

EELE 3340

Probability Theory

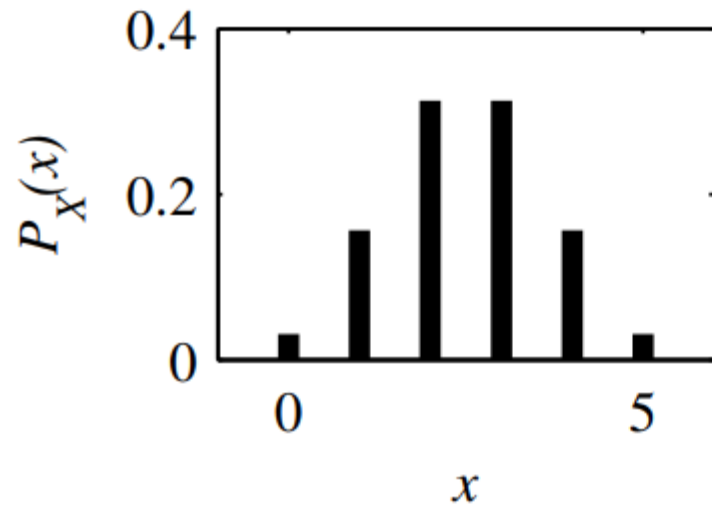


Chapter 9

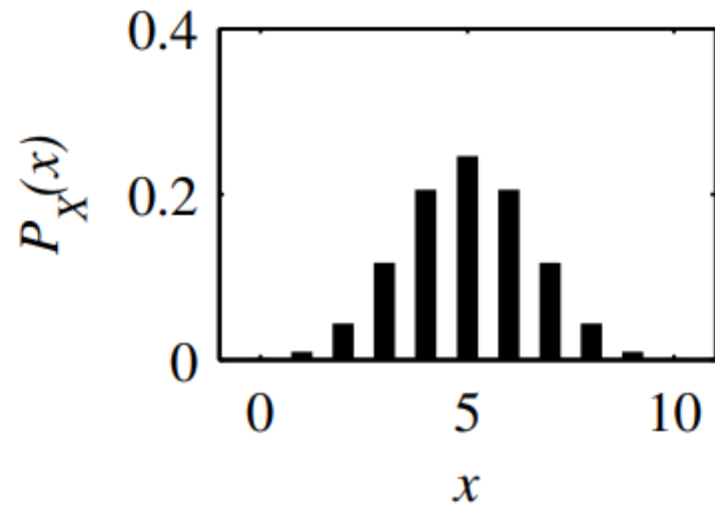
Central Limit Theorem

Dr. Mohammed El-Absi

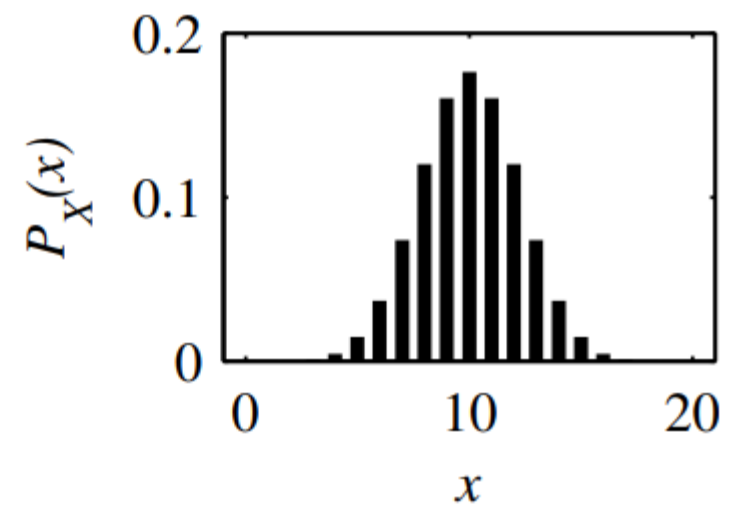
9.5 Central Limit Theorem



$n = 5$



$n = 10$



$n = 20$

The PMF of the X , the number of heads in n coin flips for $n = 5, 10, 20$. As n increases, the PMF more closely resembles a bell-shaped curve.

9.5 Central Limit Theorem

Definitions

The Central Limit Theorem (CLT)

Let $X_1, X_2, \dots, X_n$ be Independent and identically distributed (i.i.d). Random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \text{Var}(X_i) = \sigma^2 < \infty$. Then,

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad E[\bar{X}] = \mu \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

$$Z_n = \frac{\bar{X} - E[\bar{X}]}{\sqrt{\text{Var}(\bar{X})}} = \frac{\sum_{i=1}^n X_i/n - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma},$$

Converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n \rightarrow \infty} P(Z_n \leq x) = \Phi(x), \quad \text{for all } x \in \mathbb{R},$$

Where $\Phi(x)$ is the standard normal CDF. (Note: This true regardless of the distribution of X).

9.5 Central Limit Theorem

Definitions

In practice n finite, but we still can approximate $Y = X_1, X_2, \dots, X_n$ by a **Normal** random variable.

Two steps to solve problems using CLT:

$Y = X_1, X_2, \dots, X_n, \quad X_i \text{ i.i.d.}$

a) Find $E[Y] = \mu_Y = \sum_{i=1}^n E[X_i]$, and $\text{Var}(Y) = \sum_{i=1}^n \text{Var}(X_i)$.

b) Use $Y_n \sim N(\mu_{Y_n}, \text{Var}(Y_n))$, so we can use Φ function.

9.5 Central Limit Theorem

Assumptions:

- $X_1, X_2, \dots$ are iid Bernoulli(p).
- $Z_n = \frac{X_1 + X_2 + \dots + X_n - np}{\sqrt{np(1-p)}}$.

We choose $p = \frac{1}{3}$.

$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1



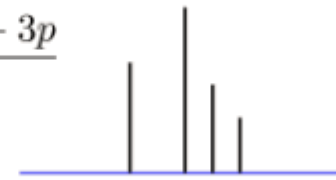
$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2



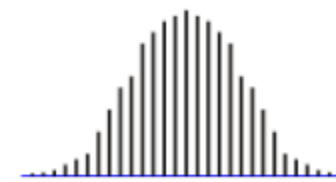
$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

PMF of Z_3



$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$

PMF of Z_{30}



9.5 Central Limit Theorem

Assumptions:

- $X_1, X_2, \dots$ are iid Uniform(0,1).
- $Z_n = \frac{X_1 + X_2 + \dots + X_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$.

$$Z_1 = \frac{X_1 - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$

PDF of Z_1



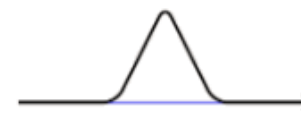
$$Z_2 = \frac{X_1 + X_2 - 1}{\sqrt{\frac{2}{12}}}$$

PDF of Z_2



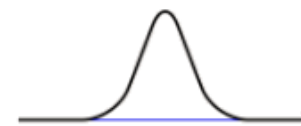
$$Z_3 = \frac{X_1 + X_2 + X_3 - \frac{3}{2}}{\sqrt{\frac{3}{12}}}$$

PDF of Z_3



$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - \frac{30}{2}}{\sqrt{\frac{30}{12}}}$$

PDF of Z_{30}



9.5 Central Limit Theorem

Example

A bank teller serves customers standing in the queue one by one. Suppose that the service time X_i for customer i has mean $EX_i = 2$ (minutes) and $\text{Var}(X_i) = 1$. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers. Find $P(90 < Y < 110)$.

$$Y = X_1 + X_2 + \dots + X_n, \quad n = 50, \quad EX_i = \mu = 2, \quad \text{Var}(X_i) = \sigma^2 = 1$$

$$\begin{aligned} P(90 < Y \leq 110) &= P\left(\frac{90 - n\mu}{\sqrt{n}\sigma} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \frac{110 - n\mu}{\sqrt{n}\sigma}\right) \\ &= P\left(\frac{90 - 100}{\sqrt{50}} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \frac{110 - 100}{\sqrt{50}}\right) = P\left(-\sqrt{2} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \sqrt{2}\right) \end{aligned}$$

By the CLT, $\frac{Y - n\mu}{\sqrt{n}\sigma}$ is approximately standard normal, so we can write

$$\begin{aligned} P(90 < Y \leq 110) &\approx \Phi(\sqrt{2}) - \Phi(-\sqrt{2}) \\ &= 0.8427 \end{aligned}$$

9.5 Central Limit Theorem

Example

There are 100 men on a plane. Let X_i be the weight (in pounds) of the i th man on the plane. Suppose that the X_i 's are i.i.d., and $EX_i = \mu = 170$ and $\sigma_{X_i} = \sigma = 30$. Find the probability that the total weight of the men on the plane exceeds 18,000 pounds.

If W is the total weight, then $W = X_1 + X_2 + \cdots + X_n$, where $n = 100$. We have

$$EW = n\mu = (100)(170) = 17000$$

$$\text{Var}(W) = 100\text{Var}(X_i) = (100)(30)^2 = 90000 \qquad \sigma_W = 300$$

$$\begin{aligned} P(W > 18000) &= P\left(\frac{W - 17000}{300} > \frac{18000 - 17000}{300}\right) \\ &= P\left(\frac{W - 17000}{300} > \frac{10}{3}\right) = 1 - \Phi\left(\frac{10}{3}\right) \quad (\text{by CLT}) \\ &\approx 4.3 \times 10^{-4}. \end{aligned}$$

Selected Problems

Questions of 9.5