

ANOVA ^{one-way} is ~~by~~ a hypothesis test where

$$H_0: \mu_0 = \mu_1 = \dots = \mu_m$$

$$H_1: \text{not } H_0$$

Z-stat in ANOVA

Z-stat equiv is F-stat

$$F\text{-stat} = \frac{\text{mean sum of sqs Btwn (SSB)}}{\text{mean sum of sqs Wtwn (SSW)}}$$

ANOVA assumptions

1. samples are independent
2. each sample is fr. a normal dist. popn
3. The popn std dev of all grps are equal \Leftrightarrow homoscedasticity

$$F = \frac{mSSB}{mSSW} \sim F\text{-dist}(SSB \text{ df}, SSW \text{ df})$$

F-stat has an F-distn

mSSB & mSSW ^{each} have Chi-distn

$$SSB = \sum_{\text{all } m} \sum_{\text{all samples in grp } m} (X_i - \bar{X}_m)^2; SSW = \sum_{\text{all } m} n_m (\bar{X}_m - \bar{X}_{\text{total}})^2$$

$n_m = \text{size of grp } m$

eg grp \rightarrow A B C
 Sample X_{1A} X_{1B} X_{1C}
 one X_{2A} \vdots \vdots
 two \vdots \vdots \vdots
 three \vdots \vdots \vdots

SSW is looks @ diff. w/in a grp

$$SSW = (X_{1A} - \bar{X}_A)^2 + \dots + (X_{n,A} - \bar{X}_A)^2 + \sum (X_{iB} - \bar{X}_B)^2$$

SSB is looks @ diff btwn grps $+ \sum (X_{iB} - \bar{X}_B)^2$

$$n_A (\bar{X}_A - \bar{X}_{\text{total}})^2 + \dots + n_C (\bar{X}_C - \bar{X}_{\text{total}})^2$$

$$mSSB = \frac{SSB}{df_B} = \frac{SSB}{N-1}$$

$$mSSW = \frac{SSW}{df_W} = \frac{SSW}{N-k}$$

df = degrees of freedom; df has to do w/ how many values you can vary & still get out 1 stat.

Khan Academy is awesome!