Population v.s. Sample

`Def:` Population is all the observations.

Sample is some observations selected from population.

Center Limit Theorm (CLT)

If you have a sequence of i.i.d r.v X_n,

$$\lim_{n\to\infty} \frac{\sum X_n}{n} \sim N(\mu, \frac{\sigma^2}{n})$$

Standard deviation of different cases

Suppose we have a discrete R.V. X with probability distribution $P(X)$, mean value μ, SD σ,

- Population SD, by definition

$$\sigma = \sqrt{\sum (x-\mu)^2 P(X=x)}$$

- Sample SD, n is the number of samples

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}}$$

  Example for lab problem 6:
  $\sqrt{((1-1.5)^2+(2-1.5)^2)/(n-1)}$, where n=2

- SD for Sample mean

  By CLT, is $\sigma/\sqrt{n}$

- SD for Sample proportion Special case of previous formula.

  $X \sim Bernouli(p)$, $P(X=1)=p$, then $SD_X = \sqrt{p(1-p)}$

Z-score

$X \sim N(\mu, \sigma^2)$, what's $P(X >= 4)$?

`Def`: $z = \frac{x-\mu}{\sigma}$, evaluates how far a value from mean for a normal distribution.

As for our problem 10, $Z = \frac{4-3.71}{0.283}$