IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE

ANOVA

Dr. A. Ramesh

DEPARTMENT OF MANAGEMENT STUDIES

Agenda

- Sample Size Calculation
- One Way ANOVA – Introduction

Determining Sample Size when Estimating μ

- ***Z* formula** $$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- **Error of Estimation (tolerable error)** $$E = \overline{X} - \mu$$
- **Estimated Sample Size** $$n = \frac{Z_{\frac{\alpha}{2}}^2 \sigma^2}{E^2} = \left(\frac{Z_{\frac{\alpha}{2}} \sigma}{E} \right)^2$$

- **Estimated σ** $$\sigma \approx \frac{1}{4} range$$

Example: Sample Size when Estimating μ

$$E = 1, \quad \sigma = 4$$

$$90\% \text{ confidence } \Rightarrow Z = 1.645$$

$$n = \frac{Z_{\frac{\alpha}{2}}^{2}\sigma^{2}}{E^{2}}$$

$$= \frac{(1.645)^{2}(4)^{2}}{1^{2}}$$

$$= 43.30 \quad \text{or} \quad 44$$

Example

$$E = 2, \quad range = 25$$

$$95\% \text{ confidence } \Rightarrow Z = 1.96$$

$$estimated\ \sigma: \quad \frac{1}{4} range = \left(\frac{1}{4}\right)(25) = 6.25$$

$$n = \frac{Z^2 \sigma^2}{E^2}$$

$$= \frac{(1.96)^2 (6.25)^2}{2^2}$$

$$= 37.52 \quad \text{or} \quad 38$$

Determining Sample Size when Estimating *P*

- Z formula $$Z = \frac{\hat{p} - P}{\sqrt{\frac{P \cdot Q}{n}}}$$
- Error of Estimation (tolerable error) $$E = \hat{p} - P$$
- Estimated Sample Size $$n = \frac{Z^2 PQ}{E^2}$$

Example

$$E = 0.03$$

$$98\% \ Confidence \Rightarrow Z = 2.33$$

$$estimated\ P = 0.40$$

$$Q = 1 - P = 0.60$$

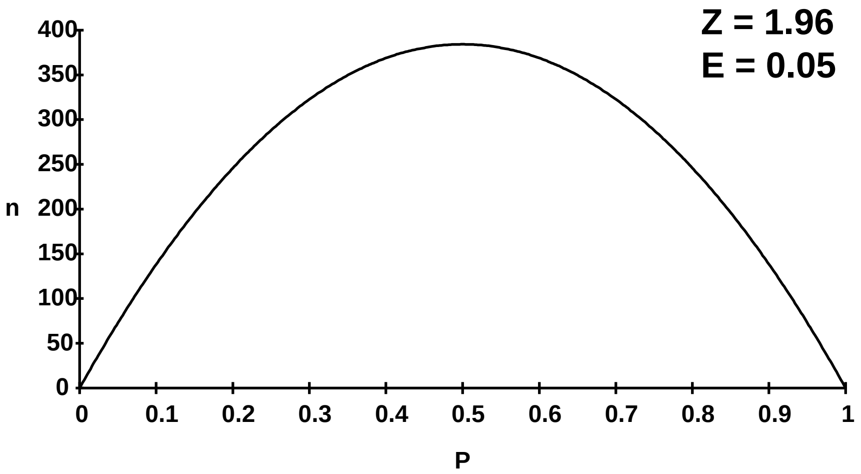
$$n = \frac{Z^2 PQ}{E^2}$$

$$= \frac{(2.33)^2 (0.40)(0.60)}{(.003)^2}$$

$$= 1{,}447.7 \ or \ 1{,}448$$

Determining Sample Size when Estimating *P* with No Prior Information

P	PQ
0.5	0.25
0.4	0.24
0.3	0.21
0.2	0.16
0.1	0.09



$$n = \frac{Z^2 \frac{1}{4}}{E^2}$$

Example

$E = 0.05$

$90\%\ Confidence \Rightarrow Z = 1.645$

$with\ no\ prior$ estimate $of\ P,\ use\ P = 0.50$

$Q = 1 - P = 0.50$

$$n = \frac{Z^2 PQ}{E^2}$$

$$= \frac{(1.645)^2 (0.50)(0.50)}{(.05)^2}$$

$$= 270.6\ or\ 271$$

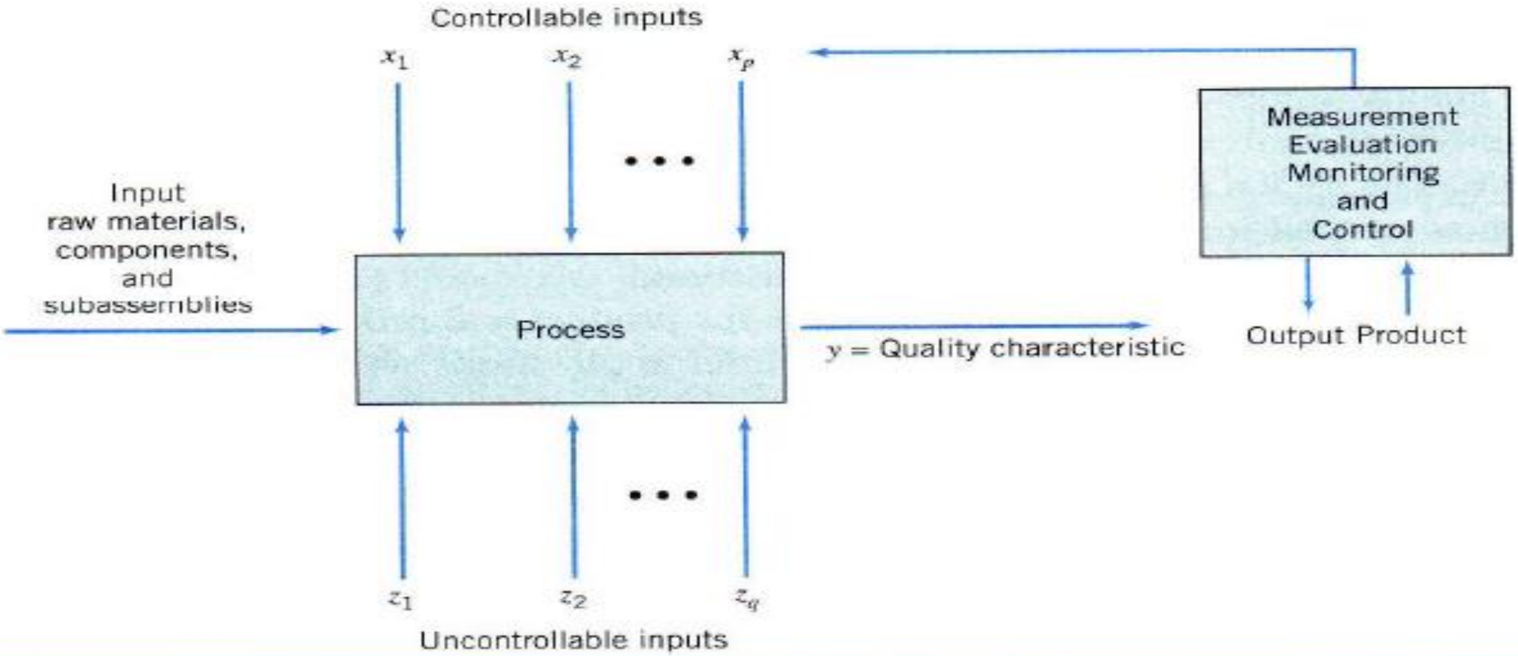
Why ANOVA?

- We could compare the means, one by one using t-tests for difference of means.
- Problem: each test contains type I error
- The total type I error is $1-(1-\alpha)^k$ where k is the number of means.
- For example, if there are 5 means and you use a=.05, you must make 10 two by two comparisons.
- Thus, the type I error is $1-(.95)^{10}$, which is .4012.
- That is, 40% of the time you will reject the null hypothesis of equal means in favor of the alternative!

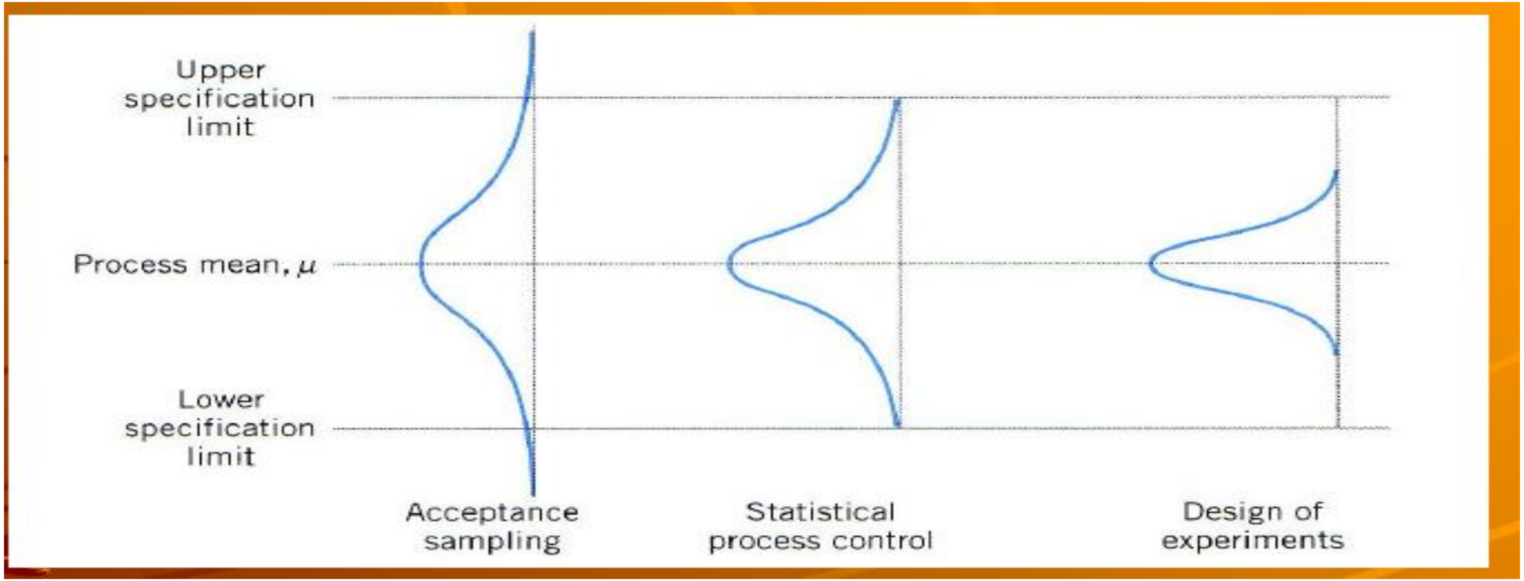
Hypothesis Testing With Categorical Data

- Chi Square tests can be viewed as a generalization of Z tests of proportions
- Analysis of Variance (ANOVA) can be viewed as a generalization of t-tests: a comparison of differences of means across more than 2 groups.
- Like Chi Square, if there are only two groups, the two analyses will produce identical results – thus a t-test or ANOVA can be used with 2 groups

Production Process inputs and outputs



Application of quality-engineering techniques and the systematic reduction of process variability



Effect of Teaching Methodology

Group 1 Black Board	Group 2 Case Presentation	Group 3 PPT
4	2	2
3	4	1
2	6	3

$$\bar{x}_1 = \frac{4+3+2}{3} = 3$$

$$\bar{x}_2 = \frac{2+4+6}{3} = 4$$

$$\bar{x}_3 = \frac{2+1+3}{3} = 2$$

$$\bar{\bar{x}} = \frac{4+3+2+2+4+6+2+1+3}{9} = 3$$

$$SST = (4-3)^2 + (3-3)^2 + (2-3)^2 + (2-3)^2 + (4-3)^2 + (6-3)^2 + (2-3)^2 + (1-3)^2 + (3-3)^2$$

$$= 1 + 0 + 1 + 1 + 1 + 9 + 1 + 4 + 0 = 18$$

$$SSB = 3(3-3)^2 + 3(4-3)^2 + 3(2-3)^2$$

$$= 0 + 3 + 3 = 6$$

$$SSE = (4-3)^2 + (3-3)^2 + (2-3)^2 + (2-4)^2 + (4-4)^2 + (6-4)^2 + (2-2)^2 + (1-2)^2 + (3-2)^2$$

$$= 1 + 0 + 1 + 4 + 0 + 4 + 0 + 1 + 1 = 12$$

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	6	2	3	1.5	0.296296	5.143253
Within Groups	12	6	2			
Total	18	8				

Thank You