Math 158 – Spring 2016
Jo Hardin
warm-up # 3

Name: ________________________________

Consider the regression model handouts concerning the birth weight data. Carry out an (one!) F test to evaluate whether, when mother's age and weight are both in the model, the smoking main effect and smoking*gained interaction are simultaneously not needed. Note that you need to write out your null and alternative hypotheses, p-value (make a sketch of the appropriate area), conclusion, and summary in the context of the problem.
You might need the following output:

```
> anova(lm(tounces ~ gained + mage))
Analysis of Variance Table

Response: tounces
           Df Sum Sq Mean Sq F value    Pr(>F)
gained      1  18856   18856  41.487 2.083e-10 ***
mage        1  10891   10891  23.963 1.195e-06 ***
Residuals 774 351780     454
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

Solution:
Consider the following model:

$$\begin{aligned}
E[Y] &= \beta_0 + \beta_1 gained + \beta_2 smoke + \beta_3 mage + \beta_4 gained \cdot smoke \\
H_0: & \quad \beta_2 = \beta_4 = 0 \\
H_a: & \quad \text{not } H_0
\end{aligned}$$

Our test statistic is calculated using the SSE from the full and reduced models:

$$\begin{aligned}
F^* &= \frac{\frac{SSE(R)-SSE(F)}{(n-3)-(n-5)}}{\frac{SSE(F)}{n-5}} \\
&= \frac{\frac{351780-346389}{2}}{449} \\
&= 6.00 \\
\text{p-value} &= P(F_{2,772} \geq 6) \\
&= 1 - pf(6, 2, 772) \\
&= 0.002595839
\end{aligned}$$

There is strong evidence that β_2 and β_4 are not simultaneously zero. That is, we should not remove both smoking and the gained*smoking interaction from the model that predicts baby's birth weight in ounces *conditional* on gained and mage being in the model.