

In a multiple regression, the  $t$  statistics are used to test hypotheses about individual coefficients. If the model is  $Y_i = \beta_0 + \beta_H H_i + \beta_K K_i + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$ , then the  $t$  statistic printed in the line for  $b_Q$  (the estimate of  $\beta_Q$ ) tests the hypothesis  $H_0(Q): \beta_Q = 0$  against the alternative  $H_1(Q): \beta_Q \neq 0$ . This must be interpreted in the context of this regression and *with* the predictors  $\{H, K, V\}$ . This is most definitely *not* independent of the tests on  $\beta_H$ ,  $\beta_K$ , or  $\beta_V$ .

Suppose instead that you wished to test the hypothesis  $H_0: \beta_H = \beta_K$  against the alternative  $H_1: \beta_H \neq \beta_K$ . This kind of inference is not automated in Minitab. We can get it, but we'll have to work for it.

As a side note, you can only consider  $\beta_H = \beta_K$  if variables  $H$  and  $K$  are in the same units.

We are going to use the “partial  $F$  test.” This is explained neatly in Chatterjee and Hadi. Let's suppose that we have a full model and a reduced model. In our example,

the full model is  $Y_i = \beta_0 + \beta_H H_i + \beta_K K_i + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$

the reduced model is  $Y_i = \beta_0 + \beta_S S_i + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$

where  $S_i = H_i + K_i$

A requirement for using this partial  $F$  test is that the reduced model must be a sub-model of the full model. This means that there is some choice for the parameters in the full model that produces the reduced model. Start from

$$Y_i = \beta_0 + \beta_H H_i + \beta_K K_i + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$$

and then make the choice that  $\beta_H = \beta_K$  to write

$$Y_i = \beta_0 + \beta_H H_i + \beta_H K_i + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$$

This can be regrouped as

$$Y_i = \beta_0 + \beta_H (H_i + K_i) + \beta_Q Q_i + \beta_V V_i + \varepsilon_i$$

Now renaming  $H_i + K_i$  as  $S_i$  and also letting  $\beta_S$  be another name for the assumed common value of  $\beta_H$  and  $\beta_K$  will produce the reduced model.



Do regressions for both the full model and the reduced model. Observe that

$$SS_{\text{regression}}(\text{full}) > SS_{\text{regression}}(\text{reduced})$$

$$SS_{\text{resid}}(\text{full}) < SS_{\text{resid}}(\text{reduced})$$

$$SS_{\text{regression}}(\text{full}) - SS_{\text{regression}}(\text{reduced}) = SS_{\text{resid}}(\text{reduced}) - SS_{\text{resid}}(\text{full})$$

The partial  $F$  statistic can be written in several ways:

$$\begin{aligned} F &= \frac{\{SS_{\text{regression}}(\text{full}) - SS_{\text{regression}}(\text{reduced})\} \div v}{\frac{SS_{\text{resid}}(\text{full})}{DF_{\text{resid}}(\text{full})}} \\ &= \frac{\{SS_{\text{resid}}(\text{reduced}) - SS_{\text{resid}}(\text{full})\} \div v}{\frac{SS_{\text{resid}}(\text{full})}{DF_{\text{resid}}(\text{full})}} \\ &= \frac{\{SS_{\text{regression}}(\text{full}) - SS_{\text{regression}}(\text{reduced})\} \div v}{MS_{\text{resid}}(\text{full})} \\ &= \frac{\{SS_{\text{resid}}(\text{reduced}) - SS_{\text{resid}}(\text{full})\} \div v}{MS_{\text{resid}}(\text{full})} \end{aligned}$$

In every form, the numerator is based on a sum squares difference that asks how much better is the fit with the full model. The value  $v$  is the number of parameters that distinguish  $H_0$  from  $H_1$ . For this example,

the full model has  $E(Y_i) = \beta_0 + \beta_H H_i + \beta_K K_i + \beta_Q Q_i + \beta_V V_i$

and it takes five parameters to specify this  $(\beta_0, \beta_H, \beta_K, \beta_Q, \beta_V)$

the reduced model has  $E(Y_i) = \beta_0 + \beta_S S_i + \beta_Q Q_i + \beta_V V_i$

and it takes four parameters to specify this  $(\beta_0, \beta_S, \beta_Q, \beta_V)$

For this example,  $v = 1$ .

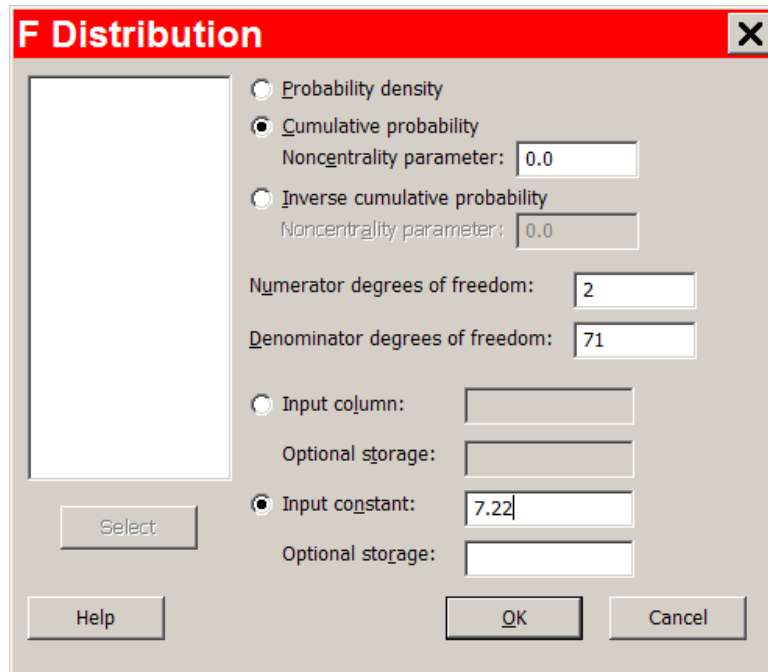
The denominator of the partial  $F$  statistic is the mean square residual from the full model. This denominator estimates  $\sigma^2$  whether  $H_0$  is true or not.

PARTIAL  $F$  TEST

The degrees of freedom for the partial  $F$  are  $(v, DF_{\text{resid}}(\text{full}))$ . The rule is to reject  $H_0$  at level  $\alpha$  if the partial  $F$  equals or exceeds  $F_{v, DF_{\text{resid}}(\text{full})}^\alpha$ , the upper  $\alpha$  point for the  $F$  distribution with  $(v, DF_{\text{resid}}(\text{full}))$  degrees of freedom.

We are out of the habit of looking up cutoff points for the  $F$  distribution because most software prints the  $p$ -value. We would just reject at level  $\alpha$  if  $p \leq \alpha$ .

The partial  $F$  is not arranged by Minitab. You have to do two runs and assemble this yourself. This also means that you won't get a  $p$ -value. You could then use printed tables of the  $F$  distribution, which give the upper 5% and 1% points. Since Minitab is already active for the work you're doing, you can get  $p$ -value yourself. Let's suppose that your partial  $F$  works out to the value 7.22 with (2, 71) degrees of freedom. Use **Calc**  $\Rightarrow$  **Probability Distributions**  $\Rightarrow$  **F** and then fill in the resulting panel as follows:



This starts with the default radio button setting at **Cumulative probability**, which is exactly what you want. The output is this:

**Cumulative Distribution Function**

F distribution with 2 DF in numerator and 71 DF in denominator

x	P( X <= x )
7.22	0.998601

The probability to the left of your 7.22 is found to be 0.9986. The probability to the right (corresponding to the rejection region for  $H_0$  versus  $H_1$ ) is then 0.0014. You can report then  $p = 0.0014$ .