Sorites Experiment 2

Contents

Description of experiment

A copy of the experiment can be found at: experiments/experiment2/insurance-inventory.html.

This was the first prior elicitation experiment that I worked on, which we ran in October of 2013. For 10 categories of items ("watch", "coffee maker", "laptop", "electric kettle", "headphones", "sweater", "camera", "backpack", "wallet", and "bike"), we asked participants to fill in guesses about the prices of those items. We gave them one piece of information for each item: whether it cost more or less than a particular cutoff (50, 100, 500, 1000, 2000, 5000, or 10000). The cutoffs were varied between participants. Each participant saw every item twice, once above the cutoff and once below.

There are two different data files for this experiment, because when I put the experiment up on mTurk the first time, I gave it the tagline "Insurance Inventory", which sounds really boring. People signed up, but really slowly. So after a day passed and I had only gotten 18 participants, I re-posted with a new tagline: "Store Rescue". 3 participants were repeats. Once we excluded repeats, there were 105 participants in the experiment.

Results

Lots of participants in high cutoff conditions commented that it was almost completely impossible for a store to have items of those categories above the cutoffs.

Here are the results for item="sweater" and cutoff="1000".
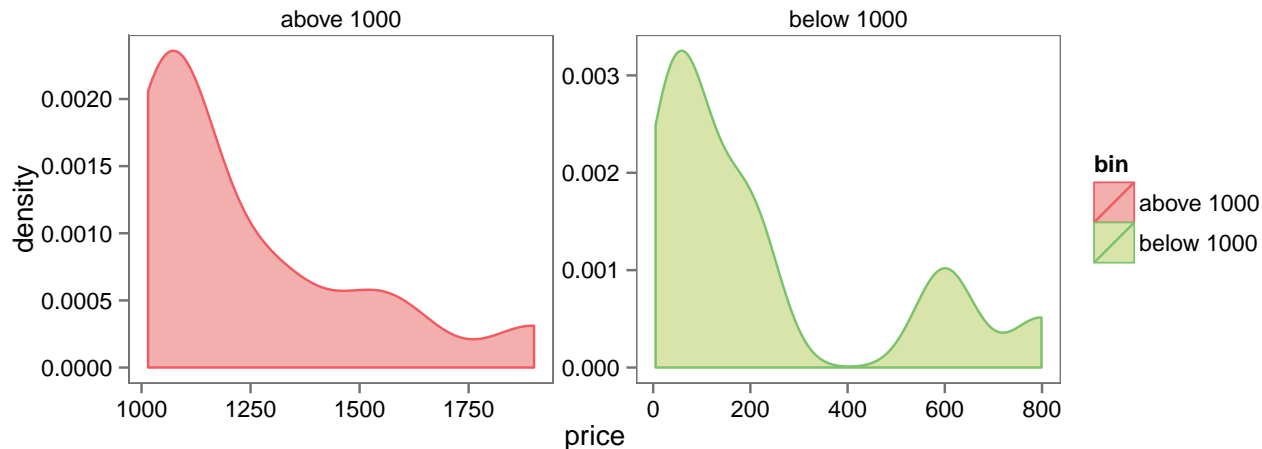
Prices below a particular range

When the cutoff was (e.g.) 50, of course all of the prices participants gave were at or below 50. When the cutoff was higher, the proportion at or below 50 decreased.
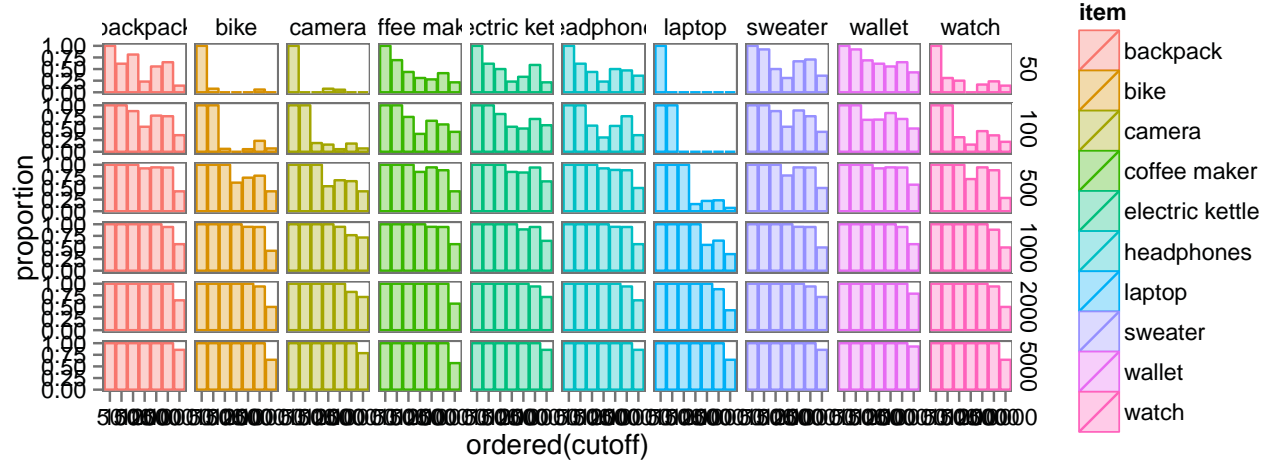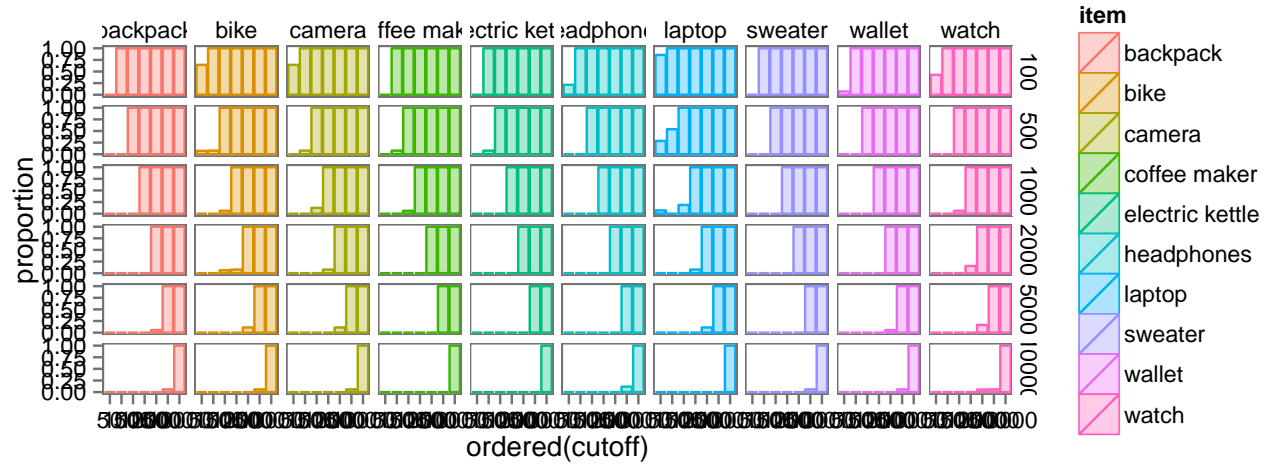

Prices below a particular range

When the cutoff was (e.g.) 50, of course all of the prices participants gave were at or below 50. When the cutoff was higher, the proportion at or below 50 decreased.

We were planning to stitch these binned free-responses together, but we ended up abandoning this strategy because a) although these questions get us better *resolution* in the tails, we still don't necessarily have a good sense of the probability *mass* in the tails, and b) the fact that participants gave mostly round numbers (which mess with the density estimators) and thought the cutoffs completely improbable for some of the items and said they had to really re-imagine the category (so it's somewhat possible that we're not even sampling from the same distribution when the cutoff is very high) makes these data harder to analyze.

The basic idea for how to analyze these data was to estimate p_k, the amount of probability mass above cutoff c_k and below cutoff c_{k+1}, by looking at the proportion of samples within that range when the cutoff was something higher or lower, e.g. $p_0 \approx (1 - \sum_{i=k}^{n} p_i)p_0^{(k)}$ (where $p_0^{(k)}$ refers to the proportion of samples between c_0 and c_1 when the cutoff in the experiment was c_k).
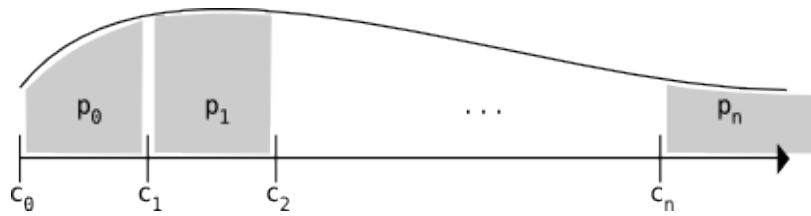
There are more resulting equations than unknowns, which means we would need to minimize the error. (Something I never actually did.) This seems do-able, but time-consuming, and we have the binned data which seems more sensible in several ways. I might come back to this later, but for now, I'll leave it unfinished.