1. **Logit model exercise**
2. The probability that a white applicant will be denied at P/I = 0.35 is 9.530634465210209 percent while the probability that a white applicant is denied at P/I = 0.30 is 7.4536896971794455 percent, where a 0.30 P/I causes the probability of the mortgage to be denied to decrease by 2.0769447680307636 percent

Ans:

**P1\_b = logisitic.cdf() use 0.35**

**P1\_a = logistic.cdf() use 0.30**

**Calculate difference**

1. The marginal effect of P/I ratio on mortgage denial probability is dependent on race.

This is because the marginal difference in percentage between black people and white people when the P/I ratio is 0.30 as opposed to 0.35 is larger, with the difference for black people being 4.991929238331526 percent compared to white people 2.0769447680307636,

as the percentage of denial at P/I ratio of 0.35 is 27.27925839964312 for black people and 22.287329161311593 at 0.30 for black people.

1. Given that black is a binary variable, a coefficient of 0.084 shows that which means that black people have an 8.762889381% higher chance of being denied a mortgage compared to white people.
2. One omitted variable bias in the model could be the wage level of the black or white people who aim to get the mortgage. This is because wages could affect the independent variable of the probability of being denied or not, since higher wages could lead to the bank having more confidence in the person purchasing the mortgage ability to pay back the mortgage and subsequently will have a higher chance of accepting the mortgage. Furthermore, wages could affect the independent variable, as black people who could be poorer than white people have less wages. As a result of this omitted variable bias, black people could have different percent higher chance of being denied or even a different lower percent chance of being denied a mortgage, based on the answer in c).
3. **Exercise model specification**
4. Since wages are expected to change by x%, I will use a log-linear model as opposed to the fixed constant change of a linear model. Thus, the equation is as follows:
5. We must add an interaction variable between the gender and smoking variables. This causes the regression equation to look like this:

The testing hypothesis would be as follows

H0:

H1:

We will use a t-test to see which hypothesis to reject. The t-statistic is as follows

, where from the t-test will we deduce the p-value from the standard normal distribution table. Assuming we are testing at the 5% significance level of α=0.05, if the p-value is < 0.05, we can reject the null hypothesis and accept the alternative hypothesis, else we cannot reject the null hypothesis that there is no effect.

1. This will change the model according as seen below.

To test whether smoking as any effect on wages, the hypothesis is as follows:

H0:

H1:

Conduct a Wald-test. It is noted that there are 4 restrictions in the given model and that there are 7 coefficients. Conduct a F-test with the degree of freedom in the numerator being equal to 4 and the degree of freedom being n-7.

Ans: **don’t include non-smokers, H0: let light = moderate = heavy = 0**

**H1: let light, moderate and heavy not equal 0**

**Use F-test, without the dataset we cannot determine the degrees of freedom**