10.1 Single-factor ANOVA

J: # of observation

i: # of group

① Notation and Assumption.

- X_{ij} = RV that denotes jth measurement taken from ith population.
- x_{ij} = the observed value of X_{ij}.
- Assumption:

  The I population or treatment distribution are all normal with the same variance σ^2

  Each X_{ij} is normally distributed with

  $E(X_{ij}) = \mu_i \qquad V(X_{ij}) = \sigma^2$

② Test Statistic.

- Mean square for treatment

$$MSTr = \frac{J}{I-1}\sum_i (\bar{X}_i - \bar{X})^2$$

- Mean square for error.

$$MSE = \frac{S_1^2 + S_2^2 + \cdots + S_I^2}{I}$$

- $F = \dfrac{MSTr}{MSE}$

③ Test

- Proposition.

  When H_0 is true

  $$E(MSTr) = E(MSE) = \sigma^2$$

  When H_0 is false

  $$E(MSTr) > E(MSE) = \sigma^2$$

④ F Distribution & F Test $\boxed{\text{Table A-9}}$

Let $F = \frac{MSTr}{MSE}$ be the test statistic in a single-factor ANOVA problem, involving I population with J random sample.

F has F distribution with $\nu_1 = I-1$ and $\nu_2 = I(J-1)$

P-value $= P(F \geq f$ when H_0 is true$)$

$= $ Area under $F_{I-1, I(J-1)}$

⑤ Sum of Square.

— Total sum of squares (SST)

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (x_{ij} - \bar{x}_{..})^2$$
$$= \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij}^2 - \frac{1}{IJ} x_{..}^2$$

— treatment sum of squares (SSTr)

$$SSTr = \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{x}_{i.} - \bar{x}_{..})^2 = \frac{1}{J} \sum_{i=1}^{I} x_{i.}^2 - \frac{1}{IJ} x_{..}^2$$

— Error sum of squares (SSE)

$$SSE = \sum_{i=1}^{I} \sum_{j=1}^{J} (x_{ij} - \bar{x}_{i.})^2$$

— $\boxed{\begin{aligned} x_{i.} &= \sum_{j=1}^{J} x_{ij} \\ x_{..} &= \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij} \end{aligned}}$

⑥ Fundemental Identity

$$SST = SSTr + SSE$$

$$MSTr = \frac{SSTr}{I-1}$$

$$MSE = \frac{SSE}{I(J-1)}$$

$$F = \frac{MSTr}{MSE}$$

① ANOVA Table

SoV	df	Sum of Squares	Mean Square	f
Treatment	$I-1$	$SSTr$	$MSTr = SSTr/(I-1)$	$\frac{MSTr}{MSE}$
Error	$I(J-1)$	SSE	$MSE = SSE/(I(J-1))$	
Total	$IJ-1$	SST		

10.2. Multiple Comparison in ANOVA

② Tukey's Procedure (T Method)

- T Method for identifying significantly Different μ

$$w = \underbrace{Q_{\alpha, I, I(J-1)}}_{\text{Table A10.}} \cdot \sqrt{MSE/J}$$

10.3. More on Single-Factor ANOVA.

① Unequal Sample size.

let $J_1, J_2, J \cdots n$ denote I sample size.

$n = \sum_i J_i$ be the total number of observations.

- $SST = \sum_{i=1}^{I} \sum_{j=1}^{J} x_{ij}^2 - \frac{1}{n} x_{..}^2$ $\quad df = n-1$
- $SSTr = \sum_{i=1}^{I} \frac{1}{J_i} x_{i.}^2 - \frac{1}{n} x_{..}^2$ $\quad df = I-1$
- $SSE = \sum_{i=1}^{I} (J_i - 1) S_i^2 = SST - SSTr$ $\quad df = n-1$

- $F = \frac{MSTr}{MSE}$, $MSTr = \frac{SSTr}{I-1}$; $MSE = \frac{SSE}{n-I}$
- P-value is area under $F_{I-1, n-1}$.
- $W_{ij} = Q_{\alpha, I, n-1} \cdot \sqrt{\frac{MSE}{2}\left(\frac{1}{J_i} + \frac{1}{J_j}\right)}$

- Test statistic for ANOVA
  - When H_0 is false, sample mean $\bar{X}_i$ spread more
  - H_0 is false $\Rightarrow F > 1$
  - If χ_1^2 and χ_2^2 are independent

$$F = \frac{\chi_1^2 / \nu_1}{\chi_2^2 / \nu_2} \qquad \nu_1 = I-1, \quad \nu_2 = I(J-1)$$

$SST = SSTr + SSE$

$MSTr = \frac{SSTr}{I-1}$

$MSE = \frac{SSE}{I(J-1)}$

$F = \frac{MSTr}{MSE}$

⑦ ANOVA Table

SoV	df	Sum of Squares	Mean Square	f
Treatment	$I-1$	$SSTr$	$MSTr = SSTr/(I-1)$	$\frac{MSTr}{MSE}$
Error	$I(J-1)$	SSE	$MSE = SSE/(I(J-1))$	
Total	$IJ-1$	SST		

⑧ Assumptions of ANOVA

- X_{ij} within any sample are independent.
  * usually come with design.

- I population comes from normal distribution.
  ($X_{i,j} \sim N(\mu_i, \sigma^2)$)
  * checked by normal probability plot.
- Population standard deviation are equal.
  * checked by $S_{max}/S_{min} < 2$