Central Limit Theorem.

Let $X_1, X_2, \cdots, X_n$ be iid r.v. with finite mean μ and finite variance σ^2. Let $\bar{X}$ be the sample-mean of $X_1, X_2, \cdots, X_n$. Then

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right) \sim N(0,1) \text{ as } n \to \infty.$$

Note: We can use C.L.T. when $n > 30$.

Note: Most of the distribution can be approximated into standard normal distribution using CLT.

Ex Let $X_1, X_2, \cdots, X_n$ be iid and taken from Bernoulli distribution with parameter p, i.e., $Ber(p)$. Define $S_n = \sum_{i=1}^{n} X_i$, then

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \sim N(0,1) \text{ as } n \to \infty.$$

Note: Bernoulli distribution is special case of binomial distribution when $n = 1$. The sum of 'n' bernoullian trial is again turns into binomial-distribution. So

$$S_n \sim bin(n, p).$$

$$\lim_{n\to\infty} \frac{S_n - np}{\sqrt{npq}} \longrightarrow N(0,1) \quad (\text{By using CLT})$$

Ex Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} Poission(\lambda)$. Then show that

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \sim N(0,1). \text{ where } n \to \infty$$

where $S_n = X_1 + X_2 + \cdots + X_n$.

Solⁿ: $S_n = X_1 + X_2 + \cdots + X_n$

$$E(S_n) = E(X_1 + X_2 + \cdots + X_n)$$
$$= E(X_1) + E(X_2) + \cdots + E(X_n)$$
$$= \lambda + \lambda + \cdots + \lambda$$
$$= n\lambda$$

$$V(S_n) = V(X_1 + X_2 + \cdots + X_n)$$ $X_n \sim iid$
$$= V(X_1) + V(X_2) + \cdots + V(X_n) + 0 + 0 + \cdots$$
$$= \lambda + \lambda + \cdots + \lambda$$
$$= n\lambda$$

(Covariance term will be zero since X_i are independent)

$$\left[V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\,Cov(X,Y) \right]$$

Thus, $$\lim_{n\to\infty} \frac{S_n - n\lambda}{\sqrt{n\lambda}} \to N(0,1) \quad \text{(By CLT)}$$

Ex: A fair die is tossed 720 times. Use CLT to find the probability of getting 100 to 140 sixes.

Solⁿ: $n = 720$; $p = \frac{1}{6}$; $q = \frac{5}{6}$

$$X \sim bin\left(720, \tfrac{1}{6}\right)$$

$$P(100 \le X \le 140) = P\left(\frac{100 - E(X)}{\sqrt{V(X)}} \le \frac{X - E(X)}{\sqrt{V(X)}} \le \frac{140 - E(X)}{\sqrt{V(X)}} \right)$$

$$\therefore E(X) = np = 720 \times \tfrac{1}{6} = 120$$
$$V(X) = npq = 720 \times \tfrac{1}{6} \times \tfrac{5}{6} = 100.$$

$$P(100 \le X \le 140) = P\left(\frac{100-120}{\sqrt{100}} \le Z \le \frac{140-120}{\sqrt{100}} \right)$$
$$= P(-2 \le Z \le 2) = \Phi(2) - \Phi(-2)$$
$$= 2\Phi(2) - 1$$
$$= 2 \times 0.9772 - 1 = 0.954$$

Ex The inside diameter of a randomly selected piston ring is a rv with $\mu = 12$ cm and $\sigma = 0.04$ cm.

a) If $\bar{x}$ is the sample mean diameter for a random-sample of $n = 16$ rings, where is the sampling distribution of $\bar{x}$ centered, and what is the σ of the $\bar{x}$ dist.?

b) Calculate $P(11.99 \leq \bar{x} \leq 12.01)$ when $n=16$.

c) How likely is it that the sample mean diameter exceeds 12.01 when $n=25$?

Soln a) We know that $E(\bar{X}) = \mu = 12$

and $V(\bar{x}) = \frac{\sigma^2}{n}$ or $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

$$= \frac{0.04}{\sqrt{16}} = \frac{0.04}{4}$$

$$= 0.01.$$

b) $n=16$, $\mu=12$, $\sigma=0.04$

$$P(11.99 < \bar{x} < 12.01) = P\left(\frac{11.99-12}{\frac{0.04}{\sqrt{16}}} < \bar{x} < \frac{12.01-12}{0.04/\sqrt{16}}\right)$$

$$= P(-1 < Z < 1)$$

$$= P(Z<1) - P(Z<-1)$$

$$= \Phi(1) - \Phi(-1)$$

$$= 0.8413 - 0.1587 = 0.6826.$$

c) $n = 25$.

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{12.01-12}{0.04/\sqrt{25}} \approx 1.25.$$

$$P(\bar{x} > 12.01) = P(Z > 1.25)$$

$$= 1 - P(Z < 1.25)$$

$$= 1 - 0.8944 = 0.1056.$$