Confidence intervals for the predictors estimates

Bookmark this page

After fitting a linear regression model we often look at the coefficients to infer relationships between the predictors and the response variable. But we should always ask, "how reliable are our model interpretations?"

Suppose our model for advertising:

$$y = 1.01x + 120$$

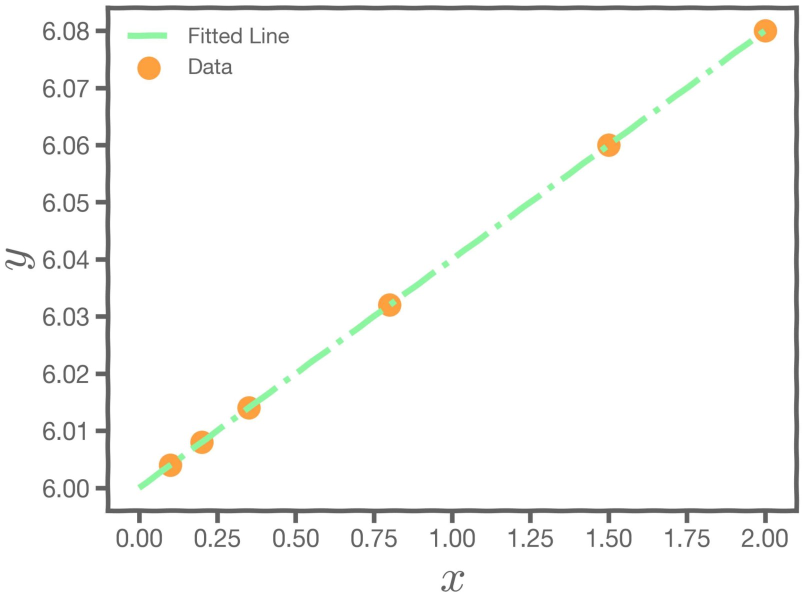
Here x are in \$1000s and y are in thousand unit sales and every unit sells for \$1.

Our interpretation might then be as follows: for every dollar invested in advertising, we get an additional 1.01 back in sales. That is, 1% profit.

Confidence intervals for the predictors estimates

Our observation, represented as $y = f(x) + \epsilon$, is interpreted either as the noise introduced by random variations in natural systems, or imprecisions of the measuring instruments, or environmental irregularities.

If we knew the exact form of $f(x)$,for example, $f(x) = \beta_0 + \beta_1(x)$, and there was no noise in the data, then the estimation of $\hat{\beta}s$ would have been exact.



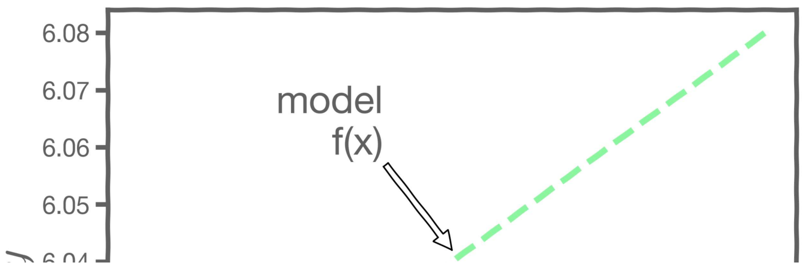
In such a case, is a 1% profit worth it? That's a business decision, and not one that linear regression can answer.

However, two things should make us mistrust of the values of $\hat{\beta}s$:

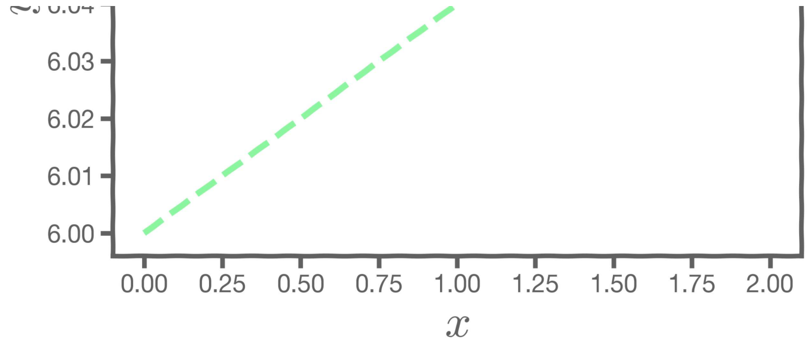
- observational error is always there - this is called **aleatoric** error or **irreducible** error
- the exact form of $f(x)$ is unknown - this is called **misspecification** error and is part of the **epistemic** error

Both errors are combined into a catch-it-all term, ϵ. Because of this, every time we measure the response y for a fixed value of x, we will obtain a different observation and hence a different estimate of $\hat{\beta}s$.

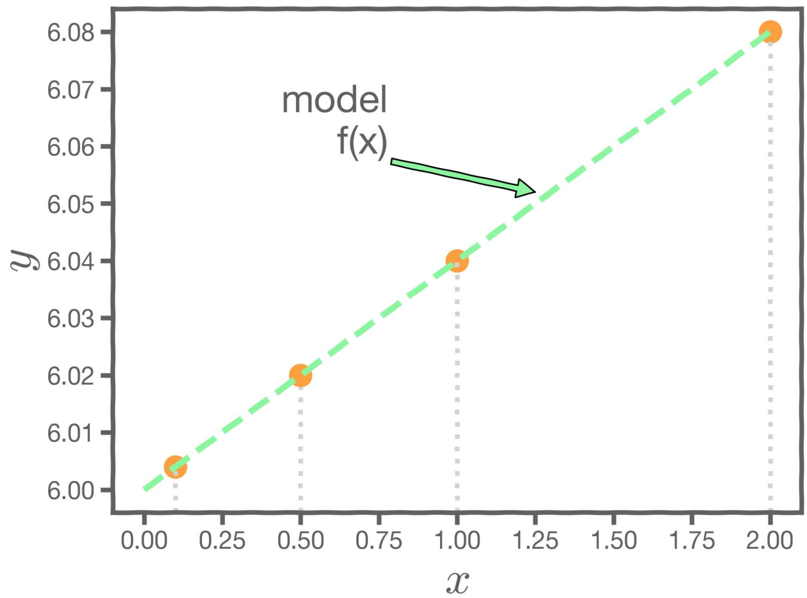
Let us consider an example. Start with a model $f(X)$, the correct relationship between input and outcome.



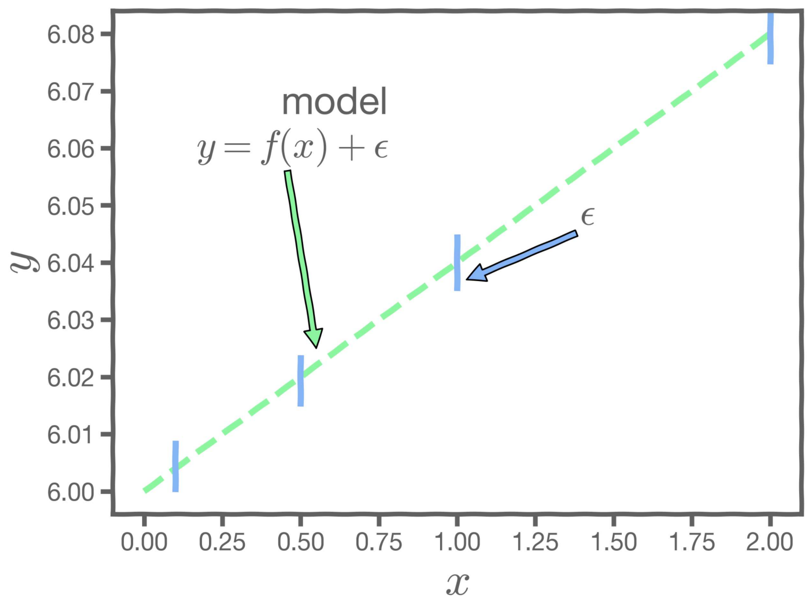
Calculator



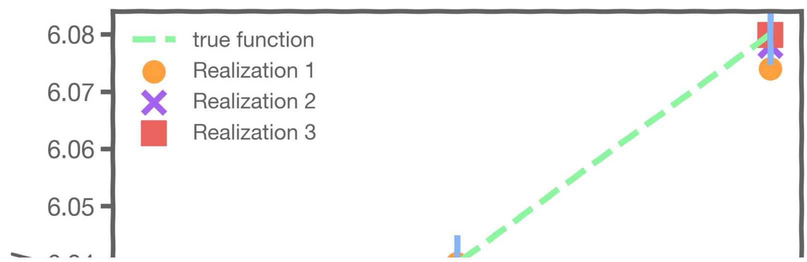
For some values of $\boldsymbol{x}$, $\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{X})$

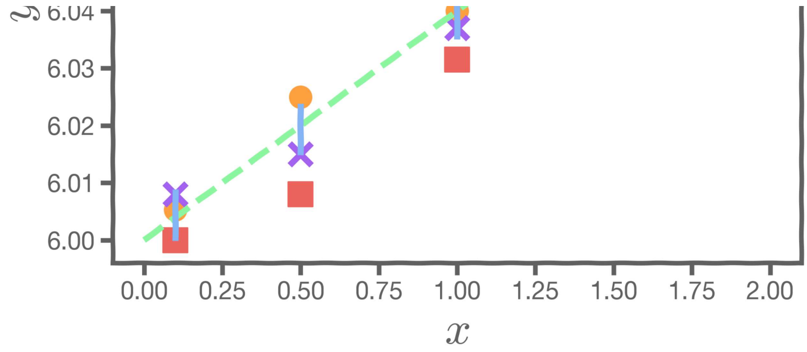


Every time we measure the response $\boldsymbol{Y}$ for a fixed $\boldsymbol{X}$ we obtain a different observation.

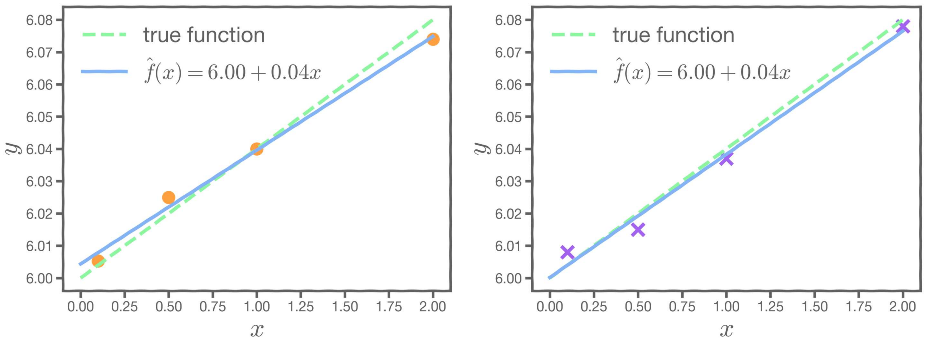


One set of observations, "one realization" yields one set of $\boldsymbol{Y}$. This is represented as orange circles in the plot. Similarly, the squares and crosses each represent a set of $\boldsymbol{Y}$ realized.



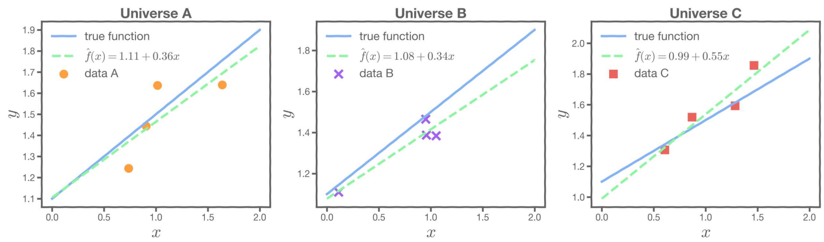


For each of these realizations, we fit a model and estimate $\hat{\beta}_0$ and $\hat{\beta}_1$. Thus, it results in one set of model parameters, $\hat{\beta}_0$ and $\hat{\beta}_1$, for each realization.

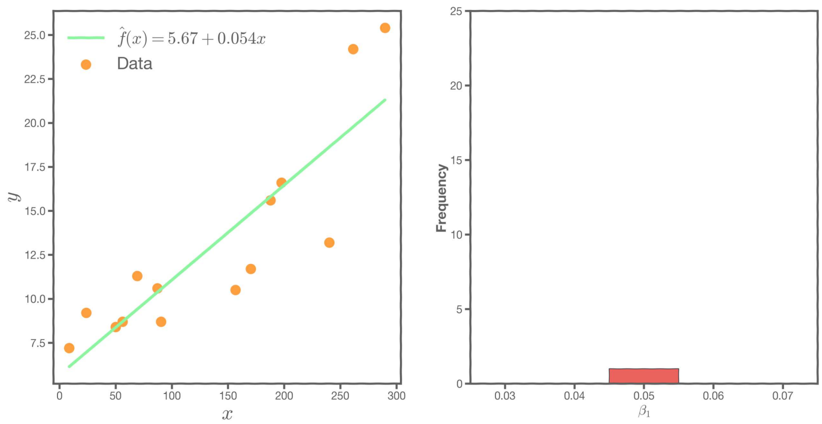


So, if we have one set of measurements of X, Y, our estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ are just for this particular realization. Given this, how do we know the truth? How do we deal with this conundrum?

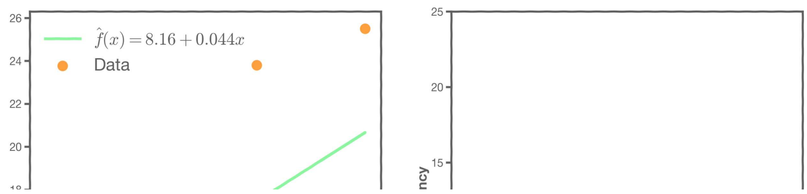
To resolve this, imagine that we have a multitude of parallel universes, and we repeat this experiment on each of the other universes.

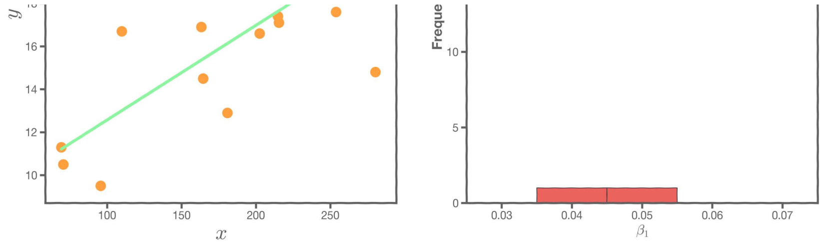


In our magical realism, we can now sample X, Y multiple times. One universe means one sample, which means one set of estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$. The graphs on the left below show the best-fit lines for each of the universes, and the right-hand graph shows the distribution of the estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$.

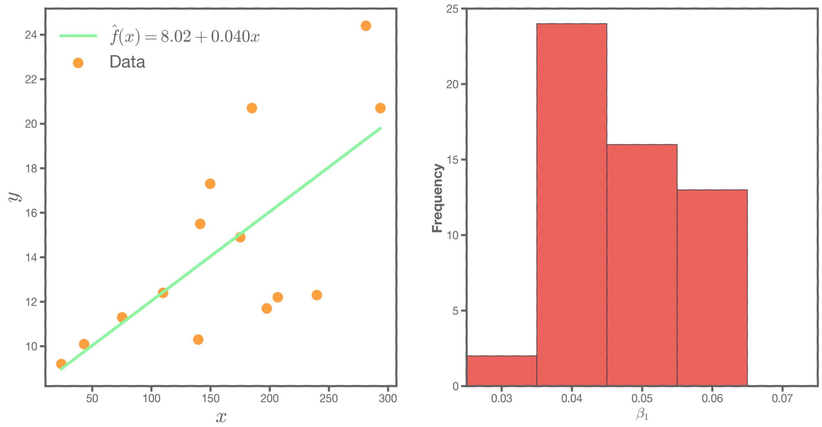


Another sample, another estimate of $\hat{\beta}_0$ and $\hat{\beta}_1$





This is repeated until we have sufficient samples of $\hat{\beta}_0$ and $\hat{\beta}_1$ to understand their distribution.



Discussion Board (External resource)

Click OK to have your username and e-mail address sent to a 3rd party application.

OK

Calculator