Sample Mean

Theorem

Let X_i be random variables such that $X_i \overset{\text{iid}}{\sim} f(x)$ with population mean $E(X_i) = \mu$ and variance $V(X_i) = \sigma^2$. The mean and variance of $\bar{X}$ are given by:

$$\begin{aligned} E(\bar{X}) &= \mu \\ V(\bar{X}) &= \frac{\sigma^2}{n} \end{aligned}$$

regardless of the distribution of $\bar{X}$.

Proof. Since $E(X_i) = \mu$ and variance $V(X_i) = \sigma^2$:

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) \\ &= \frac{1}{N}\sum_{i=1}^{N} E(X_i) \\ &= \frac{1}{N}(N\mu) \\ &= \mu \end{aligned}$$

$$\begin{aligned} V(\bar{X}) &= V\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) \\ &= \frac{1}{N^2}\sum_{i=1}^{N} V(X_i) \\ &= \frac{1}{N^2}(N\sigma^2) \\ &= \frac{\sigma^2}{N} \end{aligned}$$

■

Thus, the sample standard deviation is inversely proportional to sample size.

Theorem

Let X_i be random variables such that $X_i \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \sigma^2)$:

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{N}\right)$$

Example

The weight of eggs produced on a farm have a normal distribution of $\mathrm{N}(65, 2^2)$. For a sample of size 12, what is the probability that $\bar{X}$ is within 65 ± 1? What about an individual egg?

$$\bar{X} \sim \mathrm{N}\left(56, \frac{2^2}{12}\right) = \mathrm{N}\left(56, \frac{1}{3}\right)$$

$$
\begin{aligned}
P(64 \leq \bar{X} \leq 66) &= P\left(\frac{64-65}{\frac{1}{\sqrt{3}}} \leq Z \leq \frac{66-65}{\frac{1}{\sqrt{3}}}\right) \\
&= (-1.73 \leq Z \leq 1.73) \\
&= \Phi(1.73) - \Phi(-1.73) \\
&= 0.9582 - 0.0418 \\
&= 0.9164
\end{aligned}
$$

$$
\begin{aligned}
P(64 \leq X \leq 66) &= P\left(\frac{64-65}{2} \leq Z \leq \frac{66-65}{2}\right) \\
&= (-0.50 \leq Z \leq 0.50) \\
&= \Phi(0.50) - \Phi(0.50) \\
&= 0.6915 - 0.3085 \\
&= 0.3830
\end{aligned}
$$

Example

In the library elevator of a large university, there is a sign indicating a 16-person limit as well as a weight limit of 2500 lbs. When the elevator is full, we can think of the 16 people in the elevator as a random sample of people on campus. Suppose that the weight of the students, faculty, and staff is normally distributed with a mean weight of 150 lbs and a standard deviation of 27 lbs. What is the probability that the total weight of a random sample of 16 people in the elevator will exceed the weight limit?

$$E(16\bar{X}) = 16E(\bar{X}) = 16 \cdot 150 = 2400 \text{ lbs}$$

$$V(16\bar{X}) = 16^2 V(\bar{X}) = 16^2 \cdot \frac{27^2}{16} = 16 \cdot 27^2$$

$$\sigma = 4 \cdot 27 = 108 \text{ lbs}$$

$$
\begin{aligned}
P(16\bar{X} > 2500) &= 1 - P(16\bar{X} \leq 2500) \\
&= 1 - P\left(Z \leq \frac{2500-2400}{108}\right) \\
&= 1 - P(Z \leq 0.93) \\
&= 1 - \Phi(0.93) \\
&= 1 - 0.8238 \\
&= 0.1762
\end{aligned}
$$

Theorem: Central Limit Theorem (CLT)

Let $X_i \overset{\text{iid}}{\sim} f(x)$ for any distribution $f(x)$ (discrete or continuous) such that both μ and σ^2 are finite. If n is large (≥ 30) then:

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Example

Suppose that the salaries of all SJSU employees follow an exponential distribution with the average salary is 45 (in thousands of dollars, which means that $\lambda = \frac{1}{45}$). We draw a random sample of size 30 from the population, and let $\bar{X}$ be the sample mean. Find $P(\bar{X} > 55)$.

$$\bar{X} \approx \mathrm{N}\left(45, \frac{45^2}{30}\right)$$

$$\begin{aligned} P(\bar{X} > 55) &= 1 - P(\bar{X} \leq 55) \\ &= 1 - P\left(Z \leq \frac{55 - 45}{\frac{45}{\sqrt{30}}}\right) \\ &= 1 - P(Z \leq 1.22) \\ &= 1 - \Phi(1.22) \\ &= 1 - 0.8888 \\ &= 0.1112 \end{aligned}$$

Note that the exact value is 0.1157.

Note that the normal approximation to the binomial distribution is a direct consequence of the CLT:

Theorem

Let $X \sim \mathrm{B}(n, p)$. If n is large ($np, n(1-p) \geq 10$) then:

$$X \approx \mathrm{N}(np, np(1-p))$$

Proof. Let $X_i \sim \mathrm{Bernoulli}(p)$. Then:

$$X = \sum_{i=1}^{n} X_i \sim \mathrm{B}(n, p)$$

But according to the CT:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X} - p}{\frac{\sqrt{p(1-p)}}{\sqrt{n}}} = \frac{X - np}{\sqrt{np(1-p)}} \approx \mathrm{N}(0, 1)$$

∎