All Central Limit Theorem Problems Are the Same

Tianchen Liu

Here we are, Central Limit Theorem, the grand finale of the course (Markov Chains taught in CS70 is like a sweet little encore). I was confused when I first learned it, so I would like to use this note to show some insights I have on CLT and how to crack CLT problems.

1 CLT: ELI5

The definition and math of CLT is written perfectly clear on the notes. I'm here to give a explained-like-I'm-five version:

If I repeat the same (*i.i.d.*) r.v. a bunch of times, no matter what the distribution behind the r.v. is, the r.v. that represents the mean/summation of all the *i.i.d.* r.v.'s is going to look like (converge to) a normal distribution.

2 All CLT Problems Are the Same

In math language, the statement above would be

$$\frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} \approx \mathcal{N}(0,1)$$

S_n is the summation of all the i.i.d. r.v.'s. $\frac{S_n}{n}$ would therefore denote the sample mean.

All CLT problems are essentially the same: Solve for one variable with all other variables given.

- Given a probability/confidence and number of experiments, what's the (widest) confidence interval?
  - You run a casino that has a machine which rolls 1000 dice simultaneously. A customer wins 3x their ticket price if the sum of all dice values are more than v away from 350, otherwise the customer gets nothing. If you want to make money on the long run, what's the constrain of v? *Remark: The probability/confidence is given implicitly in this problem.*
- Given a confidence interval and number of experiments, what's the probability/confidence?
  - I flip a fair coin 1000 times, approximate the probability that I get at most 510 heads using CLT. *Remark: this problem doesn't have a "confidence interval" but you know what to plug into the* Φ *function.*
  - Spring 2020 Midterm 3 (Final) 6(c).
- Given confidence interval and probability/confidence, how many experiments are needed?
  - *Sp19 Final 7(3)* The number of people coming to a McDonald's per hour follow a poisson distribution, with parameter at most 10. How many hours do I need to sit in a McDonald's so that I can construct a 95% confidence interval with width 2?
  - Sp18 Final 2(e).

Whenever you see a CLT problem, first categorize the problem. Everything should be plug-and-chug from there.