Homework 10
QMB 3200: Advanced and Quantitative Methods
Fall 2019

Probit and Logistic Regression Analysis

Submitted to
Dr. Jim Dewey
Florida Polytechnic University

Submitted by
Luiz Gustavo Fagundes Malpele
Department of Data Science
Florida Polytechnic University

December 2nd, 2019

1. Robust-Poisson regression model selection:

The first step to find the best regression model is to analyze the variables correlation and exclude those that account for high degree of collinearity. Using the correlation table below as a reference (1), it can be said that the combinations days and cold (r = 0.9685), and weight and female (r = - 0.7656) have high degree of correlation since r is higher than 0.5. This way, the variables days and weight will not be used on the final regression model.

Correlation table for cold, days, weight, female, and supplement (1)

	cold	days	weight	female	supple~t
cold	1.0000				
days	0.9685	1.0000			
weight	0.0800	0.0235	1.0000		
female	-0.1333	-0.0657	-0.7656	1.0000	
supplement	-0.2667	-0.3365	0.1197	0.0000	1.0000

Since the variable weight is being excluded from the analysis, other related variable was like squared weight, and weight * supplement was created, but these still presented high degree of collinearity, so they are also not being used on this regression model. Furthermore, another correlation table was created, this one accounted only for the variables that are going to be used on the regression model:

Correlation table for cold, female, and supplement (2)

(obs=60)

	cold	female	supple~t
cold	1.0000		
female	-0.1333	1.0000	
supplement	-0.2667	0.0000	1.0000

After knowing which variables presented high degree of collinearity, the next step was to compare model which contained the predictors and choose the one that would best contain that would best predict the variable cold, so Leave-One-Out Cross-Validation was performed with three different Robust-Logit regression were performed, and the results can be observer on table (3):

1. Robust-Logistics regression of days on supplement, and female;
2. Robust-Logistics regression of days on supplement, and weight;
3. Robust-Logistics regression of days on supplement, female, and weight;

RMSE value comparison for 3 different Poisson regression models (3)

	(1)	(2)	(3)
	cold	cold	cold
cold			
0.female	0		0
	(.)		(.)
1.female	-0.579		-0.497
	(0.547)		(0.842)
0.supplement	0	0	0
	(.)	(.)	(.)
1.supplement	-1.116*	-1.168*	-1.129*
	(0.547)	(0.552)	(0.555)
weight		0.00735	0.00159
		(0.008)	(0.012)
_cons	0.847	-0.554	0.566
	(0.476)	(1.248)	(2.211)
RMSE	0.5024	0.5036	0.5111
N	60	60	60

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

After running all regression, it was possible to compare the RMSE. There is not such a statistically significant difference between the RMSE values, which indicate low variance between the three models, so the chosen one is the first model which do not account for the weight variable as it presents high degree of collinearity with the female variable and it is less significant than female for this analysis.

The Logistic regression below refers to model 1, which account for female and supplement variables, but it is still necessary to check the statistical significance of the variables, and to do so, the Chi-Square test will be used.

Robust-Logistics Regression of cold on supplement and female (4)

```
Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.859982
Iteration 2:   log pseudolikelihood = -38.852883
Iteration 3:   log pseudolikelihood = -38.852882

Logistic regression                             Number of obs     =         60
                                                Wald chi2(2)      =       4.97
                                                Prob > chi2       =     0.0834
Log pseudolikelihood = -38.852882               Pseudo R2         =     0.0658
```

cold	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.female	-.5790338	.5472366	-1.06	0.290	-1.651598	.4935302
1.supplement	-1.115562	.5472366	-2.04	0.041	-2.188126	-.0429978
_cons	.8472978	.4757966	1.78	0.075	-.0852464	1.779842

The p-values for supplement ($\widehat{\beta_1} = 0.041$) which is below 0.05, so it is statistically significant for this analysis, but female ($\widehat{\beta_2} = 0.29$) which is above 0.05, but since these two variables interact with each other, the Chi-Square test is the most appropriate to measure the statistical significance of them:

Chi-Square tests to evaluate the statistical significance of female and supplement (5)

```
 ( 1)  [cold]1.female = 0
 ( 2)  [cold]1.supplement = 0

           chi2(  2) =    4.97
         Prob > chi2 =    0.0834
```

While conducing these tests and assuming a conservative α = 0.1, it can be said that there is enough evidence to say that the combination of supplement and female have a high degree statistical significance to the Logistic Regression model. This way, the regression model on table (4) is the chose one for this analysis.

By using the same parameters and testing for the Probit-Regression model of cold on female and supplement, the generate result will be like the Logistics Regression:

Robust-Logistics Regression of cold on supplement and female (6)

```
Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.857494
Iteration 2:   log pseudolikelihood = -38.852842
Iteration 3:   log pseudolikelihood = -38.852842

Probit regression                               Number of obs     =         60
                                                Wald chi2(2)      =       5.20
                                                Prob > chi2       =     0.0744
Log pseudolikelihood = -38.852842               Pseudo R2         =     0.0658
```

cold	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.female	-.3565321	.3355987	-1.06	0.288	-1.014294	.3012293
1.supplement	-.6923088	.3356543	-2.06	0.039	-1.350179	-.0344384
_cons	.5259069	.2914152	1.80	0.071	-.0452564	1.09707

RMSE value comparison for 3 different Poisson regression models (7)

	(1) cold	(2) cold
cold		
0.female	0	0
	(.)	(.)
1.female	-0.357	-0.579
	(0.336)	(0.547)
0.supplement	0	0
	(.)	(.)
1.supplement	-0.692*	-1.116*
	(0.336)	(0.547)
_cons	0.526	0.847
	(0.291)	(0.476)
RMSE	0.50238	0.50237
N	60	60

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

There is not such a statistically significant difference between the RMSE values of the Probit (1) and Logistics-Regression (2), which indicate low variance between the three models. It is necessary to investigate other factors in order to choose which predicts better the effects of female and supplement on the response variable cold.

Robust-Probit Regression of cold on supplement and female (8)

```
Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.857494
Iteration 2:   log pseudolikelihood = -38.852842
Iteration 3:   log pseudolikelihood = -38.852842

Probit regression                               Number of obs     =         60
                                                Wald chi2(2)      =       5.20
                                                Prob > chi2       =     0.0744
Log pseudolikelihood = -38.852842               Pseudo R2         =     0.0658
```

cold	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.female	-.3565321	.3355987	-1.06	0.288	-1.014294	.3012293
1.supplement	-.6923088	.3356543	-2.06	0.039	-1.350179	-.0344384
_cons	.5259069	.2914152	1.80	0.071	-.0452564	1.09707

Chi-Square tests to evaluate the statistical significance of female and supplement (9)

```
 ( 1)  [cold]1.female = 0
 ( 2)  [cold]1.supplement = 0

           chi2(  2) =    5.20
         Prob > chi2 =    0.0744
```

While conducting again the Chi-Square test and assuming a conservative $\alpha = 0.1$, it can be said that there is enough evidence to say that the combination of supplement and female have a high degree statistical significance to the Probit Regression model. Lastly, the p-value for the Probit model is lower than for the Logistics-Regression, so it has a higher degree of significance, and thus, the chose model while comparing both is the Robust-Probit regression of cold on female and supplement.

2. Marginal effects of gender and weight

Marginal effect of female, supplement, and weigh on cold (10)

```
Average marginal effects                        Number of obs     =         60
Model VCE    : Robust

Expression   : Pr(cold), predict()
dy/dx w.r.t. : 1.female 1.supplement weight

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0.female     |  (base outcome)
-------------+----------------------------------------------------------------
1.female     |
      female |
          0  |  -.1135985   .1939826    -0.59   0.558    -.4937974    .2666004
          1  |  -.1135985   .1939826    -0.59   0.558    -.4937974    .2666004
-------------+----------------------------------------------------------------
0.supplement |  (base outcome)
-------------+----------------------------------------------------------------
1.supplement |
      female |
          0  |  -.2707721   .1278128    -2.12   0.034    -.5212806   -.0202637
          1  |  -.2708964   .1276789    -2.12   0.034    -.5211425   -.0206503
-------------+----------------------------------------------------------------
weight       |
      female |
          0  |   .0003856    .002845     0.14   0.892    -.0051904    .0059616
          1  |   .0003858   .0028459     0.14   0.892    -.0051922    .0059637
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

For the sake of this specific analysis, the variable weight was also included. The table above shows the marginal effect of each variable on cold, in other words, how unit-change on X_i impacts the probability of getting a cold.

The table indicate that if the subject's gender is female, the probability of getting a cold decrease in 0.1136 if the subject is a woman. Furthermore, the marginal effects of weight on cold indicates that for each pound a person gains, the probability of getting a cold increase in 0.0003856 if the subject is man and 0.0003858 if the subject is a woman. Lastly, if the person is taking the experimental supplement, the probability of getting a cold decrease in 0.27077 if the subject is a man, while 0.27090 if the subject is a woman.

The weight that will be used for the marginal effect prediction model is the subject's weight, but the above information can be used to calculate the average effect of the supplement on men and women, so, for this case the mean wean weight of men and women can be calculated in order to predict the average effect of the supplement for both genders.

3. Models comparison

Robust-Probit Regression of cold on supplement and female (11)

```
Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.857494
Iteration 2:   log pseudolikelihood = -38.852842
Iteration 3:   log pseudolikelihood = -38.852842

Probit regression                               Number of obs     =         60
                                                Wald chi2(2)      =       5.20
                                                Prob > chi2       =     0.0744
Log pseudolikelihood = -38.852842               Pseudo R2         =     0.0658
```

cold	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.female	-.3565321	.3355987	-1.06	0.288	-1.014294	.3012293
1.supplement	-.6923088	.3356543	-2.06	0.039	-1.350179	-.0344384
_cons	.5259069	.2914152	1.80	0.071	-.0452564	1.09707

Robust-Poisson Regression of female, and supplement on cold (12)

```
Iteration 0:   log pseudolikelihood =  -49.44727
Iteration 1:   log pseudolikelihood =  -49.44727

Poisson regression                              Number of obs     =         60
                                                Wald chi2(2)      =       5.09
                                                Prob > chi2       =     0.0785
Log pseudolikelihood =  -49.44727               Pseudo R2         =     0.0265
```

cold	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.female	-.268264	.2560906	-1.05	0.295	-.7701923	.2336644
1.supplement	-.5465437	.2777286	-1.97	0.049	-1.090882	-.0022057
_cons	-.3315953	.1606087	-2.06	0.039	-.6463826	-.0168079

The Probit-Regression (11) model provides a more reasonable approach while handling binary data, so that is the reason why these are the best ones to predict the best fit for the probability of getting a cold given that information such as weight, gender, and supplement is available. The Poisson-robust regression model would not be as good since it provides a best fit for a dataset that involves rate data, which is not the case of this research, since response variable is not cold duration.

Furthermore, it can be observer that the statistical significance of the Probit Regression is higher than the Poisson Regression, since the Prob > Chi-Square = 0.0785, while for the Probit-Regression the Prob > Chi-Square = 0.0744. There was no significance difference on the RMSE value for both models to use it as a parameter of comparison.

In conclusion, the Probit-Regression and the Logistic-Regression models have are better when it comes to deal with dichotomous data, while the Poisson regression is better for count data, in other word, data that have been collected within a fixed period of time.

Appendix A: Do-file-for-Homework 10

```
/* QMB 3200 Homework 10 */

log using "C:\Users\luizg\Desktop\Stata\hw10.smcl"
import delimited "C:\Users\luizg\Desktop\Stata\supplement.csv"

*Question 1
correlate cold days weight female supplement
correlate cold female supplement
logit cold i.female i.supplement, robust
testparm i.female i.supplement
estimate store modell1
logit cold c.weight i.supplement, robust
estimate store modell2
logit cold c.weight i.supplement i.female, robust
estimate store modell3
loocv logit cold i.female i.supplement, robust
loocv logit cold c.weight i.supplement, robust
loocv logit cold c.weight i.supplement i.female, robust
esttab modell* using "modell2.rtf" , se(3) replace
logit cold i.female i.supplement c.female#supplement, robust

probit cold i.female i.supplement, robust
testparm i.female i.supplement
estimate store modelt1
logit cold i.female i.supplement, robust
testparm i.female i.supplement
estimate store modelt2
loocv probit cold i.female i.supplement, robust
loocv logit cold i.female i.supplement, robust
esttab modelt* using "modelt.rtf" , se(3) replace
```

```
*Question 2

probit cold i.female i.supplement c.weight, robust
margins female, dydx(*) post

*Question 3

probit cold i.female i.supplement, robust
estimate store modelr1
poisson cold i.female i.supplement, robust
estimate store modelr2
loocv probit cold i.female i.supplement, robust
loocv poisson cold i.female i.supplement, robust
esttab modelr* using "modelr.rtf" , se(3) replace

STOP

log close
```

Appendix B :Do-file-for-Homework 10

```
  ___  ____  ____  ____  ____ (R)
 /__    /   ____/   /   ____/
___/   /   /___/   /   /___/   16.0   Copyright 1985-2019 StataCorp LLC
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
                                      College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

Single-user Stata license expires 16 Mar 2020:
       Serial number:  301609236389
         Licensed to:  Luiz Gustavo Fagundes Malpele
                       Florida Polytechnic University

Notes:
      1.  Unicode is supported; see help unicode_advice.

. do "C:\Users\luizg\AppData\Local\Temp\STD6560_000000.tmp"

. /* QMB 3200 Homework 10 */
.
. log using "C:\Users\luizg\Desktop\Stata\hw10.smcl"
file C:\Users\luizg\Desktop\Stata\hw10.smcl already exists
r(602);

end of do-file

r(602);
```

```
. do "C:\Users\luizg\AppData\Local\Temp\STD6560_000000.tmp"

. /* QMB 3200 Homework 10 */
.
. log using "C:\Users\luizg\Desktop\Stata\hw10new.smcl"
------------------------------------------------------------------------------------
------------------------------------------------------------------------------
      name:  <unnamed>
       log:  C:\Users\luizg\Desktop\Stata\hw10new.smcl
  log type:  smcl
 opened on:   2 Dec 2019, 01:39:31

. import delimited "C:\Users\luizg\Desktop\Stata\supplement.csv"
(6 vars, 60 obs)

.
. *Question 1
. correlate cold days weight female supplement
(obs=60)

             |     cold     days   weight   female supple~t
-------------+---------------------------------------------
        cold |   1.0000
        days |   0.9685   1.0000
      weight |   0.0800   0.0235   1.0000
      female |  -0.1333  -0.0657  -0.7656   1.0000
  supplement |  -0.2667  -0.3365   0.1197   0.0000   1.0000


. correlate cold female supplement
(obs=60)

             |     cold   female supple~t
-------------+---------------------------
```

```
        cold |   1.0000
      female |  -0.1333   1.0000
  supplement |  -0.2667   0.0000   1.0000


. logit cold i.female i.supplement, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.859982
Iteration 2:   log pseudolikelihood = -38.852883
Iteration 3:   log pseudolikelihood = -38.852882

Logistic regression                               Number of obs     =         60
                                                  Wald chi2(2)      =       4.97
                                                  Prob > chi2       =     0.0834
Log pseudolikelihood = -38.852882                 Pseudo R2         =     0.0658

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |  -.5790338   .5472366    -1.06   0.290    -1.651598    .4935302
1.supplement |  -1.115562   .5472366    -2.04   0.041    -2.188126   -.0429978
       _cons |   .8472978   .4757966     1.78   0.075    -.0852464    1.779842
------------------------------------------------------------------------------

. testparm i.female i.supplement

 ( 1)  [cold]1.female = 0
 ( 2)  [cold]1.supplement = 0

           chi2(  2) =    4.97
         Prob > chi2 =    0.0834
```

```
. estimate store modell1

. logit cold c.weight i.supplement, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -39.022879
Iteration 2:   log pseudolikelihood = -39.017653
Iteration 3:   log pseudolikelihood = -39.017653

Logistic regression                             Number of obs     =         60
                                                Wald chi2(2)      =       4.84
                                                Prob > chi2       =     0.0891
Log pseudolikelihood = -39.017653               Pseudo R2         =     0.0618

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      weight |   .0073456   .0078945     0.93   0.352    -.0081274    .0228186
1.supplement |  -1.167747   .5515119    -2.12   0.034     -2.24869   -.0868033
       _cons |   -.554055   1.247558    -0.44   0.657    -2.999224    1.891114
------------------------------------------------------------------------------

. estimate store modell2

. logit cold c.weight i.supplement i.female, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.852195
Iteration 2:   log pseudolikelihood = -38.844996
Iteration 3:   log pseudolikelihood = -38.844996
```

```
Logistic regression                             Number of obs     =         60
                                                Wald chi2(3)      =       5.00
                                                Prob > chi2       =     0.1716
Log pseudolikelihood = -38.844996               Pseudo R2         =     0.0660

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      weight |   .0015939   .0121767     0.13   0.896    -.0222721    .0254599
1.supplement |  -1.128648    .555406    -2.03   0.042    -2.217224   -.0400722
    1.female |  -.4972672   .8417202    -0.59   0.555    -2.147009    1.152474
       _cons |   .5661651   2.211146     0.26   0.798    -3.767602    4.899933
------------------------------------------------------------------------------

. estimate store modell3

. loocv logit cold i.female i.supplement, robust


 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50236573
Mean Absolute Errors     |   .47954825
Pseudo-R2                |   .01832682
-----------------------------------------

. loocv logit cold c.weight i.supplement, robust


 Leave-One-Out Cross-Validation Results
```

```
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50361447
Mean Absolute Errors     |   .48225649
Pseudo-R2                |   .0147318
-----------------------------------------

. loocv logit cold c.weight i.supplement i.female, robust



 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .51106126
Mean Absolute Errors     |   .48802338
Pseudo-R2                |   .00619336
-----------------------------------------

. esttab modell* using "modell2.rtf" , se(3) replace
(output written to modell2.rtf)

. logit cold i.female i.supplement c.female#supplement, robust

note: 1.supplement#c.female omitted because of collinearity
Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.715784
Iteration 2:   log pseudolikelihood = -38.705467
Iteration 3:   log pseudolikelihood = -38.705465

Logistic regression                               Number of obs     =         60
                                                  Wald chi2(3)      =       5.17
```

```
                                                  Prob > chi2       =     0.1599
Log pseudolikelihood = -38.705465                 Pseudo R2         =     0.0693

------------------------------------------------------------------------------------
                    |               Robust
               cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
--------------------+---------------------------------------------------------------
           1.female |  -.2876821   .7665316    -0.38   0.707    -1.790056
1.214692
       1.supplement |  -1.417066   .7932042    -1.79   0.074    -2.971717
.1375861
                    |
supplement#c.female |
                 0  |   -.590387   1.098478    -0.54   0.591    -2.743365
1.562591
                 1  |          0  (omitted)
                    |
              _cons |     1.0116   .5888014     1.72   0.086     -.142429
2.16563
------------------------------------------------------------------------------------

.
. probit cold i.female i.supplement, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.857494
Iteration 2:   log pseudolikelihood = -38.852842
Iteration 3:   log pseudolikelihood = -38.852842

Probit regression                                 Number of obs     =         60
                                                  Wald chi2(2)      =       5.20
                                                  Prob > chi2       =     0.0744
```

```
Log pseudolikelihood = -38.852842             Pseudo R2         =     0.0658

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |  -.3565321   .3355987    -1.06   0.288    -1.014294    .3012293
1.supplement |  -.6923088   .3356543    -2.06   0.039    -1.350179   -.0344384
       _cons |   .5259069   .2914152     1.80   0.071    -.0452564     1.09707
------------------------------------------------------------------------------

. testparm i.female i.supplement

 ( 1)  [cold]1.female = 0
 ( 2)  [cold]1.supplement = 0

           chi2(  2) =    5.20
         Prob > chi2 =    0.0744

. estimate store modelt1

. logit cold i.female i.supplement, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.859982
Iteration 2:   log pseudolikelihood = -38.852883
Iteration 3:   log pseudolikelihood = -38.852882

Logistic regression                             Number of obs     =         60
                                                Wald chi2(2)      =       4.97
                                                Prob > chi2       =     0.0834
Log pseudolikelihood = -38.852882               Pseudo R2         =     0.0658
```

```
------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |  -.5790338   .5472366    -1.06   0.290    -1.651598    .4935302
1.supplement |  -1.115562   .5472366    -2.04   0.041    -2.188126   -.0429978
       _cons |   .8472978   .4757966     1.78   0.075    -.0852464    1.779842
------------------------------------------------------------------------------

. testparm i.female i.supplement

 ( 1)  [cold]1.female = 0
 ( 2)  [cold]1.supplement = 0

           chi2(  2) =    4.97
         Prob > chi2 =    0.0834

. estimate store modelt2

. loocv probit cold i.female i.supplement, robust


 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50238303
Mean Absolute Errors     |   .47954802
Pseudo-R2                |   .01831338
-----------------------------------------

. loocv logit cold i.female i.supplement, robust
```

```
 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50236573
Mean Absolute Errors     |   .47954825
Pseudo-R2                |   .01832682
-----------------------------------------

. esttab modelt* using "modelt.rtf" , se(3) replace
(output written to modelt.rtf)

.
. *Question 2
.
. probit cold i.female i.supplement c.weight, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.849082
Iteration 2:   log pseudolikelihood = -38.844223
Iteration 3:   log pseudolikelihood = -38.844223

Probit regression                               Number of obs     =         60
                                                Wald chi2(3)      =       5.23
                                                Prob > chi2       =     0.1557
Log pseudolikelihood = -38.844223               Pseudo R2         =     0.0660

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |  -.3035246   .5202637    -0.58   0.560    -1.323223    .7161735
```

```
1.supplement |  -.7009031   .3410033    -2.06   0.040    -1.369257   -.0325489
      weight |   .0010375   .0076341     0.14   0.892    -.0139252    .0160001
       _cons |    .343039   1.382863     0.25   0.804    -2.367322      3.0534
------------------------------------------------------------------------------

. margins female, dydx(*) post

Average marginal effects                        Number of obs     =         60
Model VCE    : Robust

Expression   : Pr(cold), predict()
dy/dx w.r.t. : 1.female 1.supplement weight

-------------------------------------------------------------------------------
              |            Delta-method
              |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
0.female      |  (base outcome)
--------------+----------------------------------------------------------------
1.female      |
       female |
           0  |  -.1135985   .1939826    -0.59   0.558    -.4937974    .2666004
           1  |  -.1135985   .1939826    -0.59   0.558    -.4937974    .2666004
--------------+----------------------------------------------------------------
0.supplement  |  (base outcome)
--------------+----------------------------------------------------------------
1.supplement  |
       female |
           0  |  -.2707721   .1278128    -2.12   0.034    -.5212806   -.0202637
           1  |  -.2708964   .1276789    -2.12   0.034    -.5211425   -.0206503
--------------+----------------------------------------------------------------
weight        |
       female |
```

```
          0  |   .0003856    .002845     0.14   0.892    -.0051904    .0059616
          1  |   .0003858   .0028459     0.14   0.892    -.0051922    .0059637
--------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.

.
. *Question 3
.
. probit cold i.female i.supplement, robust

Iteration 0:   log pseudolikelihood = -41.588831
Iteration 1:   log pseudolikelihood = -38.857494
Iteration 2:   log pseudolikelihood = -38.852842
Iteration 3:   log pseudolikelihood = -38.852842

Probit regression                               Number of obs     =         60
                                                Wald chi2(2)      =       5.20
                                                Prob > chi2       =     0.0744
Log pseudolikelihood = -38.852842               Pseudo R2         =     0.0658

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |  -.3565321   .3355987    -1.06   0.288    -1.014294    .3012293
1.supplement |  -.6923088   .3356543    -2.06   0.039    -1.350179   -.0344384
       _cons |   .5259069   .2914152     1.80   0.071    -.0452564     1.09707
------------------------------------------------------------------------------

. estimate store modelr1

. poisson cold i.female i.supplement, robust
```

```
Iteration 0:   log pseudolikelihood =  -49.44727
Iteration 1:   log pseudolikelihood =  -49.44727

Poisson regression                              Number of obs     =         60
                                                Wald chi2(2)      =       5.09
                                                Prob > chi2       =     0.0785
Log pseudolikelihood =  -49.44727               Pseudo R2         =     0.0265

------------------------------------------------------------------------------
             |               Robust
        cold |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    1.female |   -.268264   .2560906    -1.05   0.295    -.7701923    .2336644
1.supplement |  -.5465437   .2777286    -1.97   0.049    -1.090882   -.0022057
       _cons |  -.3315953   .1606087    -2.06   0.039    -.6463826   -.0168079
------------------------------------------------------------------------------

. estimate store modelr2

. loocv probit cold i.female i.supplement, robust


 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50238303
Mean Absolute Errors     |   .47954802
Pseudo-R2                |   .01831338
-----------------------------------------

. loocv poisson cold i.female i.supplement, robust
```

```
 Leave-One-Out Cross-Validation Results
-----------------------------------------
         Method          |    Value
-------------------------+---------------
Root Mean Squared Errors |   .50134207
Mean Absolute Errors     |   .47823749
Pseudo-R2                |   .02049683
-----------------------------------------


. esttab modelr* using "modelr.rtf" , se(3) replace
(output written to modelr.rtf)

.
. STOP
command STOP is unrecognized
r(199);

end of do-file

r(199);

.
```