The MEANS Procedure

| Variable | N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|---|
| Age | 147 | 56.9659864 | 7.3488560 | 37.0000000 | 67.0000000 |
| Years_smoked | 147 | 27.8503401 | 13.9756901 | 0 | 50.0000000 |
| Cigarettes_per_day | 147 | 15.7482993 | 9.7037230 | 0 | 45.0000000 |

The FREQ Procedure

| Lung_Cancer | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 0 | 98 | 66.67 | 98 | 66.67 |
| 1 | 49 | 33.33 | 147 | 100.00 |

| Sex | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 0 | 111 | 75.51 | 111 | 75.51 |
| 1 | 36 | 24.49 | 147 | 100.00 |

| Socioeconomic_Status | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 0 | 102 | 69.39 | 102 | 69.39 |
| 1 | 45 | 30.61 | 147 | 100.00 |

| Birdkeeping | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 0 | 80 | 54.42 | 80 | 54.42 |
| 1 | 67 | 45.58 | 147 | 100.00 |

The FREQ Procedure

| Cigarettes_per_day | Frequency |
|---|---|
| 20 | 38 |
| 15 | 27 |
| 0 | 17 |
| 10 | 15 |
| 25 | 15 |
| 30 | 7 |
| 12 | 5 |
| 40 | 4 |
| 6 | 3 |
| 1 | 2 |
| 2 | 2 |
| 4 | 2 |
| 5 | 2 |
| 8 | 2 |
| 18 | 2 |
| 3 | 1 |
| 16 | 1 |
| 37 | 1 |
| 45 | 1 |

## 11: Logistic regression, with lung cancer as the response variable

The LOGISTIC Procedure

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

| Number of Observations Read | 147 |
|---|---|
| Number of Observations Used | 147 |

| Response Profile | | |
|---|---|---|
| Ordered Value | Lung_Cancer | Total Frequency |
| 1 | 0 | 98 |
| 2 | 1 | 49 |

Probability modeled is Lung_Cancer=1.

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

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 189.135 | 168.198 |
| SC | 192.126 | 189.131 |
| -2 Log L | 187.135 | 154.198 |

| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 32.9367 | 6 | <.0001 |
| Score | 27.9620 | 6 | <.0001 |
| Wald | 22.5697 | 6 | 0.0010 |

| Analysis of Maximum Likelihood Estimates | | | | | |
|---|---|---|---|---|---|
| Parameter | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept | 1 | -1.9375 | 1.8042 | 1.1531 | 0.2829 |
| Sex | 1 | 0.5613 | 0.5312 | 1.1166 | 0.2907 |
| Socioeconomic_Status | 1 | 0.1054 | 0.4688 | 0.0506 | 0.8221 |
| Birdkeeping | 1 | 1.3626 | 0.4113 | 10.9766 | 0.0009 |
| Age | 1 | -0.0398 | 0.0355 | 1.2552 | 0.2626 |
| Years_smoked | 1 | 0.0729 | 0.0265 | 7.5677 | 0.0059 |
| Cigarettes_per_day | 1 | 0.0260 | 0.0255 | 1.0390 | 0.3081 |

| Odds Ratio Estimates | | | |
|---|---|---|---|
| Effect | Point Estimate | 95% Wald Confidence Limits | |
| Sex | 1.753 | 0.619 | 4.964 |
| Socioeconomic_Status | 1.111 | 0.443 | 2.785 |
| Birdkeeping | 3.906 | 1.745 | 8.747 |
| Age | 0.961 | 0.896 | 1.030 |
| Years_smoked | 1.076 | 1.021 | 1.133 |
| Cigarettes_per_day | 1.026 | 0.976 | 1.079 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 77.4 | Somers' D | 0.549 |
| Percent Discordant | 22.5 | Gamma | 0.550 |
| Percent Tied | 0.0 | Tau-a | 0.246 |
| Pairs | 4802 | c | 0.775 |

---

the estimated odds of cancer for a participant of high socioeconomic status are _____ times as great, compared to a participant of low socioeconomic status.

| odds_ratio |
|---|
| 1.111155 |

---

confidence interval for last number

| lower | higher |
|---|---|
| 0.496214 | 2.0152596 |

---

**11fFor a non-smoking, bird-keeping woman of average age and low socioeconomic status, what is the estimated probability of lung cancer?**

| | probCancer |
|---|---|
| 11f: Estimated probability of lung cancer for the specified individual is | 0.0927217 |

---

**11g:For a non-smoking, non-bird-keeping woman of average age and low socioeconomic status, what is the estimated probability of lung cancer?**

| | probCancer |
|---|---|
| 11g: Estimated probability of lung cancer for a non-bird-keeping woman is | 0.025495 |

---

## Full Model Including Birdkeeping

### The LOGISTIC Procedure

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

| Number of Observations Read | 147 |
|---|---|
| Number of Observations Used | 147 |

| Response Profile | | |
|---|---|---|
| Ordered Value | Lung_Cancer | Total Frequency |
| 1 | 0 | 98 |
| 2 | 1 | 49 |

Probability modeled is Lung_Cancer=1.

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

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 189.135 | 168.198 |
| SC | 192.126 | 189.131 |
| -2 Log L | 187.135 | 154.198 |

| Testing Global Null Hypothesis: BETA=0 | | | |
|---|---|---|---|
| Test | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio | 32.9367 | 6 | <.0001 |
| Score | 27.9620 | 6 | <.0001 |
| Wald | 22.5697 | 6 | 0.0010 |

| Analysis of Maximum Likelihood Estimates | | | | | |
|---|---|---|---|---|---|
| Parameter | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept | 1 | -1.9375 | 1.8042 | 1.1531 | 0.2829 |
| Sex | 1 | 0.5613 | 0.5312 | 1.1166 | 0.2907 |
| Socioeconomic_Status | 1 | 0.1054 | 0.4688 | 0.0506 | 0.8221 |
| Birdkeeping | 1 | 1.3626 | 0.4113 | 10.9766 | 0.0009 |
| Age | 1 | -0.0398 | 0.0355 | 1.2552 | 0.2626 |
| Years_smoked | 1 | 0.0729 | 0.0265 | 7.5677 | 0.0059 |
| Cigarettes_per_day | 1 | 0.0260 | 0.0255 | 1.0390 | 0.3081 |

| Odds Ratio Estimates | | |
|---|---|---|
| Effect | Point Estimate | 95% Wald Confidence Limits |
| Sex | 1.753 | 0.619 4.964 |
| Socioeconomic_Status | 1.111 | 0.443 2.785 |
| Birdkeeping | 3.906 | 1.745 8.747 |
| Age | 0.961 | 0.896 1.030 |
| Years_smoked | 1.076 | 1.021 1.133 |
| Cigarettes_per_day | 1.026 | 0.976 1.079 |

| Association of Predicted Probabilities and Observed Responses | | | |
|---|---|---|---|
| Percent Concordant | 77.4 | Somers' D | 0.549 |
| Percent Discordant | 22.5 | Gamma | 0.550 |
| Percent Tied | 0.0 | Tau-a | 0.246 |
| Pairs | 4802 | c | 0.775 |

## Reduced Model Excluding Birdkeeping

### The LOGISTIC Procedure

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

| Number of Observations Read | 147 |
|---|---|
| Number of Observations Used | 147 |

| Response Profile | | |
|---|---|---|
| Ordered Value | Lung_Cancer | Total Frequency |
| 1 | 0 | 98 |
| 2 | 1 | 49 |

Probability modeled is Lung_Cancer=1.

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

| Model Fit Statistics | | |
|---|---|---|
| Criterion | Intercept Only | Intercept and Covariates |
| AIC | 189.135 | 177.868 |

## Model Fit Statistics

| Criterion | Intercept Only | Intercept and Covariates |
|---|---|---|
| SC | 192.126 | 195.811 |
| -2 Log L | 187.135 | 165.868 |

## Testing Global Null Hypothesis: BETA=0

| Test | Chi-Square | DF | Pr > ChiSq |
|---|---|---|---|
| Likelihood Ratio | 21.2668 | 5 | 0.0007 |
| Score | 17.0706 | 5 | 0.0044 |
| Wald | 13.6467 | 5 | 0.0180 |

## Analysis of Maximum Likelihood Estimates

| Parameter | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|
| Intercept | 1 | -0.2973 | 1.6600 | 0.0321 | 0.8579 |
| Sex | 1 | 0.7492 | 0.5050 | 2.2010 | 0.1379 |
| Socioeconomic_Status | 1 | -0.0730 | 0.4389 | 0.0277 | 0.8679 |
| Age | 1 | -0.0580 | 0.0343 | 2.8542 | 0.0911 |
| Years_smoked | 1 | 0.0796 | 0.0264 | 9.1056 | 0.0025 |
| Cigarettes_per_day | 1 | 0.0198 | 0.0242 | 0.6667 | 0.4142 |

## Odds Ratio Estimates

| Effect | Point Estimate | 95% Wald Confidence Limits | |
|---|---|---|---|
| Sex | 2.115 | 0.786 | 5.692 |
| Socioeconomic_Status | 0.930 | 0.393 | 2.197 |
| Age | 0.944 | 0.882 | 1.009 |
| Years_smoked | 1.083 | 1.028 | 1.140 |
| Cigarettes_per_day | 1.020 | 0.973 | 1.070 |

## Association of Predicted Probabilities and Observed Responses

| | | | |
|---|---|---|---|
| Percent Concordant | 72.2 | Somers' D | 0.446 |
| Percent Discordant | 27.7 | Gamma | 0.446 |
| Percent Tied | 0.1 | Tau-a | 0.199 |
| Pairs | 4802 | c | 0.723 |

## Reduced Model Excluding Birdkeeping

| G |
|---|
| -11.67 |

## Reduced Model Excluding Birdkeeping

| p_value |
|---|
| 0.0006352 |