Using Binary Logistic Regression and Binary Probit Regression to Test the Risk factors of Suicide

Abstract

Suicide is a wide public health problem. Using the interview dataset from The National Longitudinal Study of Adolescent to Adult Health (Add Health) Wave III, we aim to find the risk factors of suicide. By generating binary logistic regression and binary probit regression model, we find that age, if have depression, friend suicide, family member suicide, are significant predictors at predicting suicide intention. Specifically, every one-year increase in age decreases the probability of suicide intention by 0.005 ($p<.01$), diagnosis with depression increases the probability suicide intention of by 0.154 ($p<.001$), having friend suicide increases the probability suicide intention of by 0.092 ($p<.001$).

Background

In 2015, 42,773 in the United States died from intentional self-harm (suicide). The age-adjusted suicide rate in 2014 is 13.0 per 100,000 population. And the suicide rates have increased steadily from 1999 through 2014. (Steele, I. H., et al, 2017)

No single factor can account for the behavioral complexities of suicide. Suicide has been as an important public health issue involving psychological, biological, and societal factors. There is previous research done to seek the risk factors of suicide. (Phillips, M. R., et al, 2002) Some significant risk factors include high depression symptom score, low quality of life, high chronic stress, a blood relative with previous suicidal behavior, etc. (Mościcki, E. K., 1997)

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents, and consists of four interview waves. The interview questions contain respondents' social, economic, psychological status, and are

comprehensive and detailed. In this study, we used the dataset of Add Health Wave III. (Add Health., 2017)

The dependent variable of this study is defined as if ever seriously thought about committing suicide in the past 12 months. The independent variables are defined as age, if have bachelor's degree, have ever diagnosed with depression or not, if have any of friends tried to kill themselves during the past 12 months, if have any of family members tried to kill themselves during the past 12 months, if get health insurance.

Since the dependent variable is a binary variable, and independent variables are binary categorical variable and continuous variable, it is suitable to conduct binary logistic regression and binary probit regression for this study. Besides, binary logistic regression and binary probit regression are easy application, the interpretation is also straightforward and user-friendly.

We aim to investigate the risk factors of suicide intention. Considering previous research, we hypothesis that for age range 18-28, increased age will decrease the probability of suicide intention, having depression will increase the probability of suicide intention, having friend or family member suicide will increase the probability of suicide intention.

Method Description

The interview dataset is from The National Longitudinal Study of Adolescent to Adult Health (Add Health) Wave III.

There were 15,170 respondents completed the Wave III interview. And the Wave III datasets have 4882 subjects. All the respondents were aged between 18 and 28 years old at the interview.

After deleting subjects containing missing values on all study variables, 2117 subjects are deleted and 2756 subjects are remained in the study. The variables about if get health insurance have the most missing values, which is 2042. Considering the large numbers of missing values of this variable, the model result of health insurance may not reflect its true relationship with suicide intention.

Generalized linear model (GLM) is a generalization of ordinary linear regression that allows for dependent variables that have error distribution model other than a normal distribution, such as standard logistic distribution.

Binary logistic regression is a special case with a Bernoulli sampling model and a logit link. The probit model is a special case with a Bernoulli sampling model and an inverse normal link.

Structural model for binary probit model:

$$\eta_i = \Phi^{-1}(p_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots \beta_p X_{pi}$$

Structural model for binary logistic regression:

$$\eta_i = \log\left(\frac{\varphi_i}{1-\varphi_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots \beta_p X_{pi}$$

Results

Descriptive statistics

Table 1. Participant characteristics by if have intention to commit suicide

	Suicide (n=163, 5.91%)		**No Suicide** (n=2593, 94.09%)		
	n	%	n	%	t or $\chi 2$
Age (Mean, SD)	21.79	SD1.87	22.46	SD1.82	$t=4.58^{***}$

If have bachelor's degree					$\chi 2=3.87^{*}$
Yes	146	89.57%	2172	83.76%	
No	17	10.43%	421	16.24%	
If ever diagnosed with depression					$\chi 2=137.46^{***}$
Yes	59	36.20%	210	8.10%	
No	104	63.80%	2383	91.90%	
If have friends suicide					$\chi 2=69.34^{***}$
Yes	36	22.09%	143	5.51%	
No	127	77.91%	2450	94.49%	
If have family suicide					$\chi 2=34.70^{***}$
Yes	16	9.88%	57	2.20%	
No	146	90.12%	2529	97.80%	
If get health insurance					$\chi 2=11.26^{**}$
Yes	91	55.83%	1776	68.49%	
No	72	44.17%	817	31.51%	

*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$

For all 2748 subjects involved in the study, 2593 subjects have no intention of suicide during last 12 months, while 163 subjects have intention of suicide during last 12 months. For the subjects have no intention of suicide during last 12 months, the mean of age is 22.46 and standard deviation is 1.82. For the subjects have intention of suicide during last 12 months, the mean of age is 21.79 and standard deviation is 1.87. The result of two-sample t-test indicate the mean age of subjects with intention of suicide and subjects without intention of suicide is statistically different ($t=4.58$, $p<.001$). The Pearson's chi-squared test also indicate that there is correlation between suicide intention and if have bachelor's degree, have ever diagnosed with depression or not, if have any of friends tried to kill themselves, if have any of family members tried to kill themselves, if get health insurance.

Running probit model

Table 2. Coef. of probit model

Characteristic		Coef.	exp(b)
Age		-0.061*	0.941
If have bachelor's degree			
	Yes vs No	-0.064	0.938
If ever diagnosed with depression			
	Yes vs No	0.967***	2.630
If have friends suicide			
	Yes vs No	0.701***	2.016
If have family suicide			
	Yes vs No	0.601***	1.824
If have health insurance			
	Yes vs No	-0.176*	0.839

* p<.05, ** p<.01, *** p<.001, two-tailed test
Results are based on a binary probit regression model

Results show that the probit model has an excellent fit to data, as the likelihood ratio chi-square for the model is 163.94(df=6), p<.001. This means that at least one predictor variable has a non-zero impact on the probability of having intention of suicide. There are approximately 13.31% of variation on the outcome variable that can be explained by the six predictors included in the probit model (pseudo $R^2 = 0.1331$).

Age, depression, friend suicide, family member suicide, health insurance are significant predictors at predicting suicide intention. Results show that holding other things be equal, every one-year increase in age decreases the probability of suicide intention (p<.05), diagnosed with depression increases the probability of suicide intention (p<.001), friend suicide increases the probability of suicide intention (p<.001), family member suicide increases the probability of suicide intention (p<.001), having health insurance decreases the probability of suicide intention (p<.05).

Running logit model

Table 3. Coef. of logit model

Characteristic		Coef.	exp(b)
Age		-0.130*	0.878
If have bachelor's degree			
	Yes vs No	-0.085	0.918
If ever diagnosed with depression			
	Yes vs No	1.839***	6.290
If have friends suicide			
	Yes vs No	1.308***	3.699
If have family suicide			
	Yes vs No	1.172***	3.228
If have health insurance			
	Yes vs No	-0.362*	0.696

* p<.05, ** p<.01, *** p<.001, two-tailed test
Results are based on a binary logit regression model

Results show that the logit model has an excellent fit to data, as the likelihood ratio chi-square for the model is 160.29(df=6), p<.001. This means that at least one predictor variable has a non-zero impact on the probability of having intention of suicide. There are approximately 13.02% of variation on the outcome variable that can be explained by the six predictors included in the logit model (pseudo $R^2 = 0.1302$).

For the participants in this study, having a diagnosis with depression increases the log odds of suicide intention by 1.839 (p<.001), having a friend suicide increases the log odds of suicide intention by 1.308 (p<.001), having a family suicide increases the log odds of suicide intention by 1.172 (p<.001),

Exp(B) is also odds ratio here. For the participants in this study, those with depression are 6.290 times more likely to have suicide intention, those having friend suicide are 3.699 times more likely to have suicide intention, those having family member suicide are 3.228 times more likely to have suicide intention. Those having health insurance are 30.4% less likely to have suicide intention.

Graph the predicted probabilities to show the impact of the age

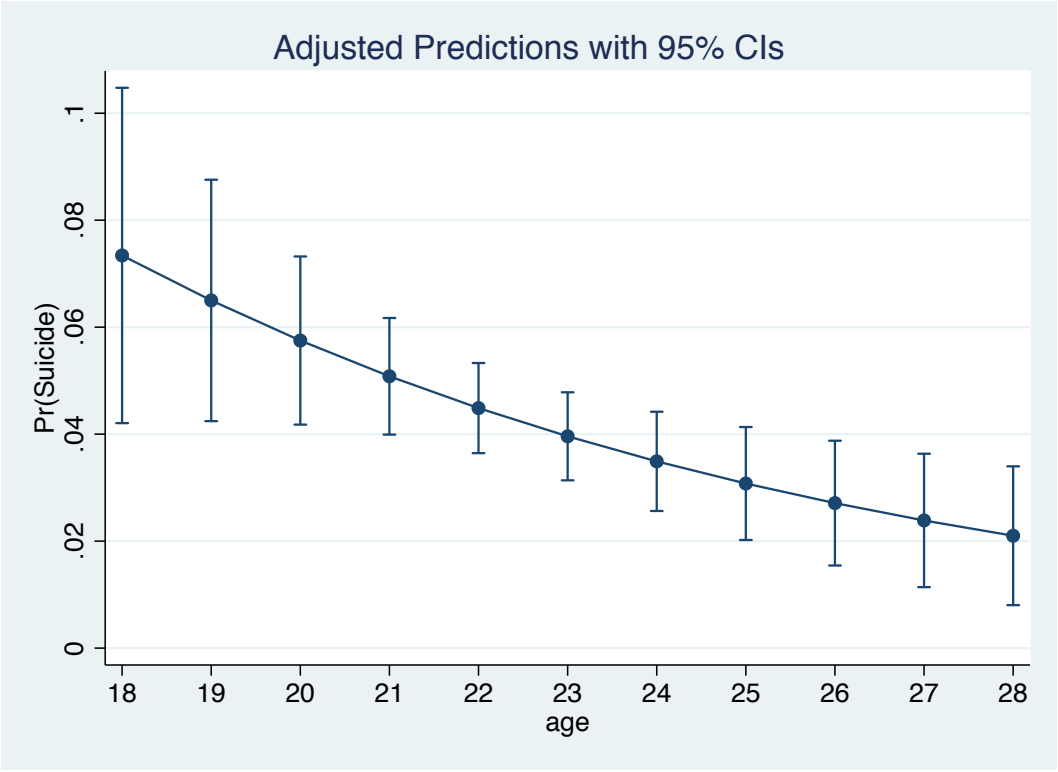

Fig 1. Predicted probabilities to show the impact of the age

Fig 1. shows the predicted probabilities to show the impact of the age in the logit model (Table 3.). While controlling for all other variables at the sample mean level, the predicted probabilities of suicide intention decreases as the age increases.

Graph the predicted probabilities to show the interaction

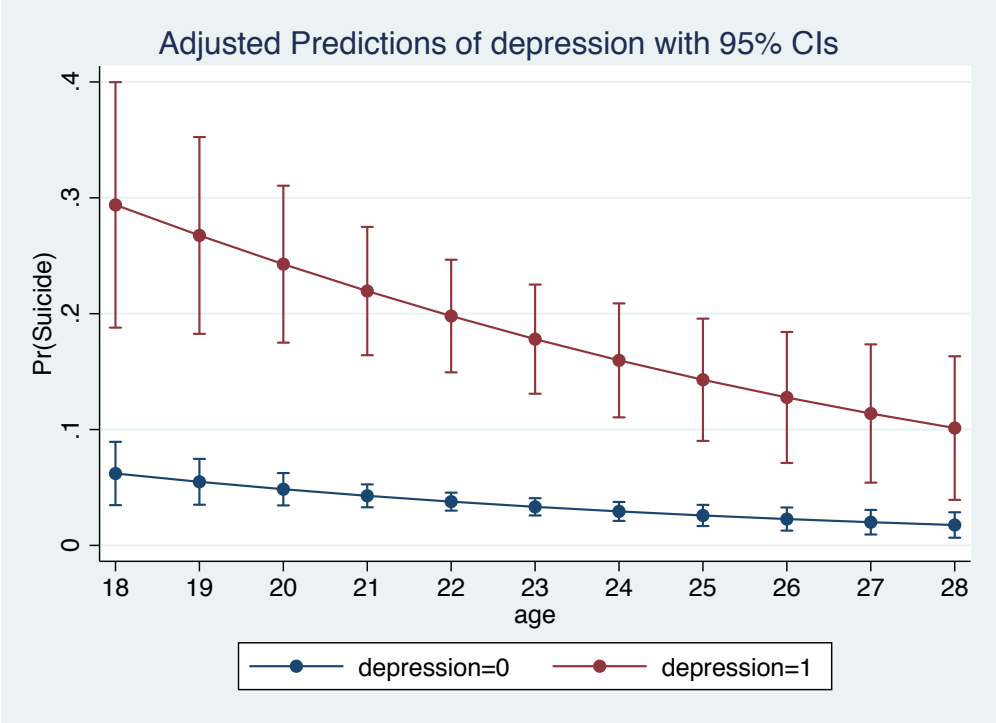

Fig 2. Predicted probabilities to show interaction between age and depression

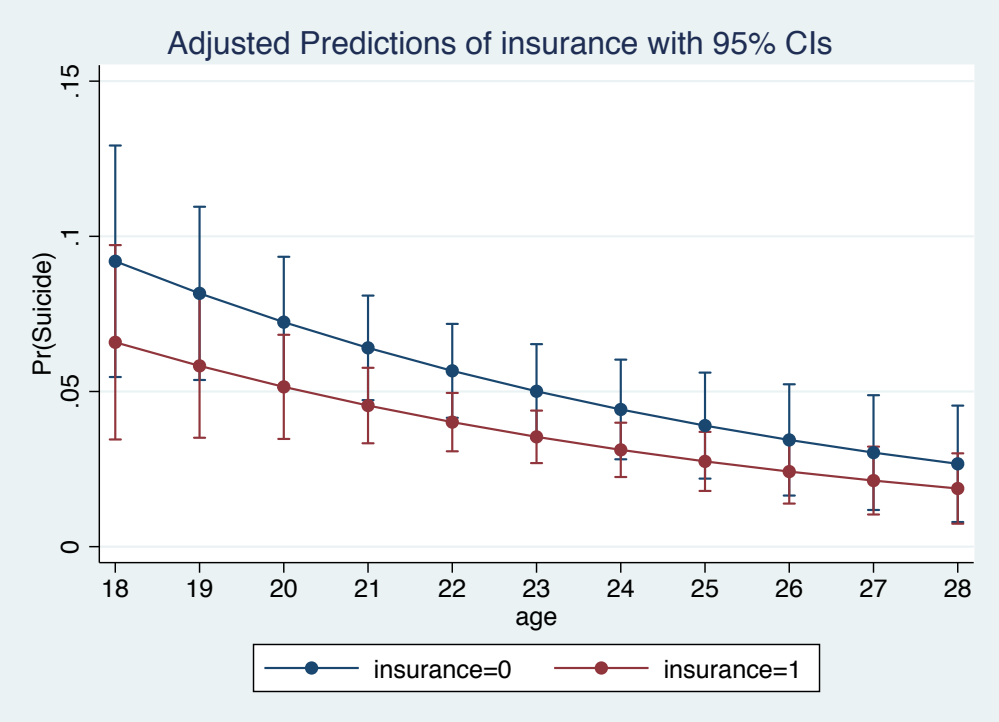

Fig 3. Predicted probabilities to show interaction between age and health insurance

Since there is no interaction in Figure 2., Figure 3., it indicates there is no interaction between age and depression, age and health insurance.

Test statistical significance for a group of variables for substantive importance

In the likelihood ratio test, I test the if friend suicide and family member suicide jointly are statistically significant. The likelihood ratio chi-square is 48.90(df=2), p<.001. Which means that the friend suicide and family member suicide as a whole are statistically significant in predicting suicide intention.

Produce marginal effects at the means (MEMs)

Table 4. Predicted Probabilities of Having Suicide Intention Using Marginal Effects at the Means

Characteristic		Coef.
Age		
	One year increase	-0.005**
Age		
	One standard deviation increase	-0.009**
If have bachelor's degree		
	Yes vs No	-0.003
If ever diagnosed with depression		
	Yes vs No	0.154***
If have friends suicide		
	Yes vs No	0.092***
If have family suicide		
	Yes vs No	0.081*
If have health insurance		
	Yes vs No	-0.016

* p<.05, ** p<.01, *** p<.001, two-tailed test
Results are based on a binary logit regression model

Using the marginal effects at the means and holding all other variables at a constant level, we find that:

- Every one-year increase in age decreases the probability of suicide intention by 0.005 ($p<.01$); equivalently, every one standard-deviation-unit increase in age (i.e., every 1.83-years increase) decreases the probability by 0.009 ($p<.01$);
- Diagnosis with depression increases the probability suicide intention of by 0.154 ($p<.001$);
- Having friend suicide increases the probability suicide intention of by 0.092 ($p<.001$);
- Having family member suicide increases the probability suicide intention of by 0.081 ($p<.05$).

Produce average marginal effects (AMEs)

Table 4. Predicted Probabilities of Having Suicide Intention Using Marginal Effects at the Means

Characteristic		Coef.
Age		
	One year increase	-0.009**
Age		
	One standard deviation increase	-0.006**
If have bachelor's degree		
	Yes vs No	-0.004
If ever diagnosed with depression		
	Yes vs No	0.159***
If have friends suicide		
	Yes vs No	0.100***
If have family suicide		
	Yes vs No	0.089*
If have health insurance		
	Yes vs No	-0.019

* p<.05, ** p<.01, *** p<.001, two-tailed test
Results are based on a binary logit regression model

Using the average marginal effects and holding all other variables at a constant level, we find that:

- Every one-year increase in age decreases the probability of suicide intention by 0.009 ($p<.01$); equivalently, every one standard-deviation-unit increase in age (i.e., every 1.83-years increase) decreases the probability by 0.006 ($p<.01$);
- Diagnosis with depression increases the probability suicide intention of by 0.159 ($p<.001$);
- Having friend suicide increases the probability suicide intention of by 0.100 ($p<.001$);
- Having family member suicide increases the probability suicide intention of by 0.089 ($p<.05$).

Produce marginal effects at representative values (MERs)

Table 5. Predicted Probabilities of Suicide Intention Using Hypothetical Scenarios of Substantively Illustrative Values

Scenerio	Pr	95% CI
age=18, have depression, have friend suicide	0.586	(0.435,0.737)
age=18, no depression, have friend suicide	0.184	(0.097,0.271)
age=18, have depression, no friend suicide	0.277	(0.174,0.280)
age=28, have depression, no friend suicide	0.094	(0.036,0.152)

The predicted probabilities for these hypothetical scenarios clearly show the joint impacts of age, depression, friend suicide on people's likelihood of having suicide intention. The worst case scenario is a 18-year-old person who have depression and friend suicide – the predicted probability of having suicide intention is 0.586, while the best case scenario is a 28-year-old person who have depression and no friend suicide.

Comparison between Binary Logistic model and Binary Probit model

The difference between Logistic and Probit models lies in this assumption about the distribution of the errors. The logit model assumes the distribution of the errors as standard logistic distribution

of errors. The probit model assumes the distribution of the errors as Normal distribution of errors. The significant variables in probit model are also significant in logit model.

Both of probit model and logit model are significant at predicting the suicide intention, and the model chi-square are very similar. The pseudo R^2 of both models are very similar (0.1331, 0.1302). Most of coefficients of probit model and logit model are also similar, and hold the same significance level. For the exp(b), the odds ratios for depression, friend suicide, family member suicide in logit model is bigger, while the odds ratios for age, bachelor degree, health insurance is smaller. In other words, the odds ratios of logit model are further away from 1.

Summary

Overall, this study tests the risk factors of suicide. Every one-year increase in age decreases the probability of suicide intention by 0.005 ($p<.01$). Diagnosis with depression, having friend suicide, having family member suicide increase the probability suicide intention of by 0.154 ($p<.001$), 0.092 ($p<.001$), ($p<.05$). However, the pseudo R^2 of logit model is 0.1302, which means approximately 13.02% of variation on the outcome variable that can be explained by the six predictors included in the logit model, which is still relatively low. Suicide is a complicated behavior, and more factors need to be included to generate a more accurate model.

The strength of this study is its easy application, because the model is simply a binary logistic model applied to a binary outcome variable. Running, interpreting, and presenting the model to readers are straightforward and user-friendly. Besides the sample is also very large, and consists of comprehensive longitudinal interview data.

The limitation of this binary logistic model is the residual term is not normally distributed, and predictions of the outcome-variable values may be out of the meaningful range of 0 and 1. The model also cannot show a nonlinear impact of independent variable on the probability of having

an event. Besides, considering the large numbers of missing values of health insurance, the model may not reflect its true relationship with suicide intention. The generalization to other age group is also a problem because all participants in this study are in age 18-28.

In conclude, we should be aware of those risk factors of suicide, such as depression, having friend's suicide, and seek public help when encountered those risk factors.

Reference

Steele, I. H., Thrower, N., Noroian, P., & Saleh, F. M. (2017). Understanding Suicide Across the Lifespan: A United States Perspective of Suicide Risk Factors, Assessment & Management. Journal of forensic sciences.

Mościcki, E. K. (1997). Identification of suicide risk factors using epidemiologic studies. Psychiatric Clinics of North America, 20(3), 499-517.

Phillips, M. R., Yang, G., Zhang, Y., Wang, L., Ji, H., & Zhou, M. (2002). Risk factors for suicide in China: a national case-control psychological autopsy study. The Lancet, 360(9347), 1728-1736.

Add Health. (2017, June 16). Retrieved December 15, 2017, from http://www.cpc.unc.edu/projects/addhealth

```
. ttest age, by(suicide)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	2,593	22.46433	.0356938	1.817583	22.39434	22.53432
1	163	21.79141	.1465227	1.870677	21.50207	22.08075
combined	2,756	22.42453	.0348078	1.827329	22.35628	22.49278
diff		.672916	.1470262		.3846233	.9612087

```
    diff = mean(0) - mean(1)                                      t =   4.5768
Ho: diff = 0                                     degrees of freedom =     2754

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000

.
. ta bachelor suicide, column chi2
```

Key
frequency *column percentage*

H3ED5	H3TO130 0	H3TO130 1	Total
0	2,172	146	2,318
	83.76	89.57	84.11
1	421	17	438
	16.24	10.43	15.89
Total	2,593	163	2,756
	100.00	100.00	100.00

```
          Pearson chi2(1) =   3.8683   Pr = 0.049
```

```
. ta depress suicide, column chi2
```

Key
frequency *column percentage*

H3ID15	H3TO130 0	H3TO130 1	Total
0	**2,383** **91.90**	**104** **63.80**	**2,487** **90.24**
1	**210** **8.10**	**59** **36.20**	**269** **9.76**
Total	**2,593** **100.00**	**163** **100.00**	**2,756** **100.00**

Pearson chi2(**1**) = **137.4610** Pr = **0.000**

```
. ta friend_sui suicide, column chi2
```

Key
frequency *column percentage*

H3TO133	H3TO130 0	H3TO130 1	Total
0	**2,450** **94.49**	**127** **77.91**	**2,577** **93.51**
1	**143** **5.51**	**36** **22.09**	**179** **6.49**
Total	**2,593** **100.00**	**163** **100.00**	**2,756** **100.00**

Pearson chi2(**1**) = **69.3428** Pr = **0.000**

```
. ta family_sui suicide, column chi2
```

Key
frequency *column percentage*

H3TO135	H3TO130 0	1	Total
0	**2,529** **97.80**	**146** **90.12**	**2,675** **97.34**
1	**57** **2.20**	**16** **9.88**	**73** **2.66**
Total	**2,586** **100.00**	**162** **100.00**	**2,748** **100.00**

Pearson chi2(**1**) = **34.7034** Pr = **0.000**

```
. ta insurance suicide, column chi2
```

Key
frequency *column percentage*

H3HS3	H3TO130 0	H3TO130 1	Total
0	**817** **31.51**	**72** **44.17**	**889** **32.26**
1	**1,776** **68.49**	**91** **55.83**	**1,867** **67.74**
Total	**2,593** **100.00**	**163** **100.00**	**2,756** **100.00**

Pearson chi2(**1**) = **11.2553** Pr = **0.001**

```
. probit suicide age i.bachelor i.depression i.friend_sui i.family_sui i.insurance

Iteration 0:   log likelihood = -615.75542
Iteration 1:   log likelihood = -537.56955
Iteration 2:   log likelihood = -533.78989
Iteration 3:   log likelihood = -533.78436
Iteration 4:   log likelihood = -533.78436

Probit regression                                 Number of obs     =      2,748
                                                  LR chi2(6)        =     163.94
                                                  Prob > chi2       =     0.0000
Log likelihood = -533.78436                       Pseudo R2         =     0.1331
```

suicide	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	**-.0617519**	**.0247958**	**-2.49**	**0.013**	**-.1103509**	**-.013153**
1.bachelor	**-.0639765**	**.133203**	**-0.48**	**0.631**	**-.3250496**	**.1970965**
1.depression	**.9567175**	**.0986448**	**9.70**	**0.000**	**.7633772**	**1.150058**
1.friend_sui	**.7011634**	**.1245836**	**5.63**	**0.000**	**.456984**	**.9453428**
1.family_sui	**.600879**	**.1840619**	**3.26**	**0.001**	**.2401243**	**.9616337**
1.insurance	**-.1761687**	**.0887079**	**-1.99**	**0.047**	**-.3500329**	**-.0023045**
_cons	**-.3545894**	**.5331958**	**-0.67**	**0.506**	**-1.399634**	**.6904551**

. fitstat

	probit
Log-likelihood	
Model	**-533.784**
Intercept-only	**-615.755**
Chi-square	
Deviance (df=2741)	**1067.569**
LR (df=6)	**163.942**
p-value	**0.000**
R2	
McFadden	**0.133**
McFadden (adjusted)	**0.122**
McKelvey & Zavoina	**0.144**
Cox-Snell/ML	**0.058**
Cragg-Uhler/Nagelkerke	**0.160**
Efron	**0.082**
Tjur's D	**0.088**
Count	**0.939**
Count (adjusted)	**-0.037**
IC	
AIC	**1081.569**
AIC divided by N	**0.394**
BIC (df=7)	**1122.999**
Variance of	
e	**1.000**
y-star	**1.168**

```
. logit suicide age i.bachelor i.depression i.friend_sui i.family_sui i.insurance

Iteration 0:   log likelihood = -615.75542
Iteration 1:   log likelihood = -543.37385
Iteration 2:   log likelihood = -535.65502
Iteration 3:   log likelihood = -535.61114
Iteration 4:   log likelihood = -535.61112

Logistic regression                             Number of obs     =      2,748
                                                LR chi2(6)        =     160.29
                                                Prob > chi2       =     0.0000
Log likelihood = -535.61112                     Pseudo R2         =     0.1302
```

suicide	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.1306338	.05187	-2.52	0.012	-.2322972	-.0289704
1.bachelor	-.0855061	.2834185	-0.30	0.763	-.6409963	.469984
1.depression	1.838997	.185611	9.91	0.000	1.475206	2.202788
1.friend_sui	1.308352	.23404	5.59	0.000	.8496419	1.767062
1.family_sui	1.171607	.3335452	3.51	0.000	.5178706	1.825344
1.insurance	-.3623541	.1814049	-2.00	0.046	-.7179011	-.0068071
_cons	-.2202613	1.111084	-0.20	0.843	-2.397946	1.957424

```
. margins, at(age=(18(1)28)) atmeans noatlegend

Adjusted predictions                            Number of obs     =      2,748
Model VCE    : OIM

Expression   : Pr(suicide), predict()
```

	Delta-method				
	Margin	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_at						
1	.0734038	.0159909	4.59	0.000	.0420622	.1047455
2	.064999	.0115148	5.64	0.000	.0424304	.0875676
3	.0574968	.00802	7.17	0.000	.0417779	.0732158
4	.0508135	.0055598	9.14	0.000	.0399164	.0617106
5	.04487	.0042973	10.44	0.000	.0364475	.0532926
6	.0395928	.0042003	9.43	0.000	.0313603	.0478252
7	.0349135	.0047329	7.38	0.000	.0256373	.0441897
8	.0307695	.0053879	5.71	0.000	.0202094	.0413296
9	.0271036	.0059508	4.55	0.000	.0154401	.038767
10	.0238636	.0063624	3.75	0.000	.0113936	.0363337
11	.0210027	.0066198	3.17	0.002	.008028	.0339773

```
. margins i.depression, at(c.age=(18(1)28)) atmeans noatlegend

Adjusted predictions                            Number of obs     =      2,748
Model VCE    : OIM

Expression   : Pr(suicide), predict()
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
_at#depression						
1 0	.0620614	.013932	4.45	0.000	.0347552	.0893676
1 1	.2938905	.0540507	5.44	0.000	.187953	.3998279
2 0	.0548784	.010092	5.44	0.000	.0350985	.0746583
2 1	.2675287	.0433352	6.17	0.000	.1825934	.3524641
3 0	.0484838	.0071194	6.81	0.000	.03453	.0624376
3 1	.2427188	.0345595	7.02	0.000	.1749835	.3104542
4 0	.0428005	.0050432	8.49	0.000	.0329161	.0526849
4 1	.2195202	.0282459	7.77	0.000	.1641592	.2748811
5 0	.037757	.0039585	9.54	0.000	.0299984	.0455156
5 1	.1979592	.0247946	7.98	0.000	.1493627	.2465558
6 0	.0332872	.0038035	8.75	0.000	.0258324	.0407419
6 1	.1780329	.0240354	7.41	0.000	.1309243	.2251414
7 0	.0293304	.0041698	7.03	0.000	.0211577	.037503
7 1	.1597129	.0250885	6.37	0.000	.1105403	.2088855
8 0	.0258313	.0046613	5.54	0.000	.0166953	.0349673
8 1	.1429502	.0269252	5.31	0.000	.0901778	.1957226

Expression : **Pr(suicide), predict()**

	Delta-method Margin	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_at#insurance						
1 0	.0919642	.0190317	4.83	0.000	.0546628	.1292655
1 1	.0658511	.0159769	4.12	0.000	.034537	.0971653
2 0	.0816214	.0142421	5.73	0.000	.0537074	.1095354
2 1	.0582567	.011818	4.93	0.000	.035094	.0814195
3 0	.0723491	.0107464	6.73	0.000	.0512866	.0934116
3 1	.0514899	.0085609	6.01	0.000	.0347109	.0682689
4 0	.0640567	.0085963	7.45	0.000	.0472083	.0809051
4 1	.0454711	.0062042	7.33	0.000	.0333112	.0576311
5 0	.0566567	.0077192	7.34	0.000	.0415273	.0717861
5 1	.0401261	.0047982	8.36	0.000	.0307218	.0495304
6 0	.0500658	.0077482	6.46	0.000	.0348797	.065252
6 1	.0353861	.004316	8.20	0.000	.026927	.0438453
7 0	.0442058	.0081934	5.40	0.000	.0281469	.0602646
7 1	.0311879	.004465	6.99	0.000	.0224367	.039939
8 0	.0390034	.008714	4.48	0.000	.0219242	.0560826

```
. mchange, atmeans

logit: Changes in Pr(y) | Number of obs = 2748

Expression: Pr(suicide), predict(pr)
```

	Change	p-value
age		
+1	**-0.005**	**0.007**
+SD	**-0.009**	**0.004**
Marginal	**-0.005**	**0.011**
bachelor		
1 vs 0	**-0.003**	**0.757**
depression		
1 vs 0	**0.154**	**0.000**
friend sui		
1 vs 0	**0.092**	**0.000**
family sui		
1 vs 0	**0.081**	**0.021**
insurance		
1 vs 0	**-0.016**	**0.060**

Predictions at base value

	0	1
Pr(y|base)	**0.957**	**0.043**

```
. mchange

logit: Changes in Pr(y) | Number of obs = 2748

Expression: Pr(suicide), predict(pr)
```

	Change	p-value
age		
+1	**-0.006**	**0.009**
+SD	**-0.011**	**0.006**
Marginal	**-0.007**	**0.012**
bachelor		
1 vs 0	**-0.004**	**0.757**
depression		
1 vs 0	**0.159**	**0.000**
friend sui		
1 vs 0	**0.100**	**0.000**
family sui		
1 vs 0	**0.089**	**0.012**
insurance		
1 vs 0	**-0.019**	**0.055**

Average predictions

	0	1

```
. mtable, ci clear at(age=18 depression=1 friend_sui=0) atmeans

Expression: Pr(suicide), predict()

   Pr(y)        ll        ul
------------------------------
   0.277     0.174     0.380

Specified values of covariates

           |                   1.                                  1.          1.
           |      age    bachelor    depres~n    friend~i      family~i     insura~e
-----------+-----------------------------------------------------------------------
   Current |       18        .159           1           0         .0266         .678

. mtable, ci clear at(age=28 depression=1 friend_sui=0) atmeans

Expression: Pr(suicide), predict()

   Pr(y)        ll        ul
------------------------------
   0.094     0.036     0.152

. mtable, ci clear at(age=18 depression=1 friend_sui=1) atmeans

Expression: Pr(suicide), predict()

   Pr(y)        ll        ul
------------------------------
   0.586     0.435     0.737

Specified values of covariates

           |                   1.                                  1.          1.
           |      age    bachelor    depres~n    friend~i      family~i     insura~e
-----------+-----------------------------------------------------------------------
   Current |       18        .159           1           1         .0266         .678

. mtable, ci clear at(age=18 depression=0 friend_sui=1) atmeans

Expression: Pr(suicide), predict()

   Pr(y)        ll        ul
------------------------------
```

Logistic regression

Number of obs	=	**2,748**	
LR chi2(**4**)	=	**111.39**	
Prob > chi2	=	**0.0000**	
Log likelihood = **-560.0619**	Pseudo R2	=	**0.0904**

suicide	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	**-.144997**	**.0510498**	**-2.84**	**0.005**	**-.2450527**	**-.0449412**
1.bachelor	**-.1529085**	**.2808671**	**-0.54**	**0.586**	**-.7033978**	**.3975809**
1.depression	**1.835938**	**.1801393**	**10.19**	**0.000**	**1.482871**	**2.189005**
1.insurance	**-.3474148**	**.1781759**	**-1.95**	**0.051**	**-.6966331**	**.0018035**
_cons	**.3155324**	**1.089495**	**0.29**	**0.772**	**-1.819838**	**2.450903**

. lrtest . base

Likelihood-ratio test — LR chi2(**2**) = **48.90**
(Assumption: . nested in base) — Prob > chi2 = **0.0000**