

CS146 session 6-2 make-up work

Pre-Class:

<https://github.com/Boyu1997/CS146/blob/master/Session-6-2/Session-6-2-pre-class.ipynb>

Class summary:

Assessment-poll:

In the assessment poll, we were asked in central limit theory apply in the ____ cases. Connor and Evgeny gives two different representations: Connor say it does not apply if we consider the entire experiment as a whole binomial distribution, $n=1$ so the central limit theory does not apply. Evgeny gives the idea if we consider the each patient as independent bernoulli trails from the same distribution, $n=s+f$ and we have central limit theory. The key distraction is if we can consider each individual patients are independent and have identical probability distribution. Rita also raise the question in the case $n>1$ but relatively small, if central lime theory apply.

Activity 1:

pdf of the binomial distribution:

$$\binom{n}{k} p^k (1 - p)^{n-k}$$

mean and variance of the binomial distribution:

Mean: np

Variance: $np(1-p)$

Activity 2:

We discussed the code in the pre-class work. We can do a sample, or as we are doing normal approximation, it is not necessary to generate sample as we can simply use the 1.96std from the mean from the mean. As we do not need to generate sample, there is no sampling error. Grope D produce a table showing the different results of running the code multiple times produce different result, and the centurial limit theory will not product a constant result and as we get to smaller data size, we are expecting more noise in the result.

Reflection Poll:

It is important to describe clearly how the result is mathematically derived as there are potentially different way of doing it can produce relatively different result, like as we discussed in activity 2, and we want to avoid situation like in the "Total Survey Error" pater that when there is a disagreement we reproducing the paper, we are uncertain if it is because metrology difference or there is a mistake in either the author or our calculation.