In-Class Problems Week 13, Wed.

Problem 1. Explaining Sampling to a Jury

In last lecture you learned why sampling 589 fish (or voters) will yield a fraction that, 95% of the time, will be within 0.04 of the actual fraction of contaminated fish (or voters who prefer Clinton over Giuliani). Notice that the size of the fish or voting population was never considered because *it did not matter*.

It seems remarkable that, whether there are a thousand, a million, or a billion voters in a country, polling only a few hundred is sufficient to be confident of an accurate estimation of average voter preference. Suppose you were going to serve as an expert witness in a trial. How would you explain why the number of people necessary to poll *does not depend on the population size*? Remember that juries do not understand formulas, so you have to provide an intuitive explanation, which is not quantitative.

Problem 2. An *International Journal of Epidemiology* has a policy that they will only publish the results of a drug trial when there were enough patients in the drug trial to be sure that the conclusions about the drug's effectiveness hold at the 95% confidence level. The editors of the Journal reason that under this policy, their readership can be confident that at most 5% of the published studies will be mistaken.

Later, the editors are astonished and embarrassed to learn that *every one* of the 20 drug trial results they published during the year was wrong. This happened even though the editors and reviewers had carefully checked the submitted data, and have no doubt that every one of the trials was *properly performed and reported* in the published paper.

The editors thought the probability of this was negligible (namely, $(1/20)^{20} < 10^{-25}$). Explain what's wrong with their reasoning and how it could be that all 20 published studies were wrong.

Problem 3. Here is a fun game. You pick a number between 1 and 6. Then you roll three fair, independent dice.

- If your number never comes up, then you lose one dollar.
- If you number comes up once, then you win one dollar.

- If your number comes up twice, then you win two dollars.
- If your number comes up three times, then you win k dollars.

(a) Compute your expected payoff as a function of k.

(b) For what value of k is this game fair?

Problem 4. Find the expectation of a variable, J, with an (n, p)-binomial distribution:

$$\mathrm{PDF}_J(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Hint: Consider

$$\frac{d\ (x+y)^n}{dx}$$

Problem 5. A couple decides to have children until they have both a boy and a girl. What is the expected number of children that they'll end up with? Assume that each child is equally likely to be a boy or a girl and genders are mutually independent.

Appendix

The ***expected value*** of a random variable R defined on a sample space, $\mathcal{S}$, is:

$$\mathrm{E}[R] = \sum_{w \in \mathcal{S}} R(w) \Pr\{w\}$$

Another helpful formula for expected values is:

$$\mathrm{E}[R] = \sum_{x \in \text{range}(R)} x \cdot \Pr\{R = x\}$$

Mean Time to Failure: If a biased coin with probability, p, of Heads is repeatedly flipped until a Head comes up, where the flips are mutually independent, then the expected number of flips is $1/p$.