October 3

Central Limit Theorem:

If $Y_1, \ldots, Y_n$ are iid draws from a distribution that is not normal, then the sampling distribution of $\overline{Y}$ is well approximated by a normal distribution with mean μ and

Variance $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n}$ when n is large. That is $\bar{Y} \sim N\left(\mu_Y, \frac{\sigma_Y^2}{n}\right)$ (approximately)

Properties of Y bar

$\overline{Y}$ is the blue estimator of μ_Y:

Best linear Unbiased Estimator

Unbiased: $E(\hat{\theta}) = \theta$, where θ is some population parameter.

$$E[\bar{Y}] = E\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i]$$

$$= \frac{1}{n} \cdot n \cdot \mu_Y$$

$$= \mu_Y$$

$$E(\bar{Y}) = \mu_Y$$

Linear: $\bar{Y}$ is a linear combination of the data

$$\bar{Y} = \frac{1}{n}(Y_1 + Y_2 + Y_3 + \ldots + Y_n)$$

Interval estimation:

CI

95% CI for $\bar{Y} \sim N(\mu, Var(\bar{Y}))$

$$CI_{95} = \{ \bar{Y} \pm 1.96 \; SE(\bar{Y}) \}$$

$$= (\bar{Y} - 1.96\, SE(\bar{Y}),\ \bar{Y} + 1.96\, SE(\bar{Y}))$$

Once the interval is estimated, the interpretation is that the CI estimator will contain μ for 95% of samples taken.

Exercise

$$SE(\bar{Y}) = \frac{S_Y}{\sqrt{n}} = \sqrt{\frac{S_Y^2}{n}} = \sqrt{100} = 10$$

$$CI = (980.4,\ 1019.6)$$