

Econometrics II TA Session #3*

Makoto SHIMOSHIMIZU[†]

Room 4, October 29, 2019

Contents

1	Qualitative Dependent Variable	2
1.1	What is “Qualitative Dependent Variable”?	2
1.2	Models for Qualitative Dependent Variable	2
2	Discrete Choice Model	2
2.1	Binary Choice Model	2
2.2	The Probit and Logit Model	3
2.3	Likelihood Function	4
2.4	Another Interpretation	5
2.5	Marginal Effect (“When x_i increase by 1%, how much y_i would increase?”)	6
3	Some Applications	6
3.1	Ordered Probit or Logit Model	6
3.2	Multinomial Logit Model	7
4	Limited Dependent Variable Model	7
4.1	Types of Limited Dependent Variable Models	8
4.2	Truncated Regression Model	8
A	Logistic Distribution: More Detail	10
A.1	Definition and Introduction	10
A.2	Genesis	11
A.3	Mean, Variance and Moment Generating Function	11
B	Derivation of Eq. (4.5)	12

*All comments welcome!

[†]E-mail: vgm067sm@student.econ.osaka-u.ac.jp

1 Qualitative Dependent Variable

1.1 What is “Qualitative Dependent Variable”?

In most cases, the dependent variable such as

- temperature
- individual income

is continuous or assumed to be continuous. In this case, y_i for $i \in \{1, \dots, n\}$ is continuous and are supposed to take any value in \mathbb{R} , i.e., $y_i \in (-\infty, \infty)$. However, we may encounter some variables which take only several values. Examples are:

- male or female;
- smoking or not smoking.

Such cases admit

$$y_i = \begin{cases} 0 & \text{if male (smoking);} \\ 1 & \text{if female (not smoking),} \end{cases}$$

which does not allow y_i to be continuous. We thus learn the case that these variables are **dependent variables**.

1.2 Models for Qualitative Dependent Variable

We here learn mainly three models used for the analysis of qualitative dependent variables:

- (i) **Discrete Choice Model**
- (ii) **Limited Dependent Variable Model**
- (iii) **Count Data Model**

2 Discrete Choice Model

In this section, the structure of a discrete choice model is explained. The analysis of individual choice that is the main focus of the microeconometrics is fundamentally about modeling discrete outcomes such as purchase decisions, voting behavior, places to live, and responses to survey questions about the strength of preferences or about the self-assessed health or well-being. Here we focus on modeling probabilities and using econometric tools to make probabilistic statements about the occurrence of these events.

2.1 Binary Choice Model

The study of **binary choice model** allows us to focus on appropriate specification, estimation and use of models for the probabilities of events, where in most cases, the “event” is an individual’s choice among a set of two alternatives.

Consider the following regression model:

$$y_i^* = X_i\beta + u_i, \quad u_i \sim (0, \sigma^2), \quad i = 1, 2, \dots, n, \quad (2.1)$$

where y^* is unobserved, but y_i is observed as 0 or 1,

$$y_i = \begin{cases} 1, & \text{if } y_i^* > 0, \\ 0, & \text{if } y_i^* \leq 0. \end{cases} \quad (2.2)$$

Consider the probability that y_i takes 1, i.e.

$$\begin{aligned} \mathbb{P}(y_i = 1) &= \mathbb{P}(y_i^* > 0) \\ &= \mathbb{P}(u_i > -X_i\beta) \\ &= \mathbb{P}(u_i^* > -X_i\beta^*) \\ &= 1 - \mathbb{P}(u_i^* \leq -X_i\beta^*) \\ &= 1 - F(-X_i\beta^*) \\ &= F(X_i\beta^*) \end{aligned} \quad (2.3)$$

where u_i^* and β^* are defined as

$$u_i^* = \frac{u_i}{\sigma}, \quad \beta^* = \frac{\beta}{\sigma}$$

The last equality of Eq. (2.3) comes from the assumption of **symmetricity of the distribution** of u_i^* , i.e., $1 - F(-x) = F(x)$. Note that we can estimate β^* but cannot estimate β and σ separately. The cumulative distribution function is given by

$$F(x) = \int_{-\infty}^x f(z)dz, \quad (2.4)$$

where $f(x)$ stands for the probability density function.

2.2 The Probit and Logit Model

Here we introduce the probit and logit model. The normal distribution has been used in many analysis, giving rise to the **probit model**,

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz =: \Phi(x). \quad (2.5)$$

The function $\Phi(x)$ is a commonly used notation for the standard normal distribution function, that is,

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz. \quad (2.6)$$

On the other hand, partly because of mathematical convenience, the **logistic distribution**,

$$F(x) = \frac{1}{1 + \exp(-x)} =: \Lambda(x), \quad (2.7)$$

has also been used in many applications. For this distribution function, the probability density function $f(x)$ becomes

$$f(x) = \frac{\exp(-x)}{(1 + \exp(-x))^2}. \quad (2.8)$$

This model is called the **logit model**. We can also consider other models for u_i^* which do not assume to be symmetric. For example, the **Gumbel distribution**,

$$\mathbb{P}(Y = 1) = \exp\{-\exp\{-X\beta\}\},$$

and **complementary log log model**,

$$\mathbb{P}(Y = 1) = 1 - \exp\{-\exp\{X\beta\}\},$$

have also been employed. Although other distributions have been suggested, the probit and logit models are still the most common frameworks used in econometric applications.

2.3 Likelihood Function

y_i for $i \in \{1, \dots, n\}$ follow the following **Bernouli distribution** $f(y_i)$ as follows:

$$\begin{aligned} f(y_i) &= (\mathbb{P}(y_i = 1))^{y_i} (1 - \mathbb{P}(y_i = 1))^{1-y_i} \\ &= (F(X_i\beta^*))^{y_i} (1 - F(X_i\beta^*))^{1-y_i} \end{aligned}$$

Here we review the definition of the Bernouli distribution.

Definition 2.1 (Bernouli Distribution). The **Bernouli** random variable $X: \Omega \rightarrow \{0, 1\}$ has a following probability mass function, denoted by $f(x)$:

$$f(x) := p^x(1-p)^{1-x}, \quad x = 0, 1.$$

Then, the mean and variance of X are:

$$\begin{aligned} \mu &:= \mathbb{E}[X] = \sum_{x=0}^1 x f(x) = 0 \times (1-p) + 1 \times p = p; \\ \sigma^2 &:= \mathbb{V}[X] = \sum_{x=0}^1 (x - \mu)^2 f(x) = (0-p)^2(1-p) + (1-p)^2 p = p(1-p). \end{aligned}$$

According to the definition of the Bernouli distribution, we obtain the likelihood function as follows:

$$L(\beta^*) = f(y_1, y_2, \dots, y_n) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n (F(X_i\beta^*))^{y_i} (1 - F(X_i\beta^*))^{1-y_i}.$$

Then the log-likelihood function becomes:

$$\log L(\beta^*) = \sum_{i=1}^n \{y_i \log F(X_i\beta^*) + (1 - y_i) \log(1 - F(X_i\beta^*))\}. \quad (2.9)$$

Solving the maximization problem of $\log L(\beta^*)$ with respect to β^* , we can derive the F.O.C. as follows:

$$\begin{aligned} \frac{\partial \log L(\beta^*)}{\partial \beta^*} &= \sum_{i=1}^n \left(\frac{y_i X_i' f(X_i\beta^*)}{F(X_i\beta^*)} - \frac{(1 - y_i) X_i' f(X_i\beta^*)}{1 - F(X_i\beta^*)} \right) \\ &= \sum_{i=1}^n \frac{X_i' f(X_i\beta^*) (y_i - F(X_i\beta^*))}{F(X_i\beta^*) (1 - F(X_i\beta^*))} = 0. \end{aligned} \quad (2.10)$$

Then the S.O.C. is given by

$$\frac{\partial^2 \log L(\beta^*)}{\partial \beta^* \partial \beta^{*t}} = \sum_{i=1}^n \frac{X_i' X_i f_i'(y_i - F_i)}{F_i(1 - F_i)} - \sum_{i=1}^n \frac{X_i' X_i f_i^2}{F_i(1 - F_i)} - \sum_{i=1}^n X_i' f_i(y_i - F_i) \frac{X_i f_i(1 - 2F_i)}{(F_i(1 - F_i))^2}, \quad (2.11)$$

and Eq. (2.11) is negative definite. When we adopt the Logit model, from Eq. (2.10) we can calculate in more detail:

$$\begin{aligned} \frac{\partial L(\beta^*)}{\partial \beta^*} &= \sum_{i=1}^n \frac{X_i' f(X_i \beta^*)(y_i - F(X_i \beta^*))}{F(X_i \beta^*)(1 - F(X_i \beta^*))} \\ &= \sum_{i=1}^n \frac{X_i' \frac{\exp(-X_i \beta^*)}{(1 + \exp(-X_i \beta^*))^2} \left(y_i - \frac{1}{1 + \exp(-X_i \beta^*)} \right)}{\frac{1}{1 + \exp(-X_i \beta^*)} \frac{\exp(-X_i \beta^*)}{1 + \exp(-X_i \beta^*)}} \\ &= \sum_{i=1}^n X_i' \left(y_i - \frac{1}{\exp(-X_i \beta^*)} \right) = 0. \end{aligned}$$

For maximization, the **method of scoring** is given by

$$\beta^{*(j+1)} = \beta^{*(j)} + \left(\sum_{i=1}^n \frac{X_i' X_i (f_i^{(j)})^2}{F_i^{(j)}(1 - F_i^{(j)})} \right)^{-1} \sum_{i=1}^n \frac{X_i' f_i^{(j)}(y_i - F_i^{(j)})}{F_i^{(j)}(1 - F_i^{(j)})}. \quad (2.12)$$

Variance of MLE $\hat{\beta}^*$ is $I(\hat{\beta}^*)^{-1}$ where

$$I(\hat{\beta}^*) = -E \left[\frac{\partial^2 \log L(\hat{\beta}^*)}{\partial \beta^* \partial \beta^{*t}} \right] = \sum_{i=1}^n \frac{X_i' X_i \hat{f}_i^2}{\hat{F}_i(1 - \hat{F}_i)}. \quad (2.13)$$

We can estimate $\hat{\beta}^*$ and test the significance of $\hat{\beta}^*$.

2.4 Another Interpretation

This maximization problem is equivalent to the nonlinear least squares estimation problem from the following regression model:

$$y_i = F(X_i \beta^*) + u_i, \quad (2.14)$$

where

$$u_i = y_i - F_i = \begin{cases} 1 - F_i & \text{w.p. } \mathbb{P}(y_i = 1); \\ 0 - F_i & \text{w.p. } \mathbb{P}(y_i = 0). \end{cases} \quad (2.15)$$

Therefore, the mean and variance of u_i are:

$$\mathbb{E}[u_i] = \underbrace{(1 - F_i)}_{\text{value}} \underbrace{F_i}_{\text{probability}} + \underbrace{(-F_i)}_{\text{value}} \underbrace{(1 - F_i)}_{\text{probability}} = 0 \quad (2.16)$$

$$\begin{aligned} \sigma_i^2 &= \mathbb{V}(u_i) = \mathbb{E}[u_i^2 - \mathbb{E}(u_i)^2] = \mathbb{E}(u_i^2) \\ &= \underbrace{(1 - F_i)^2}_{\text{value}} \underbrace{F_i}_{\text{probability}} + \underbrace{(-F_i)^2}_{\text{value}} \underbrace{(1 - F_i)}_{\text{probability}} = F_i(1 - F_i). \end{aligned} \quad (2.17)$$

Then the **weighted least squares method** solves the following minimization problem:

$$\min_{\beta^* \in \mathbb{R}^k} \sum_{i=1}^n \frac{(y_i - F(X_i\beta^*))^2}{\sigma_i^2}, \quad (2.18)$$

which corresponds to a **generalized least squares (GLS) method**. Then, the first order condition becomes:

$$\sum_{i=1}^n \frac{2X_i' f(X_i\beta^*)(y_i - F(X_i\beta^*))}{\sigma_i^2} = 0 \quad (2.19)$$

This is equivalent to the first order condition of MLE.

2.5 Marginal Effect (“When x_i increase by 1%, how much y_i would increase?”)

When we employ the OLS method ($y = x_1\beta_1 + \dots + x_k\beta_k + \epsilon$), the marginal effect becomes “ $\beta_i\%$.” When we conduct the probit or logit estimation, the result is not so straightforward as the OLS method. The model is represented as follows:

$$y_i^* = \mathbb{P}(y_i = 1) + u_i = F(X_i\beta^*) + u_i. \quad (2.20)$$

By differentiating this equation with respect to $X_{i,j}$, the j th independent variable of individual i for $j \in \{1, \dots, k\}$ and $i \in \{1, \dots, n\}$, we obtain

$$\frac{d\mathbb{P}(y_i = 1)}{dX_{i,j}} = \frac{dF(X_i\beta^*)}{dX_{i,j}} = \beta_j^* f(X_i\beta^*). \quad (2.21)$$

Where $f(X_i\beta^*)$ is probability density function of $X_i\beta^*$.

3 Some Applications

In this section, we show some applications of discrete choice models.

3.1 Ordered Probit or Logit Model

The **Ordered Probit or Logit Model** is the case where y_i is observed as $1, 2, \dots, m$ shown as follows:

$$y_i = \begin{cases} 1 & y_i^* < a_1 \\ 2 & a_1 \leq y_i^* < a_2 \\ \vdots & \\ m & a_{m-1} \leq y_i^* \end{cases} \quad (3.1)$$

Then the probability density function of y_i is given by

$$f(y_i) = (\mathbb{P}(y_i = 1))^{I_{\{i=1\}}} (\mathbb{P}(y_i = 2))^{I_{\{i=2\}}} \dots (\mathbb{P}(y_i = m))^{I_{\{i=m\}}},$$

where

$$\begin{aligned}
\mathbb{P}(y_i = j) &= \mathbb{P}(a_{j-1} \leq y_i^* < a_j) \\
&= \mathbb{P}(y_i^* < a_j) - \mathbb{P}(y_i^* < a_{j-1}) \\
&= \mathbb{P}(u_i < a_j - X_i\beta) - \mathbb{P}(u_i < a_{j-1} - X_i\beta) \\
&= F(a_j - X_i\beta) - F(a_{j-1} - X_i\beta)
\end{aligned}$$

and

$$I_{\{i=j\}} = \begin{cases} 1 & \text{if } y_i = j \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, 2, \dots, m$. Note that $a_0 = -\infty, a_m = \infty$. The Likelihood function thus becomes:

$$L(\beta) = \prod_{i=1}^n f(y_i) \quad (3.2)$$

3.2 Multinomial Logit Model

This model is constructed for an unordered choice model which applies when data are individual specific. For example,

$$y_i = \begin{cases} 0 & \text{menial;} \\ 1 & \text{blue collar;} \\ 2 & \text{craft;} \\ 3 & \text{white collar;} \\ 4 & \text{professional.} \end{cases} \quad (3.3)$$

When $y_i \in \{0, 1, \dots, m\}$, the individual has $m + 1$ choices, i.e. $j = 0, 1, 2, \dots, m$:

$$\mathbb{P}(y_i = j) = \frac{\exp(X_i\beta_j)}{\sum_{j=0}^m \exp(X_i\beta_j)} =: P_{ij} \quad (3.4)$$

for $\beta_0 = 0$ (The case where $m = 1$ corresponds to a bivariate logit model). Note that

$$\log \frac{P_{ij}}{P_{i0}} = X_i\beta_j. \quad (3.5)$$

The log likelihood function is

$$\log L(\beta_1, \beta_2, \dots, \beta_m) := \sum_{i=1}^n \sum_{j=0}^m I_{\{y_i=j\}} \log P_{ij}. \quad (3.6)$$

4 Limited Dependent Variable Model

This section is concerned with a brief exposition of “truncation” and “censoring” and then explains the Truncated regression model in detail. The truncation effects arise when one attempts to make inferences about a larger population from a sample that is drawn from a distinct subpopulation. On the other hand, the censoring of a range of values of the variable of interest introduces a distortion into conventional statistical results that are similar to that of truncation.

4.1 Types of Limited Dependent Variable Models

There are mainly three models relevant with limited dependent variable models:

- (i) **Truncated Regression Model**
- (ii) **Tobit Model**
- (iii) **Censored Data Model**

In the following, we learn how to obtain the estimator of the Truncated Regression Model, where the **truncated mean of a normal distribution** plays a fundamental role.

4.2 Truncated Regression Model

In this subsection, we are concerned with inferring the characteristics of a full population from a sample drawn from a restricted part of that population. Here we consider the following model:

$$y_i = X_i\beta + u_i, \quad u_i \sim N(0, \sigma^2)$$

and only $y_i > a$ is observed (i.e. the data $y_i \leq a$ is not observed). Then, the conditional cumulative distribution and probability density function of the error term u_i becomes

$$F(u_i|y_i > a) = F(u_i|u_i > a - X_i\beta) = \int_{a-X_i\beta}^{\infty} \frac{f(u_i)}{1 - F(a - X_i\beta)} du_i;$$

$$f(u_i|y_i > a) = f(u_i|u_i > a - X_i\beta) = \frac{f(u_i)}{1 - F(a - X_i\beta)},$$

and the conditional expectation of y_i is given by

$$\mathbb{E}[u_i|y_i > a] = \mathbb{E}[u_i|u_i > a - X_i\beta] = \int_{a-X_i\beta}^{\infty} u_i \frac{f(u_i)}{1 - F(a - X_i\beta)} du_i. \quad (4.1)$$

Then we have the following facts.

Proposition 4.1. Using the following standard normal density and distribution functions:

$$\phi(x) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}x^2\right\};$$

$$\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}z^2\right\} dz = \int_{-\infty}^x \phi(z) dz,$$

the cumulative distribution and probability density function become

$$f(x) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}x^2\right\} = \frac{1}{\sigma} \phi\left(\frac{x}{\sigma}\right); \quad (4.2)$$

$$F(x) = \int_{-\infty}^x (2\pi\sigma)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}z^2\right\} dz = \Phi\left(\frac{x}{\sigma}\right). \quad (4.3)$$

Proof. The direct calculation yields Eq. (4.2). As for Eq. (4.3), using the change of variables, we have

$$\begin{aligned} F(x) &= \int_{-\infty}^x (2\pi\sigma)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}z^2\right\} dz \\ &= \int_{-\infty}^{x/\sigma} (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}w^2\right\} dw \\ &= \Phi\left(\frac{x}{\sigma}\right), \end{aligned}$$

which proves Eq. (4.2). □

Then, for a truncated normal random variable, we have the following theorem.

Theorem 4.1 (Moments of the Truncated Normal Distribution). If $X \sim N(\mu, \sigma^2)$ and a is a constant, then

$$\mathbb{E}[X|X > a] = \mu + \sigma\lambda(\alpha); \quad (4.4)$$

$$\mathbb{V}[X|X > a] = \sigma^2[1 - \delta(\alpha)], \quad (4.5)$$

where $\alpha = (a - \mu)/\sigma$, $\lambda(\alpha) = \phi(\alpha)/[1 - \Phi(\alpha)]$ and $\delta(\alpha) = \lambda(\alpha)[\lambda(\alpha) - \alpha]$. $\lambda(\alpha) = \phi(\alpha)/[1 - \Phi(\alpha)]$ is called the **inverse Mill's ratio** or **hazard function** for the standard normal distribution.

Proof. For the probability distribution function, the following relations hold:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} \\ \iff \frac{d}{dx}f(x) &= -\frac{(x - \mu)}{\sigma^2}f(x) = -\frac{x}{\sigma^2}f(x) + \frac{\mu}{\sigma^2}f(x) \\ \iff \sigma^2 df(x) &= -xf(x)dx + \mu f(x)dx \\ \iff xf(x)dx &= \mu f(x)dx - \sigma^2 df(x). \end{aligned} \quad (4.6)$$

Thus, using Eq. (4.6), we can calculate

$$\begin{aligned} \mathbb{E}[X|X > a] &= \int_a^\infty x \frac{f(x)}{1 - \Phi(\alpha)} dx \\ &= \frac{1}{1 - \Phi(\alpha)} \int_a^\infty xf(x)dx \\ &= \frac{1}{1 - \Phi(\alpha)} \int_a^\infty \{\mu f(x)dx - \sigma^2 df(x)\} \\ &= \frac{1}{1 - \Phi(\alpha)} \left\{ \int_a^\infty \mu f(x)dx - \int_a^\infty \sigma^2 df(x) \right\} \\ &= \frac{1}{1 - \Phi(\alpha)} \left\{ \mu[F(x)]_a^\infty - \sigma^2[f(x)]_a^\infty \right\} \\ &= \frac{1}{1 - \Phi(\alpha)} \left\{ \mu(1 - \Phi(a)) - \sigma^2(-f(a)) \right\} \\ &= \mu + \sigma^2 \frac{f(a)}{1 - \Phi(\alpha)}. \end{aligned}$$

Using Eq. (4.2), the above equation yields

$$\mathbb{E}[X|X > a] = \mu + \sigma^2 \frac{\frac{1}{\sigma} \phi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)} = \mu + \sigma \frac{\phi(\alpha)}{1 - \Phi(\alpha)}, \quad (4.7)$$

which completes the proof. \square

Therefore, the conditional expectation of y_i given $y_i = X_i\beta + u_i > a$ for $i \in \{1, \dots, n\}$ is given by

$$\begin{aligned} \mathbb{E}[y_i|y_i > a] &= \mathbb{E}[X_i\beta + u_i|X_i\beta + u_i > a] \\ &= X_i\beta + \mathbb{E}[u_i|u_i > a - X_i\beta] \\ &= X_i\beta + \underbrace{\mathbb{E}[u_i]}_{=0} + \sigma \frac{\phi\left(\frac{a-X_i\beta}{\sigma}\right)}{1 - \Phi\left(\frac{a-X_i\beta}{\sigma}\right)} \\ &= X_i\beta + \sigma \frac{\phi\left(\frac{a-X_i\beta}{\sigma}\right)}{1 - \Phi\left(\frac{a-X_i\beta}{\sigma}\right)}. \end{aligned} \quad (4.8)$$

The above equation clearly shows that the mean of the truncated distribution has sample selection bias. In this case, the OLS estimator is a biased estimator, since

$$\begin{aligned} \mathbb{E}[\beta_{OLS}|y_i > a] &= \left(\sum_{i=1}^n X_i X_i' \right)^{-1} \sum_{i=1}^n X_i \mathbb{E}[y_i|y_i > a] \\ &= \left(\sum_{i=1}^n X_i X_i' \right)^{-1} \sum_{i=1}^n X_i \left[X_i\beta + \sigma \frac{\phi\left(\frac{a-X_i\beta}{\sigma}\right)}{1 - \Phi\left(\frac{a-X_i\beta}{\sigma}\right)} \right] \\ &= \beta + \sigma \left(\sum_{i=1}^n X_i X_i' \right)^{-1} \sum_{i=1}^n X_i \frac{\phi\left(\frac{a-X_i\beta}{\sigma}\right)}{1 - \Phi\left(\frac{a-X_i\beta}{\sigma}\right)} \end{aligned} \quad (4.9)$$

holds. Thus, we use the MLE for the estimation of β . We obtain the MLE by constructing the likelihood function $L(\beta, \sigma^2)$ as follows:

$$L(\beta, \sigma^2) = \prod_{i=1}^n \frac{f(y_i - X_i\beta)}{1 - F(a - X_i\beta)} = \prod_{i=1}^n \frac{1}{\sigma} \frac{\phi\left(\frac{y_i - X_i\beta}{\sigma}\right)}{1 - \Phi\left(\frac{a - X_i\beta}{\sigma}\right)} \quad (4.10)$$

and maximizing $L(\beta, \sigma^2)$ with respect to β and σ^2 .

Appendix

A Logistic Distribution: More Detail

Here we mention the more formal definition and features of the logistic distribution in more detail. The argument below is based on [2].

A.1 Definition and Introduction

The distribution of the logistic distribution is mostly simply defined in terms of its cumulative distribution function $F(x)$:

$$F(x) := \frac{1}{[1 + \exp\{-(x - \alpha)/\beta\}]} = \frac{1}{2} \left[1 + \tanh \left\{ \frac{1}{2}(x - \alpha)/\beta \right\} \right] \quad (A.1)$$

with $\beta > 0$. It can be seen that Eq. (A.1) defines a proper cumulative distribution with

$$\lim_{x \rightarrow -\infty} F(x) = 0; \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

The corresponding probability density function is

$$f(x) = \beta^{-1} \frac{[\exp\{-(x - \alpha)/\beta\}]}{[1 + \exp\{-(x - \alpha)/\beta\}]^2} = (4\beta)^{-1} \operatorname{sech}^2 \left\{ \frac{1}{2}(x - \alpha)/\beta \right\}.$$

Then, the distribution is sometimes called the sech-squared(d) distribution. Putting $\alpha \equiv 0$ and $\beta \equiv 1$ yields

$$F(x) = \frac{1}{1 + \exp(-x)}; \quad f(x) = \frac{\exp(-x)}{(1 + \exp(-x))^2},$$

which is identical to Eq. (2.7) and (2.8), respectively.

A.2 Genesis

The use of the logistic function as a growth curve can be based on the following differential equation:

$$\frac{dF}{dx} = c[F(x) - A][B - F(x)], \quad (\text{A.2})$$

where c, A and B are constants with $c > 0, B > A$. The solution of Eq. (A.2) leads to

$$F(x) = \frac{BD \exp\{x/c\} + A}{D \exp\{x/c\} + 1}, \quad (\text{A.3})$$

where D is a constant. If $D \neq 0$, as $x \rightarrow -\infty$,

$$\lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} \frac{BD \exp\{x/c\} + A}{D \exp\{x/c\} + 1} = A.$$

Also, as $x \rightarrow \infty$, by using the L'Hôpital's rule, we obtain

$$\lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} \frac{BD \exp\{x/c\} + A}{D \exp\{x/c\} + 1} = \lim_{x \rightarrow \infty} B = B.$$

When $A = 0, B = 1$, Eq. (A.3) becomes

$$F(x) = \frac{D \exp\{x/c\}}{1 + D \exp\{x/c\}} = \frac{1}{1 + D^{-1} \exp\{-x/c\}},$$

which is of the form of Eq. (A.1) with $c = \beta, D = \exp\{-\alpha/\beta\}$.

A.3 Mean, Variance and Moment Generating Function

Here we show the form of the mean, variance and moment generating function for a random variable that follows a logistic distribution (without proof). (If you want to see the proof, see [2].)

The mean and variance of the logistic distributed random variable becomes:

$$\mathbb{E}[X] = \alpha; \quad \mathbb{V}[X] = \frac{\beta^2 \pi^2}{3}. \quad (\text{A.4})$$

Then the moment generating function is given by

$$\mathbb{E}[\exp\{\theta X\}] = B(1 - \theta, 1 + \theta) = \pi \theta \operatorname{cosec} \pi \theta. \quad (\text{A.5})$$

B Derivation of Eq. (4.5)

Here we prove Eq. (4.5). Direct calculation yields

$$\begin{aligned}
\mathbb{E}[X^2|X > a] &= \int_a^\infty x^2 \frac{f(x)}{1 - \Phi(\alpha)} dx \\
&= \int_a^\infty x \left\{ \sigma^2 \left(\frac{x - \mu}{\sigma^2} \right) + \mu \right\} \frac{f(x)}{1 - \Phi(\alpha)} \\
&= \frac{1}{1 - \Phi(\alpha)} \left[\sigma^2 \int_a^\infty x \left(\frac{x - \mu}{\sigma^2} \right) f(x) dx \right] + \mu \int_a^\infty x \frac{f(x)}{1 - \Phi(\alpha)} dx \\
&= \frac{1}{1 - \Phi(\alpha)} \sigma^2 \int_a^\infty x \frac{d}{dx} \{-f(x)\} dx + \mu \mathbb{E}[X|X > a] \\
&= \underbrace{\frac{1}{1 - \Phi(\alpha)} \sigma^2 [-xf(x)]_a^\infty + \frac{1}{1 - \Phi(\alpha)} \sigma^2 \int_a^\infty \frac{d}{dx} xf(x) dx}_{\text{integration by parts}} + \mu \mathbb{E}[X|X > a] \\
&= \frac{\sigma^2}{1 - \Phi(\alpha)} [-0 + af(a)] + \frac{\sigma^2}{1 - \Phi(\alpha)} \int_a^\infty f(x) dx + \mu \mathbb{E}[X|X > a] \\
&= \frac{\sigma^2}{1 - \Phi(\alpha)} af(a) + \frac{\sigma^2}{1 - \Phi(\alpha)} (1 - \Phi(\alpha)) + \mu \mathbb{E}[X|X > a] \\
&= \frac{\sigma^2}{1 - \Phi(\alpha)} a \frac{1}{\sigma} \phi(\alpha) + \sigma^2 + \mu(\mu + \sigma\lambda(\alpha)) \\
&= \sigma a \lambda(\alpha) + \sigma^2 + \mu^2 + \mu\sigma\lambda(\alpha). \tag{B.1}
\end{aligned}$$

Thus, using Eq. (B.1), we have

$$\begin{aligned}
\mathbb{V}[X|X > a] &= \mathbb{E}[X^2|X > a] - (\mathbb{E}[X|X > a])^2 \\
&= \sigma a \lambda(\alpha) + \sigma^2 + \mu^2 + \mu\sigma\lambda(\alpha) - (\mu + \sigma\lambda(\alpha))^2 \\
&= \sigma^2 \left\{ 1 - \lambda(\alpha)^2 + \frac{a - \mu}{\sigma} \lambda(\alpha) \right\} \\
&= \sigma^2 \left\{ 1 - \lambda(\alpha) \left[\lambda(\alpha) - \frac{a - \mu}{\sigma} \right] \right\} \\
&= \sigma^2 \{1 - \lambda(\alpha) [\lambda(\alpha) - \alpha]\}, \tag{B.2}
\end{aligned}$$

which completes the proof. \square

References

- [1] Greene, W. H., *Econometric analysis Seventh Edition*. 2012, Pearson.
- [2] Johnson, N. L., Kotz, S. and Balakrishnan, N., *Continuous univariate distributions*. 1970, Houghton Mifflin Boston.