Homework 5 Solutions

http://www.stat.ucla.edu/~dinov/courses_students.dir/08/Fall/STAT13.1.dir

1. Problem 1

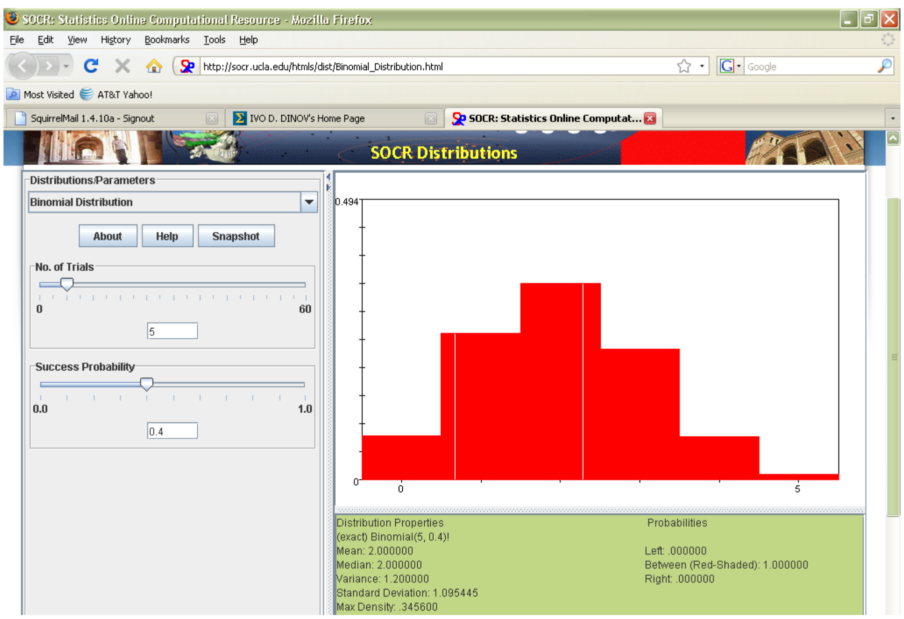
(a)

Let X = # of individuals that are mutants
n = 5
p = 0.40

p(p^=0.0) = p(X=0) = **0.08**
p(p^=0.2) = p(X=1) = **0.26**
p(p^=0.4) = p(X=2) = **0.35**
p(p^=0.6) = p(X=3) = **0.23**
p(p^=0.8) = p(X=4) = **0.08**
p(p^=1.0) = p(X=5) = **0.01**

(b)

Figure 1

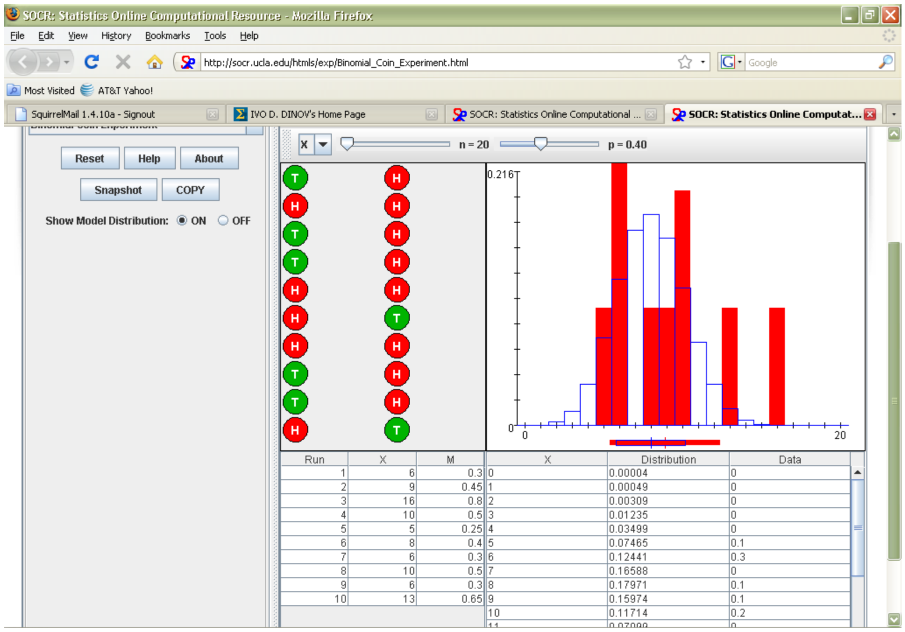


(c)

****Results will vary****

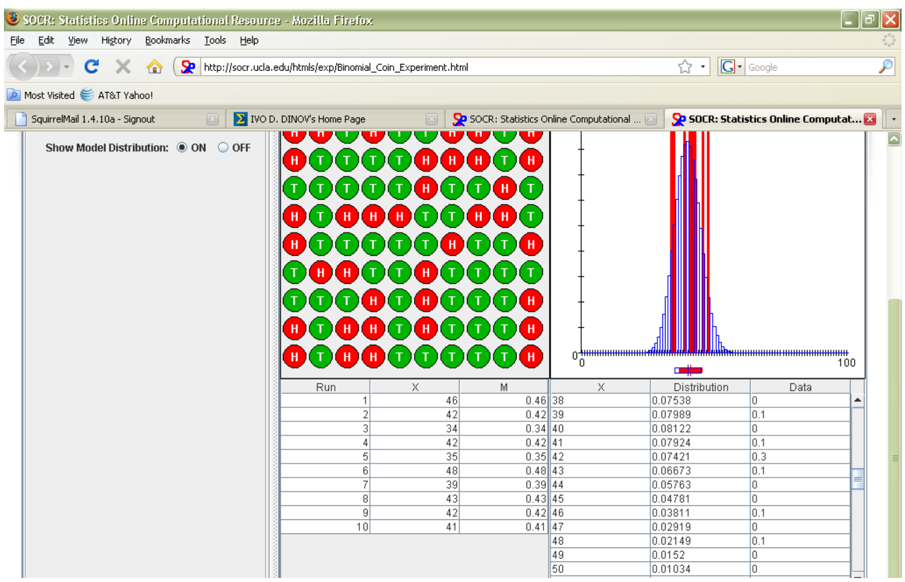
- The empirical probabilities are close for some values of X, but not very close for X=6, 7, 10, 13, and 16 (*see figure 2*).

Figure 2



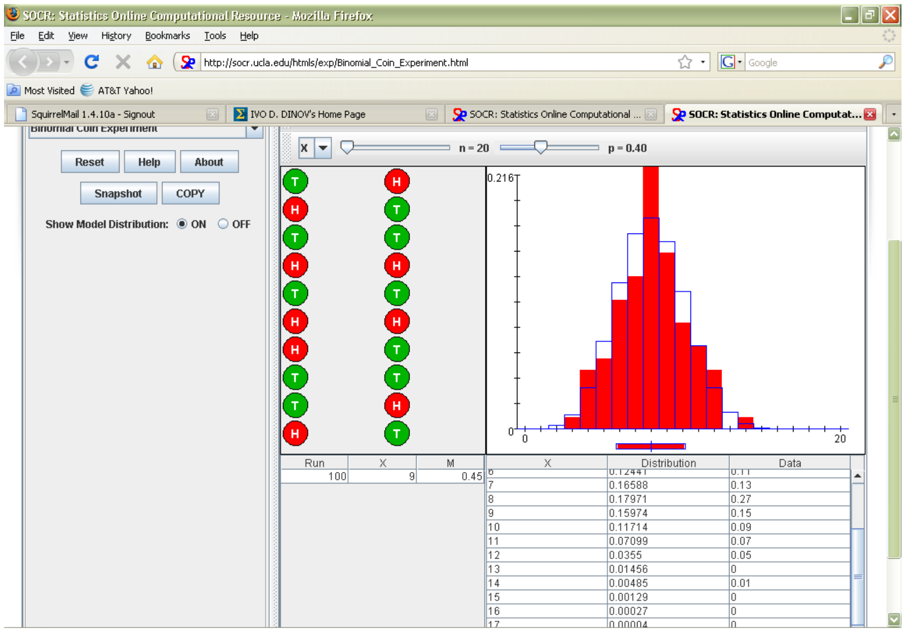
- The empirical and theoretical probabilities are expected to be similar. If we increase the number of trials without increasing the number of experiments, the probabilities and theoretical probabilities will **not** be more similar (*see figure 3*).

Figure 3



- The empirical and theoretical probabilities should become **more** similar if we run 100 experiments with 20 trials (*see figure 4*).

Figure 4



2. Problem 2

We know from the Central Limit Theorem that the sample averages should follow approximately a Normal Distribution with mean μ and variance σ^2/n

- We want to know the probability that the sample average will be between 2900 and 3100, with a sample size of n=15. We can use the SOCR Normal Distribution with mean = 3000 and sd = 400/√15 = 400/3.873 = 103.28 to find **p(2900 <= sample average <= 3100) = 0.67**
- Now, we have n=60, so we can use the SOCR Normal Distribution with mean=3000 and sd=400/√60 = 51.64 to find **p(2900 <= sample average <= 3100) = 0.95**
- The P(E) increases with an increase in the sample size. As n gets larger, the sample averages are more precise with less variation and so the probability that the sample average is close to the true mean increases.

3. Problem 3

- For n=1, we want the proportion to be closer to ½ than 9/16, so we want to know the probability that zero of the progeny are purple flowered. This is simply given by 1-9/16 = 7/16 = **0.4375**

- For n=64, we want the proportion to be closer to ½ than 9/16, so we want the number of purple flowered progeny to be closer to 32 than to 36. This happens if we have 33 purple flowered progeny at the most. So if X=# of purple flowered progeny, then we want to know p(X<=33) using the normal approximation.

$$z = \frac{33 - 64(0.5625)}{\sqrt{64 * 0.5625 * 0.4375}} = -0.76$$

Now, using the Normal distribution in SOCR (or Z-table) we can find **p(z<= -0.76) = 0.2236**

- For n=320, we want the number of purple flowered progeny to be closer to 160 than to 180. This happens if we have 169 purple flowered progeny at the most. So, if X=# of purple flowered progeny, then we want to know p(X<=169) using the normal approximation.

$$z = \frac{169 - 320(0.5625)}{\sqrt{320 * 0.5625 * 0.4375}} = -1.24$$

Now, using the Normal distribution in SOCR (or Z-table) we can find **p(z<= -1.24) = 0.1075**

4. <u>Problem 4</u>

The sample average weight is given by $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, so the total weight is just $n * \bar{x}$. If we know that n=10, and we are interested in the percentage of litters that have a total weight of 90g or more, than we know the average weight must be at least 9g

$$90g = 10 * \bar{x} \Rightarrow \bar{x} = 9g$$

We know that the sample average follows a Normal distribution, with mean = 8.3g and std. error = 1.7/√10 = 0.5376. Using the Normal Distribution in SOCR we have **p(X>=9g) = 0.096**
So, about **9.6%** of the litters will have a total weight of 90g or more.

Alternative solution:
We could have also used a Normal distribution with mean = 83g and std. dev = 5.376 and we would find **p(X>=90g) = 0.096**

5. <u>Problem 5</u>

- 95% CI

$$95\% CI = 28.7 \pm t_{0.025, df=5} * \frac{4.6}{\sqrt{6}} = 28.7 \pm 2.571 * \frac{4.6}{\sqrt{6}} = 28.7 \pm 4.83$$

$$= (23.87, 33.53)$$

- The population mean that we estimated in this case is the mean blood serum concentrations of Gentamicin of the entire population of healthy three year old Suffolk sheep 1.5 hours after being injected with the antibiotic Gentamicin at a dosage of 10 mg/kg body weight.

- It is not necessarily typical for the 95% confidence interval to nearly contain all of the values.

6. **<u>Problem 6</u>**

- 95% CI

$$95\%CI = 13 \pm t_{0.025,df=9} * \frac{12.4}{\sqrt{10}} = 13 \pm 2.262 * \frac{12.4}{\sqrt{10}} = 13 \pm 8.87$$

$$= (4.13, 21.87)$$

- The confidence interval above tells us that we can be very confident that the true mean difference in HBE levels of the entire population lies somewhere between 4.13 and 21.87 pg/mLi. For example, if we were to conduct this experiment 100 times, and calculate 100 different sample averages and 100 different 95% confidence intervals, we would expect about 95 of those confidence intervals to capture the true mean difference in HBE levels, and only about 5 of the confidence intervals will completely miss the true mean. Here, we only do it once, but we should feel very confident that the one confidence interval we have captures the true mean.