Lecture 3: Hypothesis Testing

(Text Sections 1.4.3 and 1.4.4, and parts of Chapter 6)

Example 1: t-test

Q: Does the number of days on the market significantly affect the selling price (when all the predictor variables except list price are included in the model)?

Let β_2 be the effect of days on the market on selling price. The null hypothesis is $H_0: \ \beta_2 = 0$, vs. the alternative hypothesis $H_A: \ \beta_2 \neq 0$. Under H_0,

$$t^* = \frac{\hat{\beta}_2 - 0}{\sqrt{\hat{Var}[\hat{\beta}_2]}} \sim T_{n-p}$$

where T_{n-p} is a random variable with the t-distribution on $n-p$ df.

Q: How do we interpret β_2?

Q: What is $\sqrt{\hat{Var}[\hat{\beta}_2]}$?

Let $t_{\alpha,n-p}$ be the quantile of T_{n-p} such that

$$\mathrm{P}(T_{n-p} \leq t_{\alpha,n-p}) = \alpha.$$

Then, we reject H_0 at level α if $t^* < t_{\frac{\alpha}{2},n-p}$ or if $t^* > t_{1-\frac{\alpha}{2},n-p}$ (or, equivalently, if $|t^*| > t_{1-\frac{\alpha}{2},n-p}$).

Similarly, a $100(1-\alpha)\%$ CI for β_2 is given by

$$\hat{\beta}_2 \pm t_{1-\frac{\alpha}{2},n-p}\sqrt{\hat{Var}[\hat{\beta}_2]}.$$

****In other words, if we were to repeat this experiment many times, we would expect* $100(1-\alpha)\%$ *of the CI's to contain the true value of* β_2.***

Informally, we are $100(1-\alpha)\%$ confident that the true value of β_2 lies in this interval.

Alternatively, we can compute the *p-value* associated with this test:

$$\text{p-value} = \mathrm{P}(T_{n-p} < -|t^*|) + \mathrm{P}(T_{n-p} > |t^*|) = 2\mathrm{P}(T_{n-p} > |t^*|).$$

We reject H_0 if the p-value is less than our chosen significance level (S.L.).

Relative to the commonly used S.L. of $\alpha = 0.05$, the p-value is often interpreted as

p-value≤ 0.01	strong evidence against H_0
$0.01 <$p-value≤ 0.05	some evidence against H_0
$0.05 <$p-value≤ 0.1	weak evidence against H_0
p-value> 0.1	no evidence against H_0

THESE ARE INFORMAL GUIDELINES ONLY!!!

Q: How do we formally interpret a p-value?

NOTE: This is a *2-sided test*, corresponding to $H_A: \ \beta_2 \neq 0$! If H_A were instead $\beta_2 < 0$, we would use a *1-sided test*. In this case we reject H_0 at level α if $t^* < t_{\alpha,n-p}$. The p-value associated with this test is

$$\text{p-value} = \mathrm{P}(T_{n-p} < t^*)$$

Again, we reject H_0 if the p-value is less than α.

Q: Which type of p-value (1- or 2-sided) does S-PLUS provide in its standard output?

Example 2: F-test

Does the presence of a suite affect the selling price (when Renovated is included in the model)?

We have

H_0: suite does not affect selling price
H_A: suite does affect selling price.

We can think of two models: that which includes both Suite and Renovated (the *full model*), and that with just Renovated (the *reduced model*). Then,

H_0: reduced model is adequate
H_A: full model is required.

Let RSS(R) and RSS(F) be the residual sums of squares of the reduced and full models, respectively, and let df(R) and df(F) be the associated df. Under H_0,

$$F^* = \frac{\frac{RSS(R)-RSS(F)}{df(R)-df(F)}}{\frac{RSS(F)}{df(F)}} \sim F(df(R)-df(F), df(F)),$$

where $F(a, b)$ is the F-distribution on a and b df. Let $f(1-\alpha, a, b)$ be the quantile of $F(a, b)$ such that

$$\mathrm{P}\{F(a, b) < f(1-\alpha, a, b)\} = 1-\alpha$$

We reject H_0 at level α if $F^* > f(1-\alpha, df(R) - df(F), df(F))$.

The p-value associated with this test is

$$\text{p-value} = \mathrm{P}\{F(df(R) - df(F), df(F)) > F^*\}$$

If we are testing the significance of only one factor, we can use the ANOVA table, e.g. as provided by S-PLUS. The ANOVA table gives the sums of squares (SS) and df associated with each factor in the model. Let the factor of interest (e.g. Suite) be A, and let SS(A) and df(A) be the corresponding sums of squares and df. (Recall that if A has q levels, then df(A)=$q-1$.) Then

$$F^* = \frac{\frac{SS(A)}{q-1}}{\frac{RSS(F)}{df(F)}}.$$

For a factor with only two levels, p-value based on this test will be the same as that based on the 2-sided t-test.

WARNING: If the factor of interest is A, make sure that A is the *last* factor in the list of predictors in the ANOVA table (i.e. by specifying A as the last predictor in the `lm()` command). The p-value associated with A will then correspond to the test of the effect of A given that all of the factors *above it* are in the model. *Always read the p-values in the ANOVA table from bottom to top.* In contrast, the p-values based on t-tests given in the `summary` output are associated with the test of each factor given that *all* other factors are in the model (i.e. the order doesn't matter).

Q: Is it possible to use an F-test to test whether a suite *increases* the selling price?

Exercise: Do the lot size and house size affect the selling price (when all the predictor variables are included in the model)?