9 Binary Dependent Variable (y variable): Logit and Probit Models

Reading: _____

- Many research questions on consumer choice involve analyzing qualitative dependent variables such as

- Using the linear probability model (LPM) is unrealistic because the model can result in predicted probability greater or less than one.

Can result in

$$P(y = 1|\boldsymbol{x}) > \mathbf{1} \text{ or } P(y = 1|\boldsymbol{x}) < \mathbf{0}$$

- The Logit and Probit model introduce a function (call it "G") to restrict the probability value between zero and one.
    - Logit Model – Function G(•) is the logistic function

- Probit model - Function G(•) is the standard normal cumulative distribution function (cdf).

NOW, the β in the logit and probit models are not marginal impact on probability!!To find the partial effect of each X variable on the response probability, we must rely on calculation. To find the partial impact of X_2, we calculate the following difference:

- STATA provides a post-estimation command called "mfx" to calculate such marginal effects. For instance:

```
. probit eat_rte gender age income product price place promotion

Iteration 0:   log likelihood = -99.699949
Iteration 1:   log likelihood =  -63.03342
Iteration 2:   log likelihood = -59.735085
Iteration 3:   log likelihood =  -59.68356
Iteration 4:   log likelihood = -59.683526
Iteration 5:   log likelihood = -59.683526

Probit regression                                 Number of obs   =        252
                                                  LR chi2(7)      =      80.03
                                                  Prob > chi2     =     0.0000
Log likelihood = -59.683526                       Pseudo R2       =     0.4014
```

eat_rte	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
gender	.3639538	.3001825	1.21	0.225	-.2243931	.9523007
age	-.035382	.0130412	-2.71	0.007	-.0609422	-.0098218
income	-.0000126	8.44e-06	-1.49	0.137	-.0000291	3.99e-06
product	.4826703	.1645111	2.93	0.003	.1602344	.8051062
price	-.291123	.1639812	-1.78	0.076	-.6125202	.0302742
place	1.128479	.2874352	3.93	0.000	.5651166	1.691842
promotion	.0168971	.1699563	0.10	0.921	-.3162112	.3500054
_cons	-2.638742	1.207165	-2.19	0.029	-5.004742	-.2727424

```
. mfx

Marginal effects after probit
      y  = Pr(eat_rte) (predict)
         =  .95502682
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
gender*	.0364636	.03047	1.20	0.231	-.023265	.096193	.575397
age	-.0033521	.00142	-2.36	0.018	-.006136	-.000569	27.7976
income	-1.19e-06	.00000	-1.44	0.151	-2.8e-06	4.3e-07	22501
product	.0457288	.01681	2.72	0.007	.012777	.07868	3.36984
price	-.0275814	.01534	-1.80	0.072	-.057638	.002476	3.55952
place	.1069136	.03163	3.38	0.001	.044928	.1689	4.20933
promot~n	.0016009	.01615	0.10	0.921	-.030062	.033263	2.99524

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```