18.05 Problem Set 9, Spring 2014

Problem 1. (10 pts.) **Confident coin.**
Recall the confident coin from pset 7. It was spun on its edge 250 times and came up heads
140 times. Our null hypothesis was that the coin was fair and the two-sided p-value came
out as 0.066.

(a) In this problem we want you to make one plot for each of

(i) $\theta = 0.5$, (ii) $\theta = 0.3$, (iii) $\theta = 0.1$.

Your plot should include:

- the pmf $p(x|\theta)$ of the binomial$(250, \theta)$ distribution
- the pdf of the normal distribution with the same mean and variance
- the normal distribution with the same mean (as the binomial distribution) and the rule-of-thumb variance to your plot.

The rule-of-thumb variance for a Bernoulli(θ) distribution is $1/4$, i.e. we use the variance
for $\theta = 0.5$ no matter what the true value of θ is. So the rule-of-thumb variance for a
binomial(n, θ) is just $n/4$.

Add the graph of In each case, how close are each of the normal distributions to the binomial
distribution? How do the two normal distributions differ? Based on your plots, for what
range of θ do think the rule-of-thumb normal distribution is a reasonable approximation for
binomial(n, θ) with large n?

(b) Find 80% and 95% confidence intervals for θ using the polling rule-of-thumb.

(c) Find an 80% **posterior** probability interval for θ using a flat prior, i.e. beta$(1, 1)$.
Center your interval between the 0.1 and 0.9 quantiles. Hint: Use `qbeta(p, a, b)`.

Problem 2. (10 pts.) **Height.**
Suppose μ is the average height of a college male. You measure the heights (in inches) of
twenty college men, getting data $x_1, \ldots, x_{20}$, with sample mean $\bar{x} = 69.55$ in. and sample
variance $s^2 = 14.26$ in^2. Suppose that the x_i are drawn from a normal distribution with
unknown mean μ and unknown variance σ^2.

(a) Using $\bar{x}$ and s^2, construct a 90% t–confidence interval for μ.

(b) Now suppose you are told that the height of a college male is normally distributed
with standard deviation 3.77 in. Construct a 90% z–confidence interval for μ.

(c) In (b), how many people in total would you need to measure to bring the **width** of
the 90% z–confidence interval down to 1 inch?

(d) Consider again the case of unknown variance in (a). Based on this sample variance
of 14.26 in^2, how many people in total should you expect to need to measure to bring the
width of the 90% t–confidence interval down to 1 inch? Is it guaranteed that this number
will be sufficient? Explain your reasoning.

Problem 3. (10 pts.) **Soda.**
Consider a machine that is known to fill soda cans with amounts that follow a normal
distribution with (unknown) mean μ and standard deviation $\sigma = 3$ mL. We measure the

volume of soda in a sample of bottles and obtain the following data (in mL):

$$352, 351, 361, 353, 352, 358, 360, 358, 359$$

(a) Construct a precise 95% confidence interval for the mean μ.

(b) Now construct a 98% confidence interval for the mean μ.

(c) Suppose now that σ is not known. Redo parts (a) and (b), and compare your answers to those above.

Problem 4. (10 pts.) **Cholesterol.**
In a study on cholesterol levels a sample of 12 men and women was chosen. The plasma cholesterol levels (mmol/L) of the subjects were as follows:

$$6.0, \ 6.4, \ 7.0, \ 5.8, \ 6.0, \ 5.8, \ 5.9, \ 6.7, \ 6.1, \ 6.5, \ 6.3, \ 5.8.$$

(a) Estimate the variance of the plasma cholesterol levels with a 95% confidence interval.

(b) What assumptions did you make about the sample in order to make your estimate?

Problem 5. (10 pts.) **Candy.**
A large candy manufacturer produces packs of candy targeted to weigh 52 grams. A quality control manager was concerned that the variation in the actual weights was larger than acceptable. That is, he was concerned that some packs weighed significantly less than 52-grams and some weighed significantly more than 52 grams. In order to estimate σ, the standard deviation of the weights of the nominal 52-gram packs, he took a random sample of n = 10 packs off of the factory line. The random sample yielded a sample variance of 4.2 grams.

(a) Use the random sample to derive a 95% confidence interval for σ. *Note: the interval is for σ not σ^2.*

(b) What assumptions did you make about the sample in order to make your estimate.

18.05 Introduction to Probability and Statistics

Spring 2014