# One-way Analysis of variance

- Categorical IV
- Quantitative DV
- *p* categories (groups)
- $H_0$: All population means equal
- Normal conditional distributions
- Equal variances

# Analysis means to split up

- With no IV, best predictor is the overall mean
- Variation to be explained is SSTO, sum of squared differences from the overall mean
- With an IV, best predictor is the group mean
- Variation still unexplained is SSW, sum of squared differences from the group means

$$SSTO = SSB + SSW$$

$$
\begin{aligned}
SSB &= \sum_{j=1}^{p} n_j(\overline{Y}_j - \overline{Y})^2 \\
SSW &= \sum_{j=1}^{p}\sum_{i=1}^{n_j}(Y_{i,j} - \overline{Y}_j)^2 \\
SSTO &= \sum_{j=1}^{p}\sum_{i=1}^{n_j}(Y_{i,j} - \overline{Y})^2.
\end{aligned}
$$

# ANOVA Summary Table

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | $p-1$ | $SSB$ | $MSB = SSB/(k-1)$ | $MSB/MSW$ | $p$-value |
| Error | $n-p$ | $SSW$ | $MSW = SSW/(n-k)$ | | |
| Corrected Total | $n-1$ | $SSTO$ | | | |

$$H_0 : \mu_1 = \ldots = \mu_p.$$

# $R^2$ is the proportion of variation explained by the independent variable

$$R^2 = \frac{SSB}{SSTO}$$

# Contrasts

$$\ell = a_1\mu_1 + a_2\mu_2 + \cdots + a_p\mu_p$$

$$L = a_1\overline{Y}_1 + a_2\overline{Y}_2 + \cdots + a_p\overline{Y}_p$$

$$\text{where } a_1 + a_2 + \cdots + a_p = 0$$

# Overall F-test is a test of p-1 contrasts

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

| $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|
| 1 | -1 | 0 | 0 |
| 0 | 1 | -1 | 0 |
| 0 | 0 | 1 | -1 |

$$\ell = a_1\mu_1 + a_2\mu_2 + \cdots + a_p\mu_p$$