IIT ROORKEE

FREE ONLINE EDUCATION swayam

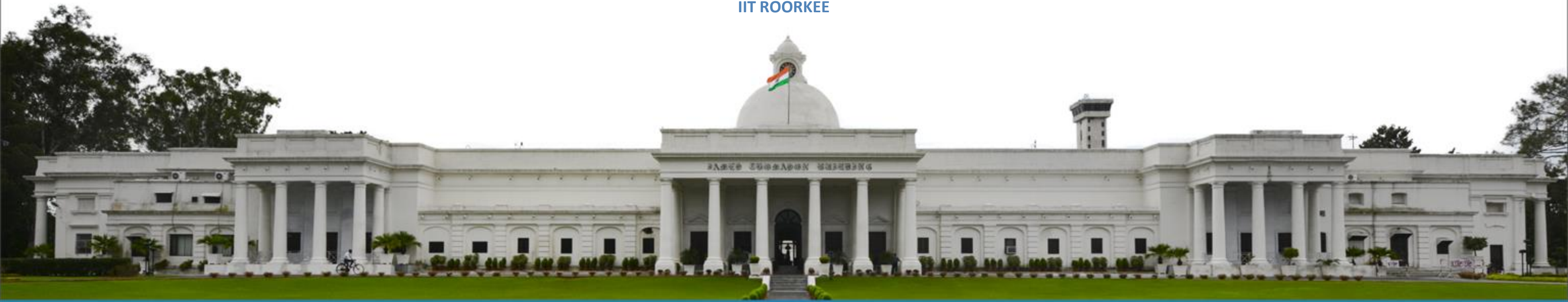
NPTEL ONLINE CERTIFICATION COURSE

Distribution of Sample Mean, proportion, and variance

Dr. A. Ramesh

DEPARTMENT OF MANAGEMENT

IIT ROORKEE



Acceptance Intervals

Goal: determine a range within which sample means are likely to occur, given a population mean and variance

- By the Central Limit Theorem, we know that the distribution of X is approximately normal if n is large enough, with mean μ and standard deviation
- Let $z_{\alpha/2}$ be the z-value that leaves area α/2 in the upper tail of the normal distribution (i.e., the interval - $z_{\alpha/2}$ to $z_{\alpha/2}$ encloses probability 1– α)
- Then

$$\boxed{\mu \pm z_{\alpha/2}\sigma_{\bar{X}}}$$

is the interval that includes X with probability 1 – α

Sampling Distributions of Sample Proportions

- Sampling Distributions
  - Sampling Distribution of Sample Mean
  - Sampling Distribution of Sample Proportion
  - Sampling Distribution of Sample Variance

Sampling Distributions of Sample Proportions

P = the proportion of the population having some characteristic

- Sample proportion ($\hat{p}$) provides an estimate of P:

$$\hat{p} = \frac{X}{n} = \frac{\text{number of items in the sample having the characteristic of interest}}{\text{sample size}}$$

- $0 \leq \hat{p} \leq 1$
- $\hat{p}$ has a binomial distribution, but can be approximated by a normal distribution when $nP(1 - P) > 5$

Sampling Distribution of $\hat{p}$

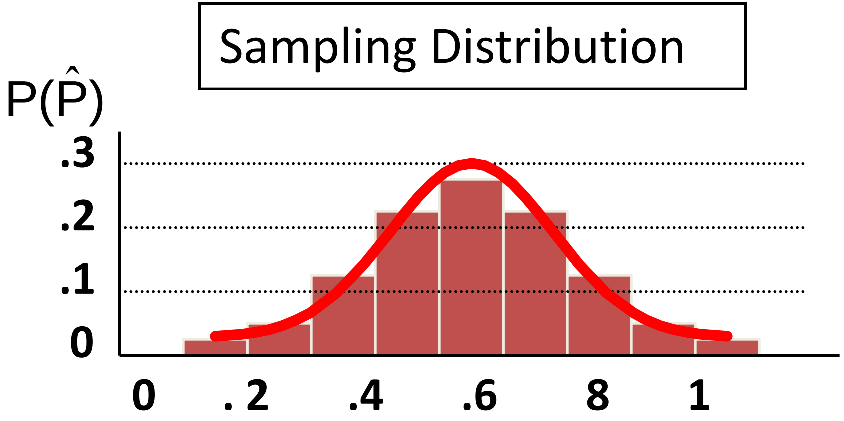
- Normal approximation:

Properties: $\boxed{E(\hat{p}) = P}$

(where P = population proportion)

And

$$\boxed{\sigma^2_{\hat{p}} = Var\left(\frac{X}{n}\right) = \frac{P(1-P)}{n}}$$



Z-Value for Proportions

Standardize $\hat{p}$ to a Z value with the formula:

$$Z = \frac{\hat{p} - P}{\sigma_{\hat{p}}} = \frac{\hat{p} - P}{\sqrt{\dfrac{P(1-P)}{n}}}$$

Example

- If the true proportion of voters who support Proposition A is P = .4, what is the probability that a sample of size 200 yields a sample proportion between .40 and .45?
- i.e.:

if P = .4 and n = 200, what is

$$P(.40 \le \hat{p} \le .45)\ ?$$

Example *(continued)*

- if $P = .4$ and $n = 200$, what is

  $P(.40 \le \hat{p} \le .45)$?

Find: $\sigma_{\hat{p}}$

$$\sigma_{\hat{p}} = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{.4(1-.4)}{200}} = .03464$$

Convert to standard normal:

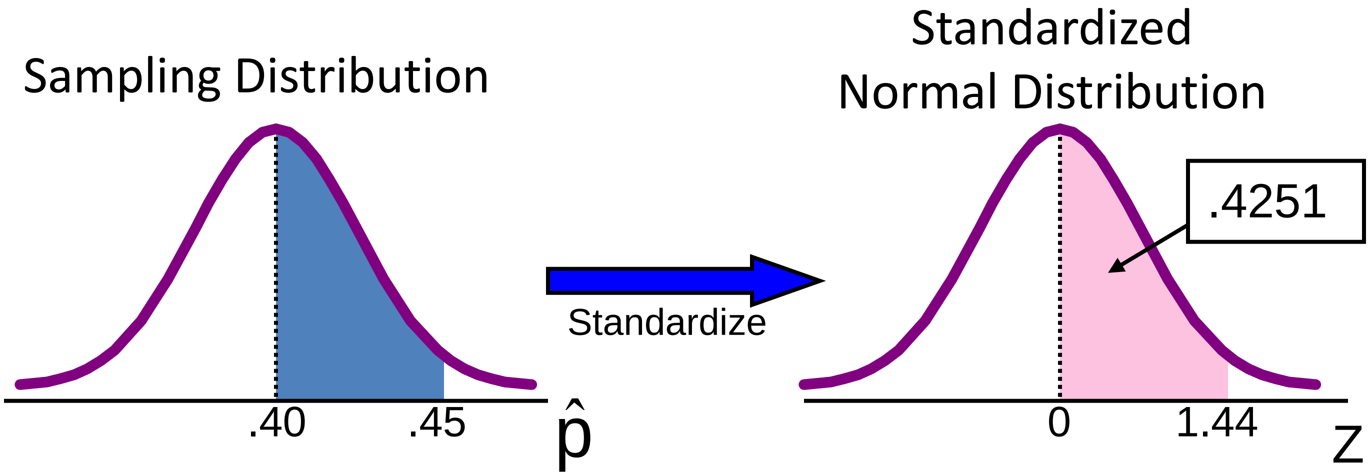
$$P(.40 \le \hat{p} \le .45) = P\left(\frac{.40-.40}{.03464} \le Z \le \frac{.45-.40}{.03464}\right)$$
$$= P(0 \le Z \le 1.44)$$

Example

(continued)

if P = .4 and n = 200, what is P(.40 ≤ $\hat{p}$ ≤ .45) ?

Use standard normal table: P(0 ≤ Z ≤ 1.44) = .4251



Sampling Distributions of Sample Variance

- Sampling Distributions
  - Sampling Distribution of Sample Mean
  - Sampling Distribution of Sample Proportion
  - Sampling Distribution of Sample Variance

Sample Variance

- Let $x_1, x_2, \ldots, x_n$ be a random sample from a population. The sample variance is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

- the square root of the sample variance is called the sample standard deviation

- the sample variance is different for different random samples from the same population

Sampling Distribution of Sample Variances

- The sampling distribution of s^2 has mean σ^2

$$E(s^2) = \sigma^2$$

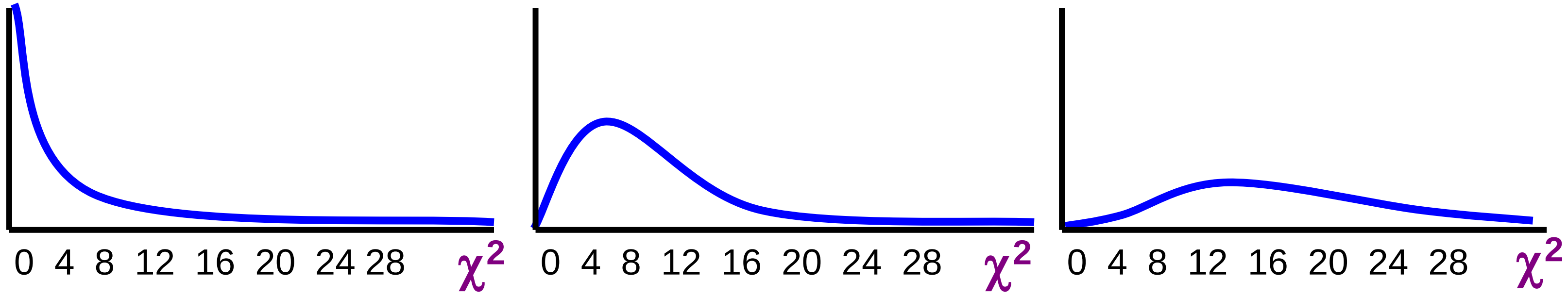
- If the population distribution is normal then

$$\frac{(n-1)s^2}{\sigma^2}$$

has a χ^2 distribution with n – 1 degrees of freedom

The Chi-square Distribution

- The chi-square distribution is a family of distributions, depending on degrees of freedom: d.f. = n – 1



Degrees of Freedom (df)

Idea: Number of observations that are free to vary after sample mean has been calculated

Example: Suppose the mean of 3 numbers is 8.0

Let $X_1 = 7$
Let $X_2 = 8$
What is X_3?

If the mean of these three values is 8.0, then X_3 must be 9
(i.e., X_3 is not free to vary)

Here, n = 3, so degrees of freedom $= n - 1 = 3 - 1 = 2$

(2 values can be any numbers, but the third is not free to vary for a given mean)

Chi-square Example

- A commercial freezer must hold a selected temperature with little variation. Specifications call for a standard deviation of no more than 4 degrees (a variance of 16 degrees^2).
- A sample of 14 freezers is to be tested
- What is the upper limit (K) for the sample variance such that the probability of exceeding this limit, given that the population standard deviation is 4, is less than 0.05?

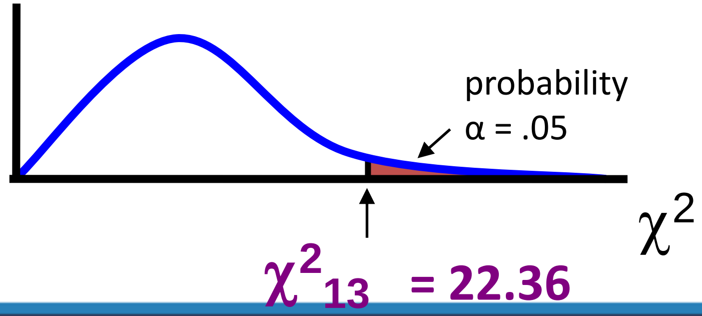
Finding the Chi-square Value

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Is chi-square distributed with (n – 1) = 13 degrees of freedom

- Use the the chi-square distribution with area 0.05 in the upper tail:

$$\chi^2_{13} = 22.36 \quad (\alpha = .05 \ \text{and} \ 14 - 1 = 13 \ \text{d.f.})$$



Chi-square Example

(continued)

χ^2_{13} = 22.36 (α = .05 and 14 – 1 = 13 d.f.)

So:

$$P(s^2 > K) = P\left(\frac{(n-1)s^2}{16} > \chi^2_{13}\right) = 0.05$$

or $$\frac{(n-1)K}{16} = 22.36$$

(where n = 14)

so $$K = \frac{(22.36)(16)}{(14-1)} = 27.52$$

If s^2 from the sample of size n = 14 is greater than 27.52, there is strong evidence to suggest the population variance exceeds 16.

Summary

- Introduced sampling distributions
- Described the sampling distribution of sample means
  - For normal populations
  - Using the Central Limit Theorem
- Described the sampling distribution of sample proportions
- Introduced the chi-square distribution
- Examined sampling distributions for sample variances
- Calculated probabilities using sampling distributions

Thank You