Homework # 8 — solutions

(1) 14.11

a. The plot certainly indicates that deposit level had an effect on the rate of return. The pattern appears close to linear, without much sigmoid shape, but this may simply be because the proportions were never extremely low or extremely high; the log-odds transformation has only modest effects within the range of values seen in these data. I therefore see no reason to think the logistic response function is inappropriate for these data.

b. $b_0 = -2.07656, \ b_1 = 0.135851$

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>Z</th>
<th>P</th>
<th>Ratio</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.0848391</td>
<td>-24.48</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>deposit</td>
<td>0.0047720</td>
<td>28.47</td>
<td>0.000</td>
<td>1.15</td>
<td>1.13</td>
<td>1.16</td>
<td></td>
</tr>
</tbody>
</table>

\[ \hat{\pi} = \left[ 1 + e^{2.07656 - (0.13585 \cdot X)} \right]^{-1} = \frac{e^{-2.07656 + (0.13585 \cdot X)}}{1 + e^{-2.07656 + (0.13585 \cdot X)}} \]

c. The fitted logistic response function is quite similar to the pattern of the observed proportions (and actually is even less sigmoid looking). The main exception is that the observed proportion returned at a deposit of 20 cents is somewhat lower than predicted by the model.

d. $\exp(b_1) = 1.1455$. [Note that this is given (rounded to 1.15) in the output in part (b).] The odds of a bottle being returned increase by 1.1455 for every additional cent of deposit. (Also see part (e) for more on this interpretation.)

e. $\hat{\pi} = -2.07656 + 0.135851 \cdot 15 = -0.03875. \ \ \ \ \hat{\pi} = e^{\hat{\pi}'}/(1 + e^{\hat{\pi}'}) = 0.4903$
To better understand the odds ratio of 1.1455 per cent of deposit, in part (d), here are estimated probabilities of bottle return for several different deposit levels:

<table>
<thead>
<tr>
<th>deposit</th>
<th>2</th>
<th>3</th>
<th>15</th>
<th>16</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>est. probability of return</td>
<td>0.14126</td>
<td>0.15856</td>
<td>0.49030</td>
<td>0.52425</td>
<td>0.86566</td>
<td>0.88069</td>
</tr>
<tr>
<td>estimated odds of return</td>
<td>0.1645</td>
<td>0.1884</td>
<td>0.9619</td>
<td>1.1019</td>
<td>6.4438</td>
<td>7.3814</td>
</tr>
</tbody>
</table>

Neither the estimated probabilities or the estimated odds increase linearly: the increments between deposits of 1 cent and 2 cents are different than those between deposits of 15 cents and 16 cents, etc. The ratios of the odds for these pairs of deposit levels, however, all are 1.1455.

f. 23.3726.

\[ \pi = 0.75 \]

Odds = 0.75 / (1 - 0.75) = 3. So \( \pi' = \text{log-odds} = \ln(3) = 1.0986 \).

\[ \pi' = b_0 + b_1 \cdot X, \]  so 1.0986 = -2.07656 + 0.135851 \( X \).

Solve this to get \( X = 23.37 \).

(2) 14.17

a. CI for \( \beta_1 \):

\[ b_1 \pm (z_{0.975} \cdot s\{b_1\}) = 0.135851 \pm (1.96 \cdot 0.004772) \]

\[ = 0.135851 \pm 0.009353 = (0.1265, 0.1452). \]

For the odds ratio:

\[ \exp(0.1265) = 1.1348, \quad \exp(0.1452) = 1.1563. \]  [Note that these are given (rounded off) in the Minitab output above.]

We are 95% confident that the odds ratio — the proportional increase in the odds of a bottle being returned, for every one-cent increase in the deposit — is between 1.1348 and 1.1563.

b. \( H_0: \beta_1 = 0, \quad H_a: \beta_1 \neq 0. \)

\[ b_1 = 0.13585, \quad s\{b_1\} = 0.004772, \quad z^* = 0.13585 / 0.004772 = 28.468. \]

\( z_{0.975} = 1.96, \) so if \( |z^*| \leq 1.96, \) do not reject \( H_0, \) otherwise do reject \( H_0. \)

\( z^* \) is much larger than 1.96 so we reject \( H_0; \) the \( P \)-value is very small, \( << 0.0005. \)

[Note that this test is in the Minitab output above.]

c. \( H_0: \beta_1 = 0. \)

Reduced model: \( \pi = [1 + e^{-\beta_0}]^{-1} \) or \( \pi' = \beta_0, \) i.e. the probability a bottle is returned is constant, independent of the deposit level. [Note that it isn’t necessary to actually fit this model, and indeed I’m not sure how this would be done.]

\( H_0: \beta_1 \neq 0. \)

Full model: \( \pi = [1 + e^{-(\beta_0 + \beta_1 X)}]^{-1} \) or \( \pi' = \ln(\pi/(1 - \pi)) = \beta_0 + \beta_1 X. \)

\[ \chi^2_{0.95; 1} = 3.8415, \]  so if \( G^2 > 3.8415, \) reject \( H_0, \) otherwise do not reject \( H_0. \)

\( G^2 = 1095.99 \) so we reject \( H_0; \) the \( P \)-value is very small, \( << 0.0005. \)

Log-Likelihood = -1531.436

Test that all slopes are zero: \( G = 1095.990, \) DF = 1, P-Value = 0.000

This result is qualitatively the same as for the Wald test in part b.