

Week 8 Reading Questions

Mercy Melo (no collaborators)

Question 1: Describe the key difference between the non-parametric model (Ch. 7.1) and the parametric model (Ch. 8.1)

Non-parametric models do not rely on the data having any sort of known probability distribution underlying the stochastic portion of the model. This allows the sample data to be analyzed without being explicitly linked to the overall population, but instead rely exclusively on the sample data to draw inferences from. Due to this, we only need to specify a deterministic model when trying to fit our data (because there is no underlying stochastic model). Once we've fitted our data to a deterministic model, we can then evaluate error using methods such as ordinary least squares, which are based solely on the sample data's fit to the specified model. Resampling using either Monte Carlo or Bootstrapping allows us to gain a deeper understanding of our sample and the model's fit. While this allows us to draw assumptions directly from the sample, it is only advantageous to use if we do not want to make estimations about the true population or if we are unsure of the underlying probability distribution of the stochastic portion of our population's data. Parametric models allow us to delve deeper into our datasets and estimate population parameters without resampling. If we are able to assess our sources of uncertainty in the data, and determine the uncertainty's distribution, it can be more ecologically-relevant to use parametric methods that include the natural stochasticity of the data.

Question 2: Interpolation and extrapolation may both be used to make predictions. What is the difference between interpolation and extrapolation?

Interpolation is the prediction of values within the measured range while extrapolation is the prediction of values outside the measured range. Let's say we have a sample of plant growth based on the amount of water provided to that plant and it looks something like this:

Water Added	Plant Growth Rate
0.0mL	0 cm/day
1.0mL	1 cm/day
1.5mL	1.5 cm/day
2.0mL	2.0 cm/day

If we wanted to predict how fast our plant would grow if it were given 1.2mL of water, that would be an example of interpolation because we have already measured the response to 1.0mL and 1.5mL. If we wanted to predict how fast our plant would grow if it were given 4.0mL of water, that would be an example of extrapolation because 4.0mL is outside of our range of measured values.

Question 3: Explain why extrapolation has more pitfalls than interpolation.

Extrapolation can be risky as the model may change shape as the range of measurements changes. What may look like a linear relationship within the measured range of values could turn out to be a limited view of a truly exponential or sigmoidal curve when zoomed out to a larger range of values. To use the previous example, let's say plant growth naturally levels off at 2.0mL added water and the plant steadily grows at a rate of 2.0cm/day despite any additional water. If this is the case (and we did not know of this

interaction previously), we would be off-base if we estimated plant growth with 4.0mL per day using the current model (we would likely estimate 4.0mL of water to elicit the plant to grow 4.0cm/day).