be run interactively or in batch mode.

We provide an example to illustrate the use of our computer program to generate odds ratios and confidence intervals. The example was alluded to earlier for comparing the effect of PDW on risk of diabetes at a level of 150 vs. 100 for black women at age 50, based on the model in Table 1. Estimation of the odds ratio and confidence interval is complex since so many terms and coefficient variances and covariances are needed. The logistic regression coefficients (Table 1) and full variance-covariance matrix for the model were outputted to a SAS data file for access by PROC IML as shown below.

Appendix provides the code and output for our example using the IML procedure. Programming code is referenced by line numbers appearing to the left of the code. Lines 6 through 12 access the SAS file created by the logistic regression program and select observations to be outputted to two separate files, one containing coefficients entitled "Betas" and the other containing the variance-covariance matrix entitled "Varcov". Lines 21 through 22 invoke the IML procedure and define processing options. In lines 26 through 36, the vector of coefficients is read from the SAS file "Betas" and printed, and in lines 38 through 53

the variance-covariance matrix is read from "Varcov" and printed. Row and column headings and formats for the printed numeric output are also assigned in these lines through the use of the MATTRIB statement. In lines 55 through 81, the vector of X differences and its transpose are specified, as well as row and column headings and labels. Computations are found in lines 83 through 115. Specification of expression 3 is found in line 86, expression 4 in line 98, expression 5 in line 93, and expression 2 in lines 109 and 114. The output contains a print of all matrices for checking purposes and is followed by computation results. When computing additional odds ratios and confidence intervals for a specific regression model, all the investigator need do is change the definition of the X vector (lines 65-68). This can be done by defining lines 55-115 as a SAS MACRO and defining the x vector (XP) as a macro variable. For help in model fitting to determine which variables should be included, the analyst may decide to program more informally with fewer names and labels.

CONCLUSION

The IML procedure is useful for calculating odds ratios and associated confidence intervals based on coefficients and a variance-covariance matrix from a logistic regression model. PROC IML is particularly efficient when logistic regression models are complex and include many independent variables, both linear and squared independent variables, or two- and three-way interactions of variables. We have also used this method, with minor modification, to compute mean estimates from linear regression and relative death rates from Cox proportional hazards regression.

REFERENCES

- Kleinbaum, DG, Kupper, LL, Morganstern, H (1982), *Epidemiologic research. Principles and quantitative methods*, Belmont, CA: Wadsworth Inc, pp. 422,432-433.
- SAS Institute Inc. (1990), *SAS/IML Software: Usage and Reference, Version 6, First Edition*, Cary, NC: SAS Institute Inc.
- SAS Institute Inc. (1989), *SAS/STAT User's Guide, Version 6, Fourth Edition*, Cary, NC: SAS Institute Inc.

SAS, SAS/IML, and SAS/STAT are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Appendix 1. Code and output for example using PROC IML

```

1. *****;
2. * GENERATE ODDS RATIOS AND *;
3. *CONFIDENCE INTERVALS FOR OR*;
4. *****;
5. *****;
6. DATA BETAS VARCOV;
7.   SET IN.SAMPLE01;
8.
9.   DROP IDNUM--B001 DENOMDF;
10.
11.   IF _N_ = 1 THEN OUTPUT BETAS;
12.   ELSE IF 2 LE IDNUM LE NCELL THEN OUTPUT VARCOV;
13.
14.   TITLE1 'COMPARING MAXIMUM PDW 150 VS. 100';
15.   TITLE2 'IN BLACK WOMEN AGE 50';
16.   TITLE3 ' ';
17.   TITLE4 ' ';
18.   TITLE5 ' ';
19.   TITLE6 'ROW VECTOR OF BETAS';
20.
21. PROC IML ;
22. RESET FUZZ LINESIZE=172 PAGESIZE=46;
23.
24.
25.
26. ** READ IN BETA MATRIX **;
27. USE BETAS;
28. READ ALL INTO BP;
29. TITLE6 'ROW VECTOR OF BETAS';
30. MATTRIB BP ROWNAME=EMPTY
31.   COLNAME = ({RACE SEX AGE FAMDIAB
32.   PDW PDWSQ EDUCAT STRATIO
33.   RACEPDW RCEPDWSQ PDWSEX
34.   PDWAGE PDWSQAGE PDWSQSEX
35.   SEXSTRAT});
36. PRINT BP;
37.
38. ** READ IN VARIANCE/COVARIANCE MATRIX **;
39. USE VARCOV;
40. READ ALL INTO V;
41. TITLE6 'VARIANCE/COVARIANCE MATRIX';
42. MATTRIB V ROWNAME = ({RACE SEX AGE FAMDIAB
43.   PDW PDWSQ EDUCAT STRATIO
44.   RACEPDW RCEPDWSQ PDWSEX
45.   PDWAGE PDWSQAGE PDWSQSEX
46.   SEXSTRAT})
47.   COLNAME = ({RACE SEX AGE FAMDIAB
48.   PDW PDWSQ EDUCAT STRATIO
49.   RACEPDW RCEPDWSQ PDWSEX
50.   PDWAGE PDWSQAGE PDWSQSEX
51.   SEXSTRAT})
52.   FORMAT=9.7;
53. PRINT V;
54.
55. *****;
56. * ENTER X VALUES BELOW *;
57. *****;
58. TITLE6 'ROW VECTOR OF X DIFFERENCES';
59. MATTRIB XP ROWNAME=EMPTY
60.   COLNAME = ({RACE SEX AGE FAMDIAB
61.   PDW PDWSQ EDUCAT STRATIO
62.   RACEPDW RCEPDWSQ PDWSEX
63.   PDWAGE PDWSQAGE PDWSQSEX
64.   SEXSTRAT}) ;
65. XP = {0 0 0 0 2.13492741246798 .169554440102069
66.   0 0 2.134927412467989 .169554440102069
67.   2.13492741246798 1.019228607524
68.   .080946422288414 .169554440102069 0};
69. PRINT /XP;
70.
71.
72. TITLE6 ' ';

```

```

73. MATTRIB X COLNAME=EMPTY
74.   ROWNAME = ({RACE SEX AGE FAMDIAB
75.   PDW PDWSQ EDUCAT STRATIO
76.   RACEPDW RCEPDWSQ PDWSEX
77.   PDWAGE PDWSQAGE PDWSQSEX
78.   SEXSTRAT})
79.   LABEL = {'COLUMN VECTOR OF X DIFFERENCES'};
80. X = (XP);
81. PRINT X;
82.
83. *** COMPUTE LOG OF ODDS RATIO ***;
84. MATTRIB LOGOR
85.   ROWNAME = {'LOG ODDS RATIO'};
86. LOGOR = BP*X;
87. TITLE6 ' ';
88. PRINT LOGOR;
89.
90. *** COMPUTE VARIANCE OF ODDS RATIO LOG ***;
91. MATTRIB VARLOGOR
92.   ROWNAME = {'VARIANCE LOG ODDS RATIO'};
93. VARLOGOR = XP*V*X;
94. PRINT VARLOGOR;
95.
96. *** COMPUTE ODDS RATIO ***;
97. MATTRIB OR ROWNAME = {'ODDS RATIO'};
98. OR = EXP(LOGOR);
99. PRINT OR;
100.
101. *** COMPUTE 95% CONFIDENCE INTERVALS ***;
102. MATTRIB SQVLOGOR
103.   ROWNAME = {'SQUARE ROOT OF VARIANCE LOG
104.   ODDS RATIO'};
105. SQVLOGOR = SQRT(ABS(VARLOGOR));
106. PRINT SQVLOGOR;
107.
108. MATTRIB UP95
109.   ROWNAME = {'UPPER 95 % CONFIDENCE LIMIT'};
110. UP95 = EXP(LOGOR + (1.96 * SQVLOGOR));
111. PRINT UP95;
112.
113. MATTRIB LOW95
114.   ROWNAME = {'LOWER 95 % CONFIDENCE LIMIT'};
115. LOW95 = EXP(LOGOR - (1.96 * SQVLOGOR));
116. PRINT LOW95;

```

COMPARING MAXIMUM PDW 150 VS. 100
IN BLACK WOMEN AGE 50

ROW VECTOR OF BETAS

| BP | RACE | SEX | AGE | FAMDIAB | PDW | PDWSQ | EDUCAT | STRATIO | RACEPDW | RCEPDWSQ | PDWSEX | PDWAGE | PDWSQAGE | PDWSQSEX | SEXSTRAT |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|-----------|-----------|-----------|-----------|----------|
| 0.401473 | 0.4705594 | 1.0263929 | 0.9794972 | 0.3334383 | 0.2024862 | -0.193331 | 0.1688836 | 0.2626923 | -0.121516 | 0.273118 | 0.1156945 | -0.057983 | -0.231006 | 0.3997185 | |

VARIANCE/COVARIANCE MATRIX

| V | RACE | SEX | AGE | FAMDIAB | PDW | PDWSQ | EDUCAT | STRATIO | RACEPDW | RCEPDWSQ | PDWSEX | PDWAGE | PDWSQAGE | PDWSQSEX | SEXSTRAT |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| RACE | 0.0765578 | -.0151781 | 0.0084871 | 0.0132415 | 0.0167958 | -.0122462 | 0.0021566 | -.0010868 | -.0328136 | 0.0025226 | -.0116512 | -.0017981 | 0.0000311 | 0.0117977 | -.0057562 |
| SEX | -.0151781 | 0.0303453 | -.0012538 | 0.0004509 | -.0190292 | 0.0123604 | -.0031401 | 0.0021472 | 0.0124008 | -.0016489 | 0.0053102 | 0.0066644 | -.0018300 | -.0088726 | 0.0119179 |
| AGE | 0.0084871 | -.0012538 | 0.0067409 | 0.0052682 | 0.0071047 | -.0015155 | 0.0022432 | 0.0003508 | 0.0018824 | -.0003962 | -.0053698 | -.0025845 | 0.0000836 | 0.0014970 | 0.0012072 |
| FAMDIAB | 0.0132415 | 0.0004509 | 0.0052682 | 0.0151716 | 0.0074637 | -.0008018 | 0.0021205 | -.0002196 | -.0089062 | 0.0009806 | -.0030324 | -.0026558 | 0.0003745 | 0.0000507 | 0.0060093 |
| PDW | 0.0167958 | -.0190292 | 0.0071047 | 0.0074637 | 0.0437124 | -.0159817 | -.0005651 | -.0042952 | -.0207971 | 0.0030061 | -.0235358 | -.0131389 | 0.0023721 | 0.0127624 | 0.0034759 |
| PDWSQ | -.0122462 | 0.0123604 | -.0015155 | -.0008018 | -.0159817 | 0.0119388 | -.0002959 | 0.0016699 | 0.0108248 | -.0018100 | 0.0092477 | 0.0039957 | -.0010907 | -.0105339 | 0.0009540 |
| EDUCAT | 0.0021566 | -.0031401 | 0.0022432 | 0.0021205 | -.0005651 | -.0002959 | 0.0085407 | 0.0003122 | 0.0032437 | 0.0003928 | 0.0044106 | 0.0004253 | -.0001018 | -.0010369 | -.0000456 |
| STRATIO | -.0010868 | 0.0021472 | 0.0003508 | -.0002196 | -.0042952 | 0.0016699 | 0.0003122 | 0.0037019 | 0.0034591 | -.0006704 | 0.0033812 | -.0006041 | 0.0002106 | -.0017292 | -.0048961 |
| RACEPDW | -.0328136 | 0.0124008 | 0.0018824 | -.0089062 | -.0207971 | 0.0108248 | 0.0032437 | 0.0034591 | 0.0644891 | -.0122853 | 0.0057000 | 0.0049043 | -.0011600 | -.0082168 | -.0019032 |
| RCEPDWSQ | 0.0025226 | -.0016489 | -.0003962 | 0.0009806 | 0.0030061 | -.0018100 | 0.0003928 | -.0006704 | -.0122853 | 0.0036436 | 0.0003042 | -.0006237 | 0.0002084 | 0.0010377 | 0.0018320 |
| PDWSEX | -.0116512 | 0.0053102 | -.0053698 | -.0030324 | -.0235358 | 0.0092477 | 0.0044106 | 0.0033812 | 0.0057000 | 0.0003042 | 0.0297740 | 0.0026452 | 0.0000630 | -.0119243 | -.0038241 |
| PDWAGE | -.0017981 | 0.0066644 | -.0025845 | -.0026558 | -.0131389 | 0.0039957 | 0.0004253 | -.0006041 | 0.0049043 | -.0006237 | 0.0026452 | 0.0105309 | -.0019894 | -.0020136 | 0.0035348 |
| PDWSQAGE | 0.0000311 | -.0018300 | 0.0000836 | 0.0003745 | 0.0023721 | -.0010907 | -.0001018 | 0.0002106 | -.0011600 | 0.0002084 | 0.0000630 | -.0019894 | 0.0006468 | 0.0004031 | -.0013250 |
| PDWSQSEX | 0.0117977 | -.0088726 | 0.0014970 | 0.0000507 | 0.0127624 | -.0105339 | -.0010369 | -.0017292 | -.0082168 | 0.0010377 | -.0119243 | -.0020136 | 0.0004031 | 0.0107281 | 0.0002702 |
| SEXSTRAT | -.0057562 | 0.0119179 | 0.0012072 | 0.0060093 | 0.0034759 | 0.0009540 | -.0000456 | -.0048961 | -.0019032 | 0.0018320 | -.0038241 | 0.0035348 | -.0013250 | 0.0002702 | 0.0364345 |

RDW VECTOR OF X DIFFERENCES

| XP | RACE | SEX | AGE | FAMDIAB | PDW | PDWSQ | EDUCAT | STRATIO | RACEPDW | RCEPDWSQ | PDWSEX | PDWAGE | PDWSQAGE | PDWSQSEX | SEXSTRAT |
|----|------|-----|-----|---------|-----------|-----------|--------|---------|-----------|-----------|-----------|-----------|-----------|-----------|----------|
| 0 | 0 | 0 | 0 | 0 | 2.1349274 | 0.1695544 | 0 | 0 | 2.1349274 | 0.1695544 | 2.1349274 | 1.0192286 | 0.0809464 | 0.1695544 | 0 |

COLUMN VECTOR OF X DIFFERENCES

| | |
|----------|-----------|
| RACE | 0 |
| SEX | 0 |
| AGE | 0 |
| FAMDIAB | 0 |
| PDW | 2.1349274 |
| PDWSQ | 0.1695544 |
| EDUCAT | 0 |
| STRATIO | 0 |
| RACEPDW | 2.1349274 |
| RCEPDWSQ | 0.1695544 |
| PDWSEX | 2.1349274 |
| PDWAGE | 1.0192286 |
| PDWSQAGE | 0.0809464 |
| PDWSQSEX | 0.1695544 |
| SEXSTRAT | 0 |

LOGOR
LOG ODDS RATIO 1.9435692

VARLOGOR
VARIANCE LOG ODDS RATIO 0.2551266

OR
ODDS RATIO 6.9836323

SQVLOGOR
SQUARE ROOT OF VARIANCE LOG ODDS 0.5051005

UP95
UPPER 95 % CONFIDENCE LIMIT 18.794537

LOW95
LOWER 95 % CONFIDENCE LIMIT 2.5949626