1 Logistic Regression: Odds Ratios and Log-Odds

- Suppose that in a sample of 100 men, 90 drank wine in the previous week, while in a sample of 100 women only 20 drank wine in the same period.
- The odds of a man drinking wine are 90 to 10, or 9:1, while the odds of a woman drinking wine are only 20 to 80, or 1:4 = 0.25:1.
- The odds ratio is thus 9/0.25, or 36, showing that men are much more likely to drink wine than women.
- The detailed calculation is:

$$\frac{0.9/0.1}{0.2/0.8} = \frac{0.9 \times 0.8}{0.1 \times 0.2} = \frac{0.72}{0.02} = 36$$

- This example also shows how odds ratios are sometimes sensitive in stating relative positions: in this sample men are $90/20 = 4.5$ times more likely to have drunk wine than women, but have 36 times the odds.
- The logarithm of the odds ratio, the difference of the logits

of the probabilities, tempers this effect, and also makes the measure symmetric with respect to the ordering of groups.

- For example, using natural logarithms, an odds ratio of 36/1 maps to 3.584, and an odds ratio of 1/36 maps to -3.584.

Logistic Regression: Logits

The logit transformation is given by the following formula:

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

The inverse of the logit transformation is given by the following formula:

$$\pi_i = \text{logit}^{-1}(\eta_i) = \frac{e^{\eta_i}}{1+e^{\eta_i}}$$

Example 1

Given that $\pi_i = 0.2$, compute η_i.

$$\eta_i = \log\left(\frac{0.2}{1-0.2}\right) = \log\left(\frac{0.2}{0.8}\right)$$

$$\eta_i = \log(0.25) = -1.386$$

Example 2

Given that $\eta_i = 2.3$, compute π_i.

$$\pi_i = \frac{e^{2.3}}{1 + e^{2.3}} = \frac{9.974}{1 + 9.974} = 0.908$$

Logits

In logistic regression, the logit may be computed in a manner similar to linear regression:

$$\eta_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots$$

Example 2

Let us suppose that the probability of survival of a marine species of fauna is dependent on pollution, depth and water temperature. Suppose the logit for the logistic regression was computed as follows:

$$\eta_i = 0.14 + 0.76x_1 - 0.093x_2 + 1.2x_3$$

Variables	case 1	case 2
Pollution(x_1)	6.0	1.9
Depth (x_2)	51	99
Temp (x_3)	3.0	2.9

Compute the probability of success for both case 1 and case 2.

case 1

$$\eta_1 = 0.14 + (0.76 \times 6) - (0.093 \times 51) + (1.2 \times 3) = 3.557$$

case 2

$$\eta_2 = 0.14 + (0.76 \times 1.9) - (0.093 \times 99) + (1.2 \times 2.9) = -4.143$$

The probabilities for success are therefore:

$$\pi_1 = \frac{e^{3.557}}{1 + e^{3.557}} = \frac{35.057}{1 + 35.057} = 0.972$$

$$\pi_2 = \frac{e^{-4.143}}{1 + e^{-4.143}} = \frac{0.0158}{1 + 0.0158} = 0.0156$$