

## **Homework #4**

(Due Monday, February 1<sup>st</sup>)

### **Completing an Incomplete Study**

Roughly design a study to determine the number of something using incomplete data. Your study should identify at least three biases, limitations and/or complications of your data and explain how you would try to estimate the effect of and then correct for each one. You should utilize the method of similar fractions somewhere in your analysis.

*A detailed example of this kind of study is included below.*

- a. Explain in 2-3 sentences what your study is aiming to measure, predict or estimate and why you (or someone) would care to know the answer.
- b. List what data you would collect/be able to get with a brief explanation of each (if needed). Then either collect that data. Or, if that's not feasible, make up some fake data for analysis.
- c. Identify at least three biases, limitations and/or complications of that data and give a brief explanation of each.
- d. For each of these biases, limitations and/or complications explain how you would try to estimate the effect of each one and correct your result to account for it. Again, make up numbers if you need to for this correction.
- e. Use the corrected numbers to estimate the number you were aiming at.

Example Study:

## Dog-sitting Feasibility Estimation...

a. Basic Description and Motivation:

My study aims to estimate how much money I could make if I was trying to start a dog-sitting service. I'm interested in this because I'm interested in making some extra money and I really like dogs.

b. Data:

- Population of Santa Cruz county = **253,137**
  - As of the 2008 census (as passed down to us by the all knowing Google)
- Average number of people who own dogs in Santa Cruz = ???
  - I could sit on a bench on Westcliff and count how many people who walk by have dogs versus how many don't (over 10 minutes)  
**30 with dogs of 75 people total** (*I made these numbers up*)
- Typical cost of dog-sitting services = **\$25/dog**
  - I looked this up on websites around Santa Cruz and it looked like \$25 was an average number.
- How often people go away without their dog = ???
  - I'd estimate that people go away for vacations and random other stuff about **once every three months**.

c. Biases, limitations and complications:

- 1 It looks like the Santa Cruz population is slowly growing so the population is probably a bit higher now than it was in 2008.
- 2 The numbers I counted on Westcliff is probably biased (too high) because people with dogs would probably go on more walks than people without dogs.
- 3 I probably couldn't charge \$25/dog because I'm not as hardcore as the sites that I looked at to get that price (I don't have fancy rooms for the dogs and stuff).

d. Corrections:

- 1 Google had a graph which made it look like the SC population is leveling out, so I think this is a small correction – I'd round the population up to **260,000**.
- 2 I'd guess that this is a fairly big effect, so I'd knock the number of people I saw on Westcliff with dogs down by half, to **15 people with dogs**.
- 3 I'd start off with a low price hoping to generate repeat customers so I would knock my price down to **\$15/dog**.

e. And the answer is:

Given the above (corrected) data, we can use the method of similar fractions to determine the approximate number of dog owners in Santa Cruz:

$$\frac{N_{DogOwners}}{Pop_{SantaCruz}} = \frac{People\ with\ Dogs}{Total\ People} = \frac{15}{75} = 0.20$$

$$\begin{aligned}\rightarrow N_{DogOwners} &= Pop_{SantaCruz} \times 0.20 \\ &= 260,000 \times 0.20\end{aligned}$$

$$\rightarrow N_{DogOwners} = 52,000$$

Next we can use this to calculate how many people will use us, assuming we only get 1/10 of the business in competition with the other dog-sitters:

$$N_{Customers} = N_{DogOwners} \times \frac{1}{10} = 5,200$$

Next we can change this into how often we'll have customers... assuming each of these people leave town once every three months we'll have 5,200 dogs every ~90 days (3 months). That's 5,200dogs / 90days = 58 dogs per day. That's not realistic (I can't take care of that many dogs) BUT it looks like there's enough people and dogs to support my business. If I could handle that many dogs I'd make about:

$$58 \text{ dogs/day} \times (\$15/\text{dog}) = \$870 \text{ per day (unrealistic)}$$

So what if people only use a professional dog-sitter 1/10 of the time? That's still 5.8 (~6) dogs every day, which is probably still more than I would want to handle. So the amount of money I'd make is probably limited by how many dogs I'm willing to take care of (about 3 per day):

$$3 \text{ dogs/day} \times (\$15/\text{dog}) = \$45 \text{ per day (more realistic)}$$