

- About 77.8% of high school students attend college.
- Variables: Grades, family income/size, parents attended college or not, sexuality and ethnicity.

Probit model explaining whether one will enroll in college:

Coefficients:

	Estimate	Std. Error	z value	p-value
(Intercept)	2.693704	0.285643	9.430	< 2e-16
GRADES	-0.294552	0.027452	-10.730	< 2e-16
FAMINC	0.005392	0.001976	2.728	0.006366
FAMSIZ	-0.053108	0.037349	-1.422	0.155047
PARCOLL	0.476540	0.141154	3.376	0.000735
FEMALE	0.023788	0.101626	0.234	0.814926
BLACK	0.610896	0.215740	2.832	0.004631

AIC: 846.44

We can see that **GRADES, FAMINC, PARCOLL, BLACK variables are statistically significant.**

Original expectations of the coefficients before running the model are:

GRADES, FAMINC, FEMALE, BLACK will be significant, not sure about FAMSIZ and PARCOLL.

Suspect that FAMSIZ will have a negative effect and PARCOLL will have a positive effect.

The difference between actual results and the original expectation:

Not surprised that GRADES, FAMINC, BLACK are all significant. Also live up to the original expectation that FAMSIZ has a negative effect.

But quite out of our expectation that FEMALE variable statistically is not significant at all, and PARCOLL variable actually is very significant.

Also, note that intercepts are statistically significant as well, meaning that there are still lots of other factors that haven't been taken into consideration but actually affect whether one will attend college or not.

- Use a certain assumption individual to see the probability of attending college:

The result of running the prediction in the probit model is **0.924101.**

⇒  $\Phi(0.9241) = 0.8221$  (approximately)

**Conclusion: 82.21% probability** that a black female high school graduates with grade=10, average family income, 4 members family, one parent that attended college, will go to college.

- What about using Logit model?

Coefficients:

	Estimate	Std. Error	z value	p-value
(Intercept)	4.541835	0.519120	8.749	< 2e-16
GRADES	-0.517438	0.050464	-10.254	< 2e-16
FAMINC	0.013042	0.003897	3.346	0.000819
FAMSIZ	-0.088256	0.064466	-1.369	0.170987
PARCOLL	0.835371	0.262977	3.177	0.001490
FEMALE	0.034279	0.176918	0.194	0.846365
BLACK	1.073344	0.372292	2.883	0.003938

AIC: 843.95

Although the coefficients estimate and standard errors are very different from those of the Probit model, the z values and p-values are quite similar. The results of whether variables have positive or negative effects on COLLEGE are also the same. In terms of whether variables are statistically significant or not, we also get the same result.

**The conclusion that we get from Logit model is basically the same as the Probit model, even though the estimation itself does not look similar at all.**