8.1 二值因变量模型与一般线性回归模型的本质区别是什么？

8.2 二值因变量模型的连接函数需要满足什么条件？

8.3 probit 模型和 logit 模型有什么区别，那个更好一些？

8.4 为什么要采用数值方法求二值因变量模型对数似然函数的最大值点和最大值？

8.5 为什么二值因变量模型中参数显著性检验采用的是 z 统计量（z-Statistic）而不是 t 统计量（t-Statistic）？

8.6 什么是完全分离？完全分离会给二值因变量模型的估计带来什么问题？如何处理？

8.7 在分析二值因变量模型的估计结果时，需要注意什么问题？

8.8 1973 年密西根州 Troy 市打算增加税收，用于公立学校的建设。为此就是否同意该项加税计划进行公民投票[①]。表 8.1 给出了 95 个家庭的投票数据。

表 8.4

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0	1	0	0	10	1	9.77	7.05
0	0	1	0	0	8	0	10.0	7.05
0	1	0	0	0	4	0	10.0	7.05
…	…	…	…	…	…	…	…	…
0	1	0	0	0	26	0	9.77	6.75
0	0	0	0	1	18	0	10.2	7.5

，其中 y 表示是否同意，$y=1$ 表示同意，$y=0$ 表示不同意；x_1 表示是否有 1 到 2 个孩子在公立学校上学，1 表示是，0 表示否；x_2 表示是否有 3 到 4 个孩子在公立学校上学，1 表示是，0 表示否；x_3 表示是否有 5 个或更多孩子在公立学校上学，1 表示是，0 表示否；x_4 表示是否有孩子在私立学校上学，1 表示是，0 表示否；x_5 表示在该地区生活的年数；x_6 表示该地区的学校数（包括私立和公立）；x_7 表示家庭收入的自然对数；x_8 表示每年缴纳财产税的自然对数。

（1）对公民投票选择建立 logit 模型和 probit 模型，并用 EViews 进行估计。

（2）对估计结果进行分析；

（3）比较 logit 模型和 probit 模型估计上的差别。

8.9* 设 $\Phi(x)$ 和 $\Lambda(x)$ 分别为标准正态分布函数和逻辑概率分布函数，$0\le\lambda\le1$ 为实数。

（1）证明 $F(x)=\lambda\Phi(x)+(1-\lambda)\Lambda(x)$ 为分布函数，满足 $1-F(-x)=F(x)$；

（2）写出以 $F(x)$ 为连接函数、以 $Z=\beta_0+\beta_1X_1+\beta_2X_2$ 为指标函数的二元选择模型。

◆ 参考答案

1. 一般线性回归模型是分析自变量对因变量 Y 的解释能力，二值因变量模型是分析自变量对因变量 $Y=1$ 概率的解释能力。二值因变量模型为非线性模型。
2. 二值因变量模型的连接函数必须是概率分布函数，并且满足对称性条件 $F(x)=1-F(-x)$。
3. probit 模型和 logit 模型采用了不同的连接函数。当自变量值很大或者很小时，logit 模型更能反映出自变量变化对概率的影响。此外，logit 模型可以变换为线性模型。

[①] S. 平狄克、L. 鲁宾菲尔德，《计量经济模型与经济预测》，钱小军 译，机械工业出版社，1999 年。

与 probit 模型相比，logit 模型好一些。

4. 二值因变量模型的对数似然函数十分复杂，不能通过导数等于 0 的方法求最大值点，需要数值迭代方法求最大值点和最大值。
5. 二值因变量模型参数的极大似然估计的精确分布并不知道，只知道渐近服从正态分布。以此为基础构造的参数显著性检验的统计量，其精确分布未知，只能采用渐近分布进行显著性检验，检验统计量不服从 t 分布，不是 t-统计量。
6. 完全分离是指存在自变量的一个线性组合 $\beta_0+\beta_1X_1+\beta_2X_2$ 和一个临界值 C，将样本数据分为两个子集 $\mathrm{A}_1=\{i\,|\,Y_i=1,\ \beta_0+\beta_1X_{1i}+\beta_2X_{2i}>C\}$ 和 $\mathrm{A}_2=\{i\,|\,Y_i=0,\ \beta_0+\beta_1X_{1i}+\beta_2X_{2i}\ge C\}$。样本存在完全分离时，二值因变量模型无法估计，或者估计结果不可信。分析造成完全分离的原因，通过去掉一些样本或者去掉一些自变量来消除完全分离。
7. 二值因变量模型是非线性模型，自变量变化对因变量取 1 概率的影响不是一个常数，而是与自变量所有样本值有关。可以选择自变量样本均值带入进行计算，也可以计算自变量对概率的相对影响。
8. 略。

9* （1）首先证明 $F(x)$ 满足分布函数的性质：（i）单调递增；（ii）$\lim_{x\to-\infty}F(x)=0$，$\lim_{x\to+\infty}F(x)=1$。由于 $G(x)$ 和 $\Lambda(x)$ 为分布函数，满足性质（i）和（ii）根据单调函数的性质和极限运算法则，可以证明 $F(x)$ 满足（i）和（ii），是分布函数。由于 $G(x)$ 和 $\Lambda(x)$ 满足对称性，因此

$$
\begin{aligned}
F(-x)&=\lambda\Phi(-x)+(1-\lambda)\Lambda(-x)\\
&=\lambda[1-\Phi(x)]+(1-\lambda)[1-\Lambda(x)]\\
&=1-[\lambda\Phi(x)+(1-\lambda)\Lambda(x)]\\
&=1-F(x)
\end{aligned}
$$

（2）将 $F(x)$ 带入（8.9）得出

$$
\begin{aligned}
l=&\sum_{i=1}^{N}Y_i\ln[\lambda\Phi(\beta_0+\beta_1X_{1i}+\beta_2X_{2i})+(1-\lambda)\Lambda(\beta_0+\beta_1X_{1i}+\beta_2X_{2i})]\\
&+\sum_{i=1}^{N}(1-Y_i)\ln\{1-[\lambda\Phi(\beta_0+\beta_1X_{1i}+\beta_2X_{2i})+(1-\lambda)\Lambda(\beta_0+\beta_1X_{1i}+\beta_2X_{2i})]\}
\end{aligned}
$$