# Logistic Regression for Predicting Diabetes

## The LOGISTIC Procedure

| Model Information | |
|---|---|
| Data Set | WORK.ANALYSIS_DATA |
| Response Variable | Diabetes_bin |
| Number of Response Levels | 2 |
| Model | binary logit |
| Optimization Technique | Fisher's scoring |

| | |
|---|---|
| Number of Observations Read | 2198 |
| Number of Observations Used | 2198 |

| Response Profile | | |
|---|---|---|
| Ordered Value | Diabetes_bin | Total Frequency |
| 1 | 0 | 2185 |
| 2 | 1 | 13 |

Probability modeled is Diabetes_bin=1.

| Model Convergence Status |
|---|
| Convergence criterion (GCONV=1E-8) satisfied. |

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 161.312 | 130.527 |
| SC | 167.007 | 153.308 |
| -2 Log L | 159.312 | 122.527 |

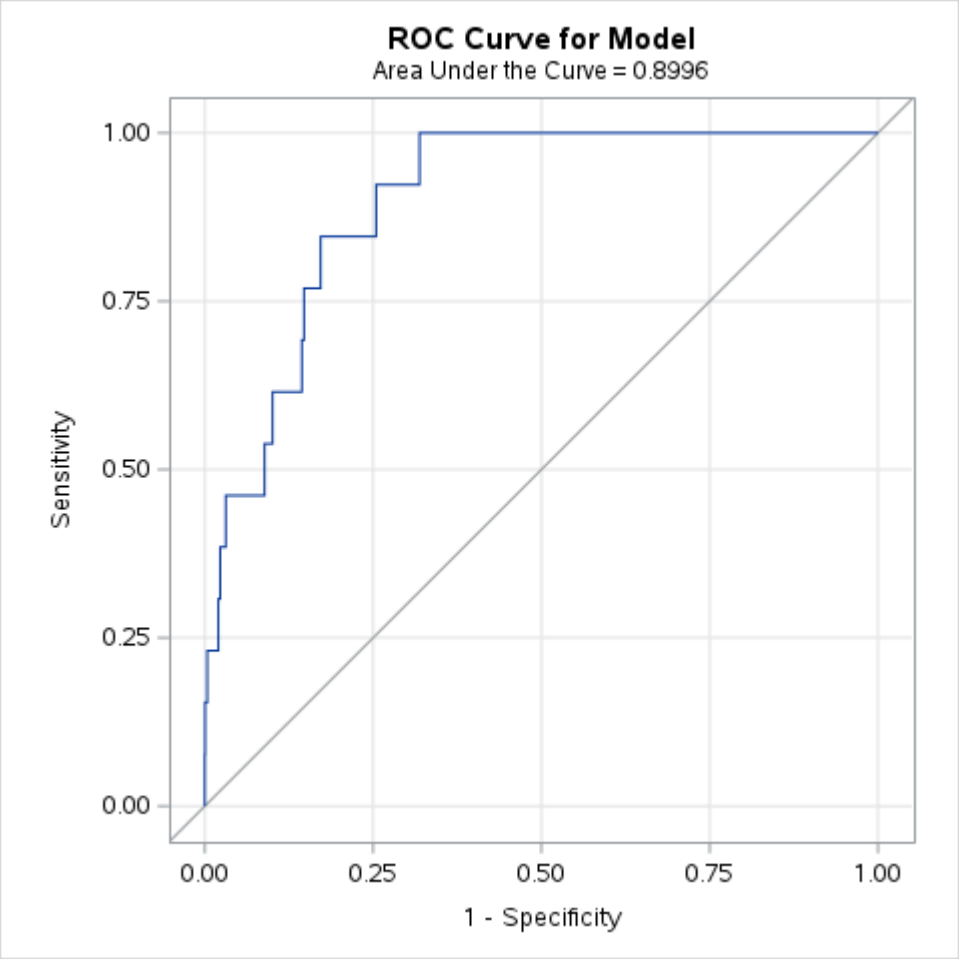
| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 36.7856 | 3 | <.0001 |
| Score | 48.2870 | 3 | <.0001 |
| Wald | 37.7487 | 3 | <.0001 |

| Analysis of Maximum Likelihood Estimates | | | | | |
|---|---|---|---|---|---|
| Parameter | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept | 1 | -6.7411 | 1.4162 | 22.6584 | <.0001 |
| Log_Insulin | 1 | 2.0611 | 0.3831 | 28.9516 | <.0001 |
| Log_HDL | 1 | -0.8753 | 1.1254 | 0.6049 | 0.4367 |
| Log_LDL | 1 | -2.7164 | 0.9561 | 8.0724 | 0.0045 |

| Odds Ratio Estimates | | | |
|---|---|---|---|
| Effect | Point Estimate | 95% Wald Confidence Limits | |
| Log_Insulin | 7.855 | 3.707 | 16.641 |
| Log_HDL | 0.417 | 0.046 | 3.783 |
| Log_LDL | 0.066 | 0.010 | 0.431 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 90.0 | Somers' D | 0.799 |
| Percent Discordant | 10.0 | Gamma | 0.799 |
| Percent Tied | 0.0 | Tau-a | 0.009 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Pairs | 28405 | c | 0.900 |




| Partition for the Hosmer and Lemeshow Test | | | | | |
|---|---|---|---|---|---|
| | | Diabetes_bin = 1 | | Diabetes_bin = 0 | |
| Group | Total | Observed | Expected | Observed | Expected |
| 1 | 221 | 0 | 0.05 | 221 | 220.95 |
| 2 | 221 | 0 | 0.12 | 221 | 220.88 |
| 3 | 222 | 0 | 0.19 | 222 | 221.81 |
| 4 | 220 | 0 | 0.27 | 220 | 219.73 |
| 5 | 221 | 0 | 0.37 | 221 | 220.63 |
| 6 | 220 | 0 | 0.49 | 220 | 219.51 |
| 7 | 220 | 1 | 0.67 | 219 | 219.33 |
| 8 | 220 | 1 | 1.00 | 219 | 219.00 |
| 9 | 220 | 4 | 1.75 | 216 | 218.25 |
| 10 | 213 | 7 | 8.09 | 206 | 204.91 |

| Hosmer and Lemeshow Goodness-of-Fit Test | | |
|---|---|---|
| Chi-Square | DF | Pr > ChiSq |
| 4.7034 | 8 | 0.7888 |

## ROC Curve for Diabetes Prediction Model

**The LOGISTIC Procedure**

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 90.0 | Somers' D | 0.799 |
| Percent Discordant | 10.0 | Gamma | 0.799 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Tied | 0.0 | Tau-a | 0.009 |
| Pairs | 28405 | c | 0.900 |

## Logistic Regression for Predicting Metabolic Syndrome HD

### The LOGISTIC Procedure

| Model Information | |
|---|---|
| Data Set | WORK.ANALYSIS_DATA |
| Response Variable | Mets_HD_bin |
| Number of Response Levels | 2 |
| Model | binary logit |
| Optimization Technique | Fisher's scoring |

| | |
|---|---|
| Number of Observations Read | 2198 |
| Number of Observations Used | 2198 |

| Response Profile | | |
|---|---|---|
| Ordered Value | Mets_HD_bin | Total Frequency |
| 1 | 0 | 1928 |
| 2 | 1 | 270 |

**Probability modeled is Mets_HD_bin=1.**

| Model Convergence Status |
|---|
| Convergence criterion (GCONV=1E-8) satisfied. |

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 1639.701 | 1030.044 |
| SC | 1645.396 | 1052.825 |
| -2 Log L | 1637.701 | 1022.044 |

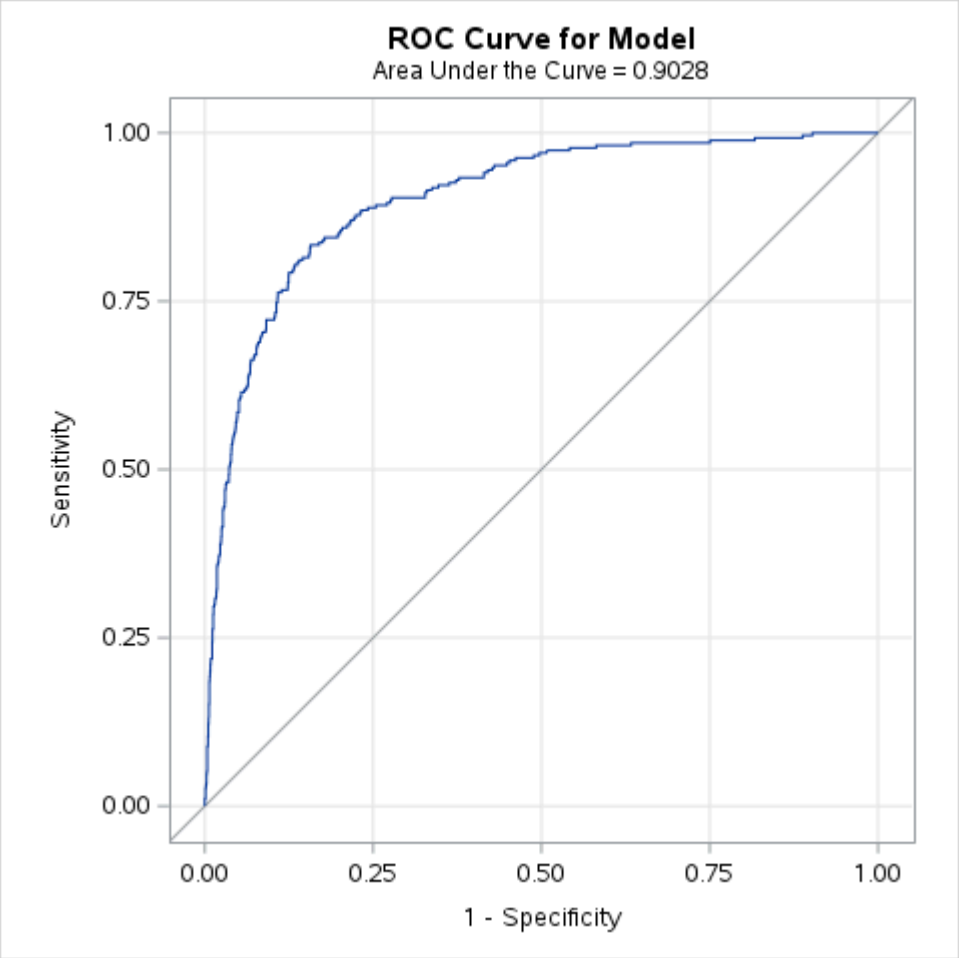
| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 615.6570 | 3 | <.0001 |
| Score | 563.3535 | 3 | <.0001 |
| Wald | 332.6732 | 3 | <.0001 |

| Analysis of Maximum Likelihood Estimates | | | | | |
|---|---|---|---|---|---|
| Parameter | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept | 1 | -8.5434 | 0.5786 | 217.9899 | <.0001 |
| Log_Insulin | 1 | 2.4829 | 0.1710 | 210.8547 | <.0001 |
| Log_HDL | 1 | -4.6947 | 0.3812 | 151.6709 | <.0001 |
| Log_LDL | 1 | 1.5406 | 0.3227 | 22.7983 | <.0001 |

| Odds Ratio Estimates | | | |
|---|---|---|---|
| Effect | Point Estimate | 95% Wald Confidence Limits | |
| Log_Insulin | 11.976 | 8.566 | 16.744 |
| Log_HDL | 0.009 | 0.004 | 0.019 |

| Odds Ratio Estimates | | | |
|---|---|---|---|
| Effect | Point Estimate | 95% Wald Confidence Limits | |
| Log_LDL | 4.667 | 2.480 | 8.784 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 90.3 | Somers' D | 0.806 |
| Percent Discordant | 9.7 | Gamma | 0.806 |
| Percent Tied | 0.0 | Tau-a | 0.174 |
| Pairs | 520560 | c | 0.903 |




| Partition for the Hosmer and Lemeshow Test | | | | | |
|---|---|---|---|---|---|
| | | Mets_HD_bin = 1 | | Mets_HD_bin = 0 | |
| Group | Total | Observed | Expected | Observed | Expected |
| 1 | 220 | 2 | 0.37 | 218 | 219.63 |
| 2 | 220 | 1 | 1.14 | 219 | 218.86 |
| 3 | 220 | 1 | 2.20 | 219 | 217.80 |
| 4 | 221 | 2 | 3.64 | 219 | 217.36 |
| 5 | 220 | 7 | 6.10 | 213 | 213.90 |
| 6 | 220 | 11 | 10.22 | 209 | 209.78 |
| 7 | 220 | 11 | 16.34 | 209 | 203.66 |
| 8 | 220 | 28 | 29.58 | 192 | 190.42 |
| 9 | 220 | 66 | 61.88 | 154 | 158.12 |
| 10 | 217 | 141 | 138.55 | 76 | 78.45 |

| Hosmer and Lemeshow Goodness-of-Fit Test | | |
|---|---|---|
| Chi-Square | DF | Pr > ChiSq |

| Hosmer and Lemeshow Goodness-of-Fit Test | | |
|---|---|---|
| Chi-Square | DF | Pr > ChiSq |
| 11.3417 | 8 | 0.1831 |

## ROC Curve for Metabolic Syndrome HD Prediction Model

**The LOGISTIC Procedure**

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 90.3 | Somers' D | 0.806 |
| Percent Discordant | 9.7 | Gamma | 0.806 |
| Percent Tied | 0.0 | Tau-a | 0.174 |
| Pairs | 520560 | c | 0.903 |