|  |  |
| --- | --- |
| Confidence interval calculations  Using numeric calculations | Abstract  TODO: list of methods, approach, ext…  Daniel Abutbul  Statistics Exercise |

Contents

[Motivation 2](#_Toc106370744)

[Working example 2](#_Toc106370745)

# Motivation

For many applications, we encounter the problem of creating predictions based on observed data. While a very general case can be formulated, if one wishes to practice, understand, implement and finally use such prediction, it is best first to solve the simplest imaginable problem, and try and solve it in as many methods as exist.

In this report, we will review and implement various methods for extrapolating the value of a straight line calibrated from observed data.

## Working example

Throughout this report, we will use the same simple example of the straight line:

The straight line formula tells us exactly the value of at any value of . For many real life situation, though, the formula and its parameters are unknown to the researcher. Instead, the researcher may have a sample of points which come from the straight line with added noise. This can represent for example a scientific experiment of measuring the value of at different ’s, with added experimental noise.

The sample is represented by a set of points , at each one we have the value . We model the noise using a normal random variable , such that is sampled normally around the true value :

It is very common and still useful to restrict ourselves to the case the noise of different points is independent of noise at any different point, and identical at any point:

Finally, our very specific working example will be of the data set , and . The straight line, a possible sample points with noise from the straight line, and theirs best fit ( minimized) are given in Figure 1:

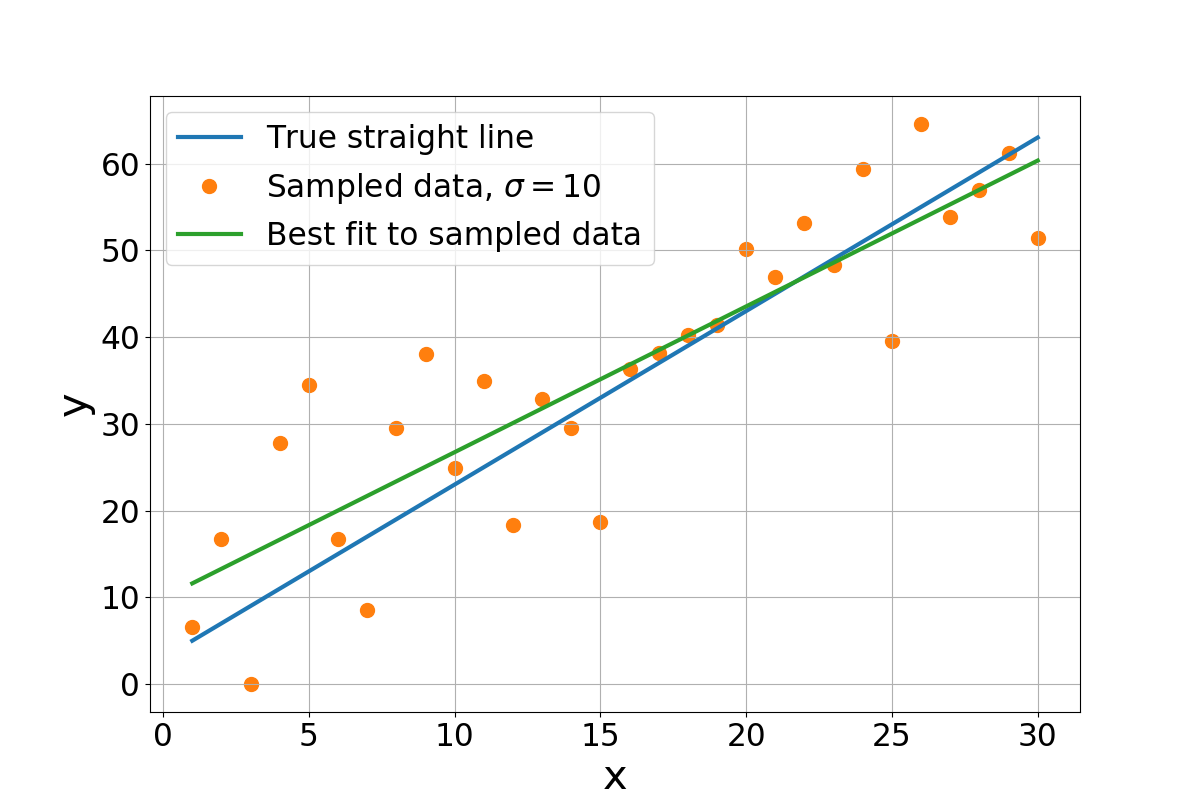


Figure 1: The straight line , sampled with normal noise ( i.i.d. random variables) and fitted with a straight line.

Our task is then to construct an extrapolation toward a value that is not in our data, say , and decide on some range around it, such that in most noise realization the true value of the straight line will be within our predicted range, as demonstrated in Figure 2:

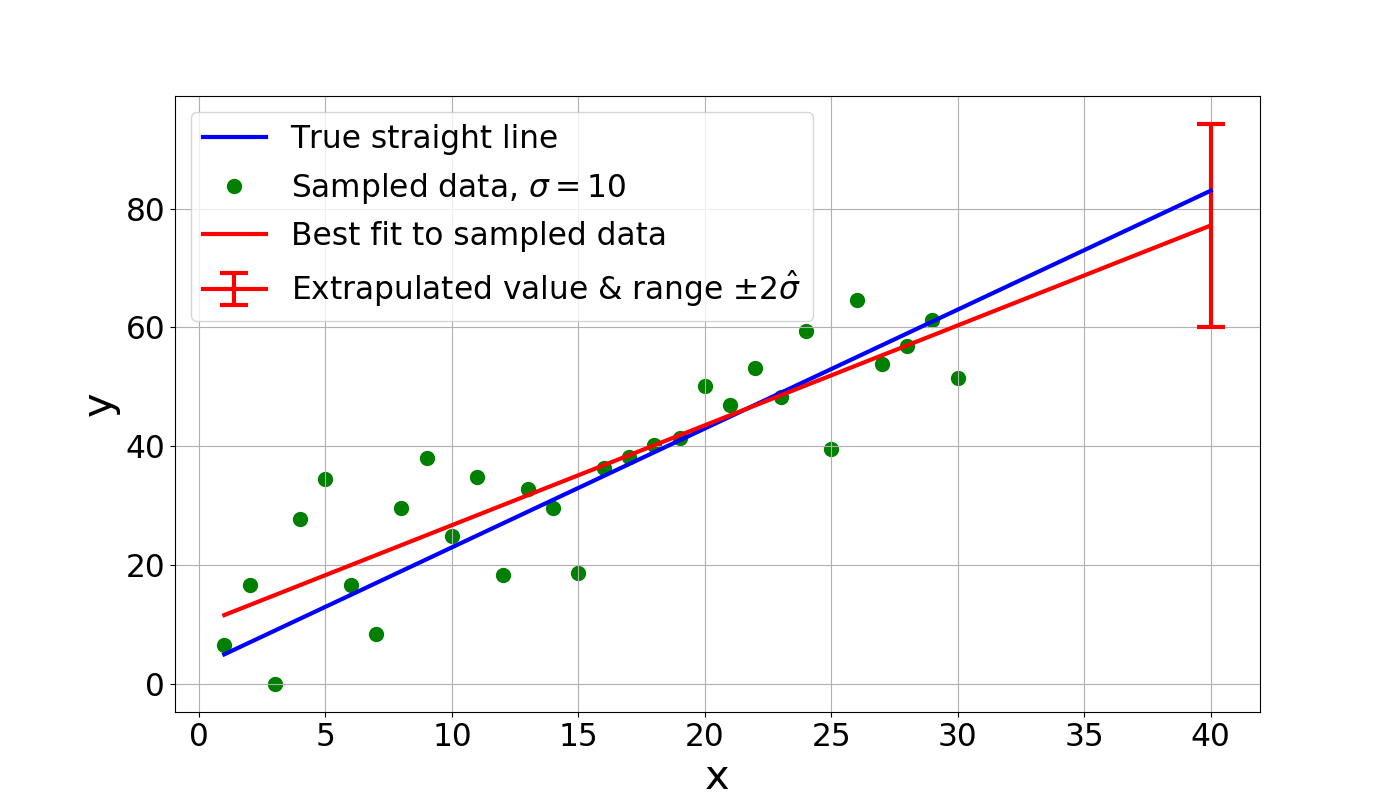


Figure 2: Extrapolating and constructing some range for from the sample in the range x=1…30