

3.3: Evaluating the Quality of the Model

The information we obtain by typing `int00.lm` shows us the regression model's basic values, but does not tell us anything about the model's quality. In fact, there are many different ways to evaluate a regression model's quality. Many of the techniques can be rather technical, and the details of them are beyond the scope of this tutorial. However, the function `summary()` extracts some additional information that we can use to determine how well the data fit the resulting model. When called with the model object `int00.lm` as the argument, `summary()` produces the following information:

```
> summary(int00.lm)
Call:
lm(formula = perf ~ clock)
Residuals:
Min       1Q       Median       3Q      Max
-634.61  -276.17  -30.83    75.38   1299.52
Coefficients:
              Estimate      Std. Error  t value Pr(>|t|)
(Intercept)   51.78709     53.31513    0.971   0.332
clock         0.58635      0.02697   21.741 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 396.1 on 254 degrees of freedom
Multiple R-squared:  0.6505, Adjusted R-squared:  0.6491
F-statistic: 472.7 on 1 and 254 DF, p-value: < 2.2e-16
```

Let's examine each of the items presented in this summary in turn.

```
> summary(int00.lm)
Call:
lm(formula = perf ~ clock)
```

These first few lines simply repeat how the `lm()` function was called. It is useful to look at this information to verify that you actually called the function as you intended.

```
Residuals:
      Min       1Q       Median       3Q      Max
-634.61  -276.17  -30.83    75.38   1299.52
```

The *residuals* are the differences between the actual measured values and the corresponding values on the fitted regression line. In Figure 3.2, each data point's residual is the distance that the individual data point is above (positive residual) or below (negative residual) the regression line. `Min` is the minimum residual value, which is the distance from the regression line to the point furthest below the line. Similarly, `Max` is the distance from the regression line of the point furthest above the line. `Median` is the median value of all of the residuals. The `1Q` and `3Q` values are the points that mark the first and third quartiles of all the sorted residual values.

How should we interpret these values? If the line is a good fit with the data, we would expect residual values that are normally distributed around a mean of zero. (Recall that a normal distribution is also called a Gaussian distribution.) This distribution implies that there is a decreasing probability of finding residual values as we move further away from the mean. That is, a good model's residuals should be roughly balanced around and not too far away from the mean of zero. Consequently, when we look at the residual values reported by `summary()`, a good model would tend to have a median value near zero, minimum and maximum values of roughly the same magnitude, and first and third quartile values of roughly the same magnitude. For this model,

the residual values are not too far off what we would expect for Gaussian-distributed numbers. In Section 3.4, we present a simple visual test to determine whether the residuals appear to follow a normal distribution.

```

Coefficients:
            Estimate      Std. Error    t value    Pr(>|t|)
(Intercept)  51.78709      53.31513      0.971      0.332
clock        0.58635       0.02697     21.741     <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

This portion of the output shows the estimated coefficient values. These values are simply the fitted regression model values from Equation 3.2. The Std. Error column shows the statistical `standard error` for each of the coefficients. For a good model, we typically would like to see a standard error that is at least five to ten times smaller than the corresponding coefficient. For example, the standard error for `clock` is 21.7 times smaller than the coefficient value ($0.58635/0.02697 = 21.7$). This large ratio means that there is relatively little variability in the slope estimate, a_1 . The standard error for the intercept, a_0 , is 53.31513, which is roughly the same as the estimated value of 51.78709 for this coefficient. These similar values suggest that the estimate of this coefficient for this model can vary significantly.

The last column, labeled `Pr(>|t|)`, shows the probability that the corresponding coefficient is *not* relevant in the model. This value is also known as the *significance* or p-value of the coefficient. In this example, the probability that `clock` is not relevant in this model is 2×10^{-16} a tiny value. The probability that the intercept is not relevant is 0.332, or about a one-inthree chance that this specific intercept value is not relevant to the model. There is an intercept, of course, but we are again seeing indications that the model is not predicting this value very well.

The symbols printed to the right in this summary that is, the asterisks, periods, or spaces are intended to give a quick visual check of the coefficients' significance. The line labeled `Signif. codes:` gives these symbols' meanings. Three asterisks (***) means $0 < p \leq 0.001$, two asterisks (**) means $0.001 < p \leq 0.01$, and so on.

R uses the column labeled `t value` to compute the p-values and the corresponding significance symbols. You probably will not use these values directly when you evaluate your model's quality, so we will ignore this column for now.

```

Residual standard error: 396.1 on 254 degrees of freedom
Multiple R-squared:  0.6505, Adjusted R-squared:  0.6491
F-statistic: 472.7 on 1 and 254 DF, p-value: < 2.2e-16

```

These final few lines in the output provide some statistical information about the quality of the regression model's fit to the data. The Residual standard error is a measure of the total variation in the residual values. If the `residuals` are distributed normally, the first and third quantiles of the previous residuals should be about 1.5 times this `standard error`.

The number of `degrees of freedom` is the total number of measurements or *observations* used to generate the model, minus the number of coefficients in the model. This example had 256 unique rows in the data frame, corresponding to 256 independent measurements. We used this data to produce a regression model with two coefficients: the slope and the intercept. Thus, we are left with ($256 - 2 = 254$) degrees of freedom.

The `Multiple R-squared` value is a number between 0 and 1. It is a statistical measure of how well the model describes the measured data. We compute it by dividing the total variation that the model explains by the data's total variation. Multiplying this value by 100 gives a value that we can interpret as a percentage between 0 and 100. The reported R^2 of 0.6505 for this model means that the model explains 65.05 percent of the data's variation. Random chance and measurement errors creep in, so the model will never explain all data variation. Consequently, you should not ever expect an R^2 value of exactly one. In general, values of R^2 that are closer to one indicate a better-fitting model. However, a good model does not necessarily require a large R^2 value. It may still accurately predict future observations, even with a small R^2 value.

The `Adjusted R-squared value` is the R^2 value modified to take into account the number of predictors used in the model. The adjusted R^2 is always smaller than the R^2 value. We will discuss the meaning of the adjusted R^2 in Chapter 4, when we present regression models that use more than one predictor.

The final line shows the `F-statistic` . This value compares the current model to a model that has one fewer parameters. Because the one-factor model already has only a single parameter, this test is not particularly useful in this case. It is an interesting statistic for the multi-factor models, however, as we will discuss later.

This page titled [3.3: Evaluating the Quality of the Model](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [David Lilja](#) ([University of Minnesota Libraries Publishing](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.