CLT: samples of size $n$ drawn from a population with mean $\mu$ and s.d. $\sigma$, then sampling distribution of $\bar{X}$ will be approx. Normal, with $\mu_{\bar{X}} = \mu$ and sd $\sigma_{\bar{X}} = \sigma / \sqrt{n}$.

Note: if population is normal, then $\bar{X}$ is exactly normal.

CLT examples

1) IQ scores are Normally distributed with mean 100, and s.d. 15 points.

a) Probability a randomly selected person will have a score below 97. $\Pr(X < 97)$

\[
\Pr(X < 97) = \Pr\left(\frac{X - \mu}{\sigma} < \frac{97 - \mu}{\sigma}\right) = \Pr\left(Z < \frac{97 - 100}{15}\right) = \Pr(Z < -0.2) = 0.4207
\]
b) Random sample of 100 people selected, what is prob mean score is less than 97?

$$\Pr(\bar{X} < 97) = \Pr \left( \frac{\bar{X} - \mu_x}{\sigma_x} < \frac{97 - \mu_x}{\sigma_x} \right)$$

$$\mu_x = 100, \quad \sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{100}}$$

$$= \Pr \left( Z < \frac{97 - 100}{15/\sqrt{100}} \right) = \Pr(Z < -2)$$

$$= 0.0228$$

2) Batch of seeds has mean time to germination of 10.4 days, with s.d. 2.13 days. What is prob a sample of 49 seeds will have mean germ. time between 10 and 11 days?

$$\Pr(10 < \bar{X} < 11) = \Pr \left( \frac{10 - \mu_x}{\sigma_x} < \frac{\bar{X} - \mu_x}{\sigma_x} < \frac{11 - \mu_x}{\sigma_x} \right)$$

$$= \Pr(-1.31 < Z < 1.97)$$

$$\mu_x = 10.4, \quad \sigma_x = 2.13, \quad \sqrt{49}$$
Normal Approximation to Binomial

Binomial distr.
1) fixed # trials
2) ind. trials
3) 2 poss. outcomes
4) constant probability

parameters: \( n \) and \( p \)

Approx. works when \( np \geq 5 \) and \( n(1-p) \geq 5 \). If this holds, then Binomial R.V. has distribution that can be approximated by Normal distribution, with mean \( \mu = np \), and s.d. \( \sigma = \sqrt{np(1-p)} \).
Binomial is discrete, normal is continuous. To overcome this, we use a "continuity correction." We move the count by \( \frac{1}{2} \), s.t. inequality still holds.

Ex: 62% of households have a computer. Random sample of 1000, what is the probability that at least 640 have a computer?

Idea: \( x \) = # households with a computer.

\[ X \sim \text{Bin}(n=1000, \ p=0.62) \]
\[ \rightarrow \mu = np = 620, \ \sigma = \sqrt{1000(0.62)(0.38)} \approx 15.35 \]

\( X \approx N(620 = \mu, \sigma = 15.35) \)

Continuity correction: replace discrete \( x \) by \( x \pm 0.5 \) in normal.
\textit{ex.} \quad \Pr(X \geq 640) = \Pr(X > 639.5) \\
\quad \text{because at least 640 includes 640.} \\
\quad = \Pr \left( \frac{X - 620}{15.35} > \frac{639.5 - 620}{15.35} \right) \\
\quad = \Pr(Z > 1.27) \\
\quad = 1 - \Pr(Z < 1.27) \\
\quad = 1 - .898 = .102

\textit{ex.} \quad \Pr(615 \leq X \leq 625) \\
\quad = \Pr(615.5 \leq X \leq 624.5) = \Pr \left( \frac{615.5 - 620}{15.35} \leq \frac{X - 620}{15.35} \leq \frac{624.5 - 620}{15.35} \right) \\
\quad = \Pr(-.29 \leq Z \leq +.29) \\
\quad = \Pr(Z < .29) - \Pr(Z < -.29) = .6141 - .3859 = .2282

\textit{ex.} \quad \text{Approx. prob. of getting a particular value} \\
\quad \Pr(X = 620) = \Pr(619.5 \leq X \leq 620.5) \\
\quad = \Pr \left( \frac{619.5 - 620}{15.35} \leq \frac{X - 620}{15.35} \leq \frac{620.5 - 620}{15.35} \right) \\
\quad = \Pr(-.03 \leq Z \leq .03) \\
\quad = \Pr(Z < .03) - \Pr(Z < -.03) \\
\quad = .512 - .488 = .024