# Binary choice – 3.2 Apply the model on data

## Michel Bierlaire

*Solution of the practice quiz.*

The formulas for the logit model are:

$$P^{\text{logit}}(\text{bicycle}; \mu) = \frac{\exp\left(\mu V_{\text{bicycle}}\right)}{\exp\left(\mu V_{\text{bicycle}}\right) + \exp\left(\mu V_{\text{metro}}\right)},$$

and

$$P^{\text{logit}}(\text{metro}; \mu) = 1 - P^{\text{logit}}(\text{bicycle}; \mu).$$

The formulas for the probit model are:

$$P^{\text{probit}}(\text{bicycle}; \sigma) = \Phi\left(\frac{V_{\text{bicycle}} - V_{\text{metro}}}{\sigma}\right),$$

and

$$P^{\text{probit}}(\text{metro}; \sigma) = 1 - P^{\text{probit}}(\text{bicycle}).$$

In order to calculate the choice probabilities with various models, we need first to calculate the utility functions for the scenario that is considered. We have

$$V_{\text{bicycle}} = 0 - 0.8 \cdot 10 = -8,$$
$$V_{\text{metro}} = 3 - 0.5 \cdot 20 - 1 \cdot 2.2 = -9.2.$$

1. The logit model with $\mu = 0.1$:

$$\begin{aligned}
P^{\text{logit}}(\text{bicycle}; 0.1) &= \frac{\exp(\mu V_{\text{bicycle}})}{\exp(\mu V_{\text{bicycle}}) + \exp(\mu V_{\text{metro}})} \\
&= \frac{\exp(-0.8)}{\exp(-0.8) + \exp(-0.92)} \\
&= 0.53,
\end{aligned}$$

and

$$P^{\text{logit}}(\text{metro}; 0.1) = \frac{\exp(V_{\text{metro}})}{\exp(V_{\text{bicycle}}) + \exp(V_{\text{metro}})}$$
$$= \frac{\exp(-0.92)}{\exp(-0.8) + \exp(-0.92)}$$
$$= 0.47.$$

2. The logit model with $\mu = 10$:

$$P^{\text{logit}}(\text{bicycle}; 10) = \frac{\exp(\mu V_{\text{bicycle}})}{\exp(\mu V_{\text{bicycle}}) + \exp(\mu V_{\text{metro}})}$$
$$= \frac{\exp(-80)}{\exp(-80) + \exp(-92)}$$
$$= 0.999994,$$

and

$$P^{\text{logit}}(\text{metro}; 10) = \frac{\exp(V_{\text{metro}})}{\exp(V_{\text{bicycle}}) + \exp(V_{\text{metro}})}$$
$$= \frac{\exp(-92)}{\exp(-80) + \exp(-92)}$$
$$= 0.000006.$$

3. The probit model with $\sigma = 0.1$:

$$P^{\text{probit}}(\text{bicycle}; 0.1) = \Phi((V_{\text{bicycle}} - V_{\text{metro}})/\sigma) = \Phi(12) \approx 1,$$

and

$$P^{\text{probit}}(\text{metro}; 0.1) = \Phi(V_{\text{metro}} - V_{\text{bicycle}}) = \Phi(-12) \approx 0.$$

4. The probit model with $\sigma = 10$:

$$P^{\text{probit}}(\text{bicycle}; 10) = \Phi((V_{\text{bicycle}} - V_{\text{metro}})/\sigma) = \Phi(0.12) = 0.55,$$

and

$$P^{\text{probit}}(\text{metro}; 10) = \Phi(V_{\text{metro}} - V_{\text{bicycle}}) = \Phi(-0.12) = 0.45.$$

The results are summarized in the following table:

|  | Logit | | Probit | |
|---|---|---|---|---|
|  | $\mu = 0.1$ | $\mu = 10$ | $\sigma = 0.1$ | $\sigma = 10$ |
| $P(\text{bicycle})$ | 0.53 | 0.999994 | 1 | 0.55 |
| $P(\text{metro})$ | 0.47 | 0.000006 | 0 | 0.45 |

This exercise illustrates the importance of the scale parameter in the calculation of the choice probability. If the $\beta$'s are set, varying the scale parameter modifies the choice probability. As the scale is normalized to one at the estimation stage, it is tempting to forget about it. But it is embedded in the values of the coefficients.

The results also illustrate the limiting cases of the models:

**Probit** There are two limiting cases of a probit model of special interest, both involving extreme values of the scale parameter. The first case is for $\sigma \to 0$:
$$\lim_{\sigma \to 0} P_n(i) = \begin{cases} 1 & \text{if} \quad V_{in} - V_{jn} > 0, \\ 0 & \text{if} \quad V_{in} - V_{jn} < 0; \end{cases}$$

this is, as $\sigma \to 0$, the variance vanishes and the choice model is deterministic. On the other hand, when $\sigma \to \infty$, the choice probabilities become $1/2$. Intuitively the model predicts equal probability of choice for each alternative irrespectively of $V_{\text{metro}}$ and $V_{\text{bicycle}}$.

**Logit** As with probit, if the $V$'s are set, there are two limiting cases of a binary logit model that are of special interest. The first case is for $\mu \to \infty$:
$$\lim_{\mu \to \infty} P_n(i) = \begin{cases} 1 & \text{if} \quad V_{in} - V_{jn} > 0, \\ 0 & \text{if} \quad V_{in} - V_{jn} < 0; \end{cases}$$

that is, as $\mu \to \infty$, the variance vanishes and the choice model is deterministic. On the other hand, when $\mu \to 0$, the choice probabilities become $1/2$.

Note that, for probit, the variance is proportional to the (squared) scale parameter, while for logit, it is inversely proportional.