# ANALYSIS OF VARIANCE (ANOVA)

**ANOVA** used to test the equality of more than two means. It is an extension of the hypothesis test for two means we have studied. The test value is computed by means of an ANOVA table.

| SOURCE | SUM OF SQUARES (SS) | Degrees of Freedom (df) | MEAN SQUARE (MS) | F VALUE or F RATIO |
|---|---|---|---|---|
| **Between means** | SSB | k - 1 | MSB | **F (test statistic)** |
| **Within the means** | SSW | n - k | MSW | not used |
| **Total** | SSTOT | n - 1 | not used | not used |

FIND:

1. **k** = number of groups with $n_i$ values in group i
2. **n** = total number of data values
3. $\bar{\mathbf{X}}$ **(grand mean)** $= \dfrac{\sum \mathbf{x}}{\mathbf{n}}$ **(mean of all the data values)**
4. Find mean $\bar{\mathbf{x}}_\mathbf{i}$ and variance $\mathbf{s_i^2}$ for each group separately
5. **SSB** = $\sum \mathbf{n_i}\left(\bar{\mathbf{x}}_\mathbf{i} - \bar{\mathbf{x}}\right)^2$
6. **SSW** = $\sum (\mathbf{n_i} - \mathbf{1})\mathbf{s_i^2}$
7. **SSTOT** = **SSB + SSW** (used as check value)
8. **Degrees of Freedom** = see table above
9. **MSB** = $\dfrac{\mathbf{SSB}}{\mathbf{k-1}}$ **MSW** = $\dfrac{\mathbf{SSW}}{\mathbf{n-k}}$
10. **F (test statistic)** = $\dfrac{\mathbf{MSB}}{\mathbf{MSW}}$

---

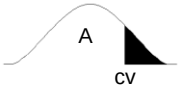
## HYPOTHESIS TEST FOR ANOVA

**$H_o$:** $\mu_1 = \mu_2 = \ldots = \mu_k$ (k is number of populations) (treated as right-tailed test)
**$H_1$:** at least one population mean is different from the others

**Use** given $\alpha$ level.

**Critical value**: from Table V (pages A-6 to A-9). **Use:** **given $\alpha$**
**df numerator $= \mathbf{k-1}$**
**df denominator $= \mathbf{n-k}$**

**Test value:** use F-ratio for ANOVA table.




**Decision Rule:** If test statistic > critical value, then reject $H_o$, otherwise accept $H_O$

**ANOVA Example #1:** In planning for future staffing, the ages of 19 hospital staff members were analyzed. Three groups (nurses, doctors, and x-ray techs) were chosen. At $\alpha = .05$, can it be concluded that the average ages of the three groups differ?

| Nurses | Doctors | X-ray Techs |
|---|---|---|
| 23 | 60 | 33 |
| 25 | 36 | 28 |
| 26 | 29 | 35 |
| 35 | 56 | 29 |
| 42 | 32 | 23 |
| 22 | 54 | 41 |
| | 58 | |

| | | | |
|---|---|---|---|
| **Sum of data** $\sum x_1$ | 173 | | |
| **Mean** $\overline{x_i}$ | 28.83 | | |
| **St. deviation** $s_i$ | 7.94 | | |
| **No. of values** $n_i$ | 6 | | |

| |
|---|
| **k =** |
| **n =** |
| $\bar{x}$ = |

| SOURCE | SUM OF SQUARES (SS) | Degrees of Freedom (df) | MEAN SQUARE (MS) | F VALUE |
|---|---|---|---|---|
| **Between means** | | | | |
| **Within a mean** | | | | not used |
| **Total** | | | not used | not used |

**$H_o$:**

**$H_1$:**

**$\alpha$ =**

**Critical value:** **Test value:**

**Decision:**

---

**Calculations:**

**ANOVA Example #2:** Workers are randomly assigned to four machines on an assembly line. The number of defective parts produced by each worker for one day is recorded. At α=.01, test the claim that the mean number of defective parts produced by the workers is the same

| | Machine 1 | Machine 2 | Machine 3 | Machine 4 |
|---|---|---|---|---|
| | 3 | 3 | 5 | 9 |
| | 2 | 3 | 7 | 9 |
| | 0 | 2 | 8 | 8 |
| | 4 | 0 | 6 | 8 |
| | 4 | 1 | 4 | 1 |
| | 3 | 4 | 5 | 2 |
| | 5 | 7 | 5 | 0 |
| **Sum of data $\sum x_i$** | 21 | 20 | | |
| **Mean $\bar{x}$** | 3 | 2.857 | | |
| **std. dev. $s_i$** | 1.633 | 2.268 | | |
| **Variance $s_i^2$** | 2.667 | 5.144 | | |
| **No. of values $n_i$** | 7 | 7 | | |

| |
|---|
| **K =** |
| **N =** |
| **$\bar{x}$ =** |

| SOURCE | SUM OF SQUARES (SS) | Degrees of Freedom (df) | MEAN SQUARE (MS) | F VALUE |
|---|---|---|---|---|
| **Between means** | | | | |
| **Within a mean** | | | | not used |
| **Total** | | | not used | not used |

**$H_o$:**

**$H_1$:**

**$\alpha$ =**

**Critical value:** **Test value:**

**Decision:**

---

**Calculations:**