R3

Xuanxi Zhang

1

Consider the following letter-envelope matching problem. There are n letters labeled 1 through n and also n envelopes labeled 1 through n. The letters are permuted randomly into the envelopes (possibly by a drunk person), one letter per envelope, so that all permutations are equally likely. Let X_n be the number of letters that are placed into their matching envelopes.

1. Find the probability $\mathbb{P}(X_n = 0)$. (Hint: inclusion-exclusion principle)
2. Find the limiting distribution of X_n as $n \to \infty$. Is it a normal distribution?

For reference, here is the inclusion-exclusion principle we proved last class:

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{I \subseteq \{1,2,\cdots,n\}:|I|=k} \mathbb{P}\left(\bigcap_{i \in I} A_i\right)$$

2 7.3

Which of the following is a random variable?

1. population mean
2. population size, N
3. sample size, n
4. sample mean
5. variance of sample mean
6. largest value in sample
7. population variance
8. estimated variance of sample mean

3 7.10

True or False? If a sample from a population is large, a histogram of the values in the sample will be approximately normal, even if the population is not normal.

4 7.9

In a random sample of 1,500 voters, 55% said they planned to vote for a certain proposition, and 45% said they would vote against it. The estimated margin of victory for the proposition is thus $55\% - 45\% = 10\%$. What is the standard error of this estimated margin? What is a 95% confidence interval for the margin? (You'll need a calculator or a device that can serve as one.)

5 True or False

1. Which is more likely, 9 heads in 10 tosses of a fair coin, or 18 heads in 20 tosses? (You don't need to calculate the exact probabilities.)
2. In order to halve the width of a 95% confidence interval for a mean, by what factor should the sample size be increased?
3. True or False? A 95% confidence interval contains the sample mean with probability 0.95 .

4. Suppose you survey 100 people and ask them to rate their opinion of a product from -5 to 5 , then average their ratings and divide by the standard error. If the population mean of the ratings is 0 , approximately what type of distribution does the square of this scaled average follow?

5. Suppose we measure the heights of 10 NYU students. If we wanted to find a confidence interval for the average height of an NYU student, we'd look in a table for the CDF of what distribution?