§5.4: The Distribution of "Sample Mean" $\bar{X}$ (and "Sample Sum" T_0)

Recall: "Sample Mean" (and "Sample Sum") is everyone's favorite statistic for IID $X_1, \ldots, X_n$.

→ Make n measurements of r.v. X to get independ. and identically distributed

$$X_1, X_2, \ldots, X_n$$

Def: Sample sum is $T_0 = X_1 + X_2 + \cdots + X_n$

Sample mean is $\bar{X} = \frac{1}{n} T_0$

$$= \frac{1}{n}(X_1 + \cdots + X_n)$$

Recall: If X & Y independent then

$$E[aX + bY] = aE[X] + bE[Y]$$

$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$$

Apply this to T_0 & $\bar{X}$:

Prop: If $\mu = E[X]$ and $\sigma^2 = Var[X]$ then

- $E[T_0] = n\mu$ and $Var[T_0] = n\sigma^2$
- $E[\bar{X}] = \mu$ and $Var[\bar{X}] = \sigma^2/n$

In particular, standard dev of $\bar{X}$ is $\sigma/\sqrt{n}$

→ We will see this a lot in the next chapter...

Little Theorem: If $X \sim$ Normal, then T_0 & $\bar{X}$ are Normal too
(in fact so is any linear combination of the X_k)

Big Theorem ("Central Limit Theorem"):
No matter what distribution X has, if n is "big" (like... $n > 30$) then

$$T_0 \text{ \& } \bar{X} \approx \text{Normal}$$

("approximately Normal")

$$T_0 \approx \text{Normal}(n\mu, \sqrt{n}\,\sigma)$$

$$\bar{X} \approx \text{Normal}(\mu, \sigma/\sqrt{n})$$

"Everything is normal"

Since $Var[\bar{X}] = \sigma^2/n$, as $n \to \infty$ $Var[\bar{X}] \to 0$

this is called the

Law of Large Numbers:
As $n \to \infty$, $\bar{X} \to \mu$

Example: X has $\mu = 10$ and $\sigma = 6$. Sample X 36 times to get $X_1, \ldots, X_{36}$.

- What are distributions of T_o & $\bar{X}$?

$$E[T_o] = n\mu = 36 \cdot 10 = 360$$

$$Var[T_o] = n\sigma^2 = 36 \cdot 36 = 36^2$$

$\hookrightarrow$ std. dev. $= \sqrt{36^2} = 36$

$\hookrightarrow T_o \approx \text{Normal}(360, 36)$

$$E[\bar{X}] = \mu = 10$$

$$Var[\bar{X}] = \sigma^2/n = 36/36 = 1$$

$\hookrightarrow$ std. dev $= \sqrt{1} = 1$

$\hookrightarrow \bar{X} \approx \text{Normal}(10, 1)$

- What is probability that $T_o > 400$?

$$P(T_o > 400) = 1 - P(T_o \leq 400)$$

$$= 1 - \text{pnorm}(400, 360, 36)$$

$$\approx .133$$

- What is probability that $\bar{X} < 8$?

$$P(\bar{X} < 8) = \text{pnorm}(8, 10, 1)$$

$$\approx .023$$

- How many samples are necessary for $P(\bar{X} < 8) < 1\%$

Note: $\bar{X} \approx \text{Normal}(10, 6/\sqrt{n})$

$$\frac{\bar{X} - 10}{6/\sqrt{n}} \approx \text{Normal}(0, 1)$$

$$P\left(\frac{\bar{X}-10}{6/\sqrt{n}} < z\right) = 1\% \quad \text{if} \quad z = \text{qnorm}(.01, 0, 1)$$

$$\approx -2.33$$

$$\|$$

$$P\left(\bar{X} < \frac{6}{\sqrt{n}} z + 10\right)$$

Need $\frac{6}{\sqrt{n}} z + 10 = 8$

$$n = \left(\frac{6}{-2} \cdot z\right)^2$$

$$= 9z^2 \approx 48.7$$

$\boxed{n = 49}$ $\leftarrow$ n must be an integer.

Note: $z = \text{qnorm}(.01, 0, 1)$ is # std. dev. you must go away from mean to have $P = 1\%$