

Solutions to HW #7

Non-textbook problems

1.
 - a. cluster sampling (randomly pick the clusters, then sample all items in each cluster)
 - b. convenience sampling (note: no element of randomization here)
 - c. systematic sampling
 - d. stratified random sampling (take a simple random sample from EVERY subgroup)
 - e. simple random sample
 - f. simple random sample
2.
 - a. proportion of miscarriages for glycol ether group= $10/30=0.333$
proportion for non glycol ether = $120/750=0.16$
Glycol ether group had higher proportion of miscarriages.
 - b. No. We can only prove cause and effect in a designed experiment where subjects are randomly assigned to treatment and control groups. This is an observational study so we cannot conclude cause and effect on the basis of this study alone.
 - c. We would have to randomly assign pregnant women to be exposed to a potentially toxic chemical. We might then be causing them to have miscarriages, which would be completely unethical.
3.
 - a. Observational study. No treatment was imposed, data were simply collected.
 - b. No, cause and effect cannot be established by an observational study. A designed experiment is required to establish cause and effect. Since this is an observational study, we cannot conclude that money *causes* happiness.
4. Observational study. No treatment was imposed.
5.
 - a. Designed experiment.
 - b.
 - i. Yes, on p.4 under “Randomization and Masking” the authors indicate computer-generated randomization was used. It appears subjects were stratified by age, sex and study center.
 - ii. A placebo pill was used. You can’t use a placebo for cognitive behavioral therapy (counseling) so there was no placebo for that component of treatment.
 - iii. The study was partially double-blind. The evaluators were not aware of the treatment group assignments. The subjects in the placebo and sertraline groups were not aware whether they were on the placebo or real drug. However, the subjects in the combined treatment group knew they were taking sertraline. All subjects knew whether or not they were receiving counseling.
6.
 - a. A peer-reviewed journal means the articles are reviewed by referees who are experts in the subject matter of the article. The referees are supposed to decide if the article contains results that are important enough to be published. The peer-review process is also designed to catch any potential errors in the research before it goes to publication.
 - b. John Doe can post anything, true or not, on his website. There is no review process to check the accuracy of the information. Although mistakes may (occasionally) appear in peer-reviewed journals, the fact that these findings are scrutinized by other experts in the same field makes them more trustworthy.

- 7.10** **a** Since the question is particularly sensitive to people of different ethnic origins, you may find that the answers may not always be truthful, depending on the ethnicity of the interviewer and the person being interviewed.
- b** Notice that the percentage in favor of affirmative action increases as the ethnic origin of the interviewer changes from Caucasian to Asian to African-American. The people being interviewed may be changing their response to match what they perceive to be the response which the interviewer wants to hear.

- 7.20** **a** If the sampled populations are normal, the distribution of \bar{x} is also normal *for all values of n* .
- b** The Central Limit Theorem states that for sample sizes as small as $n = 25$, the sampling distribution of \bar{x} will be approximately normal. Hence, we can be relatively certain that the sampling distribution of \bar{x} for parts **a** and **b** will be approximately normal. However, the sample size is part **c**, $n = 8$, is too small to assume that the distribution of \bar{x} is approximately normal.
- 7.26** **a** Since the sample size is large, the sampling distribution of \bar{x} will be approximately normal with mean $\mu = 64,571$ and standard deviation $\sigma/\sqrt{n} = 4000/\sqrt{60} = 516.3978$.
- b** From the Empirical Rule (and the general properties of the normal distribution), approximately 95% of the measurements will lie within 2 standard deviations of the mean:

$$\begin{aligned}\mu \pm 2SE &\Rightarrow 64,571 \pm 2(516.3978) \\ &64,571 \pm 1032.80 \text{ or } 63,538.20 \text{ to } 65,603.80\end{aligned}$$

- c** Use the mean and standard deviation for the distribution of \bar{x} given in part **a**.

$$\begin{aligned}P(\bar{x} > 66,000) &= P\left(z > \frac{66,000 - 64,571}{516.3978}\right) \\ &= P(z > 2.77) = 1 - .9972 = .0028\end{aligned}$$

- d** Refer to part **c**. You have observed a very unlikely occurrence, assuming that $\mu = 64,571$. Perhaps your sample was not a random sample, or perhaps the average salary of \$64,571 is no longer correct.