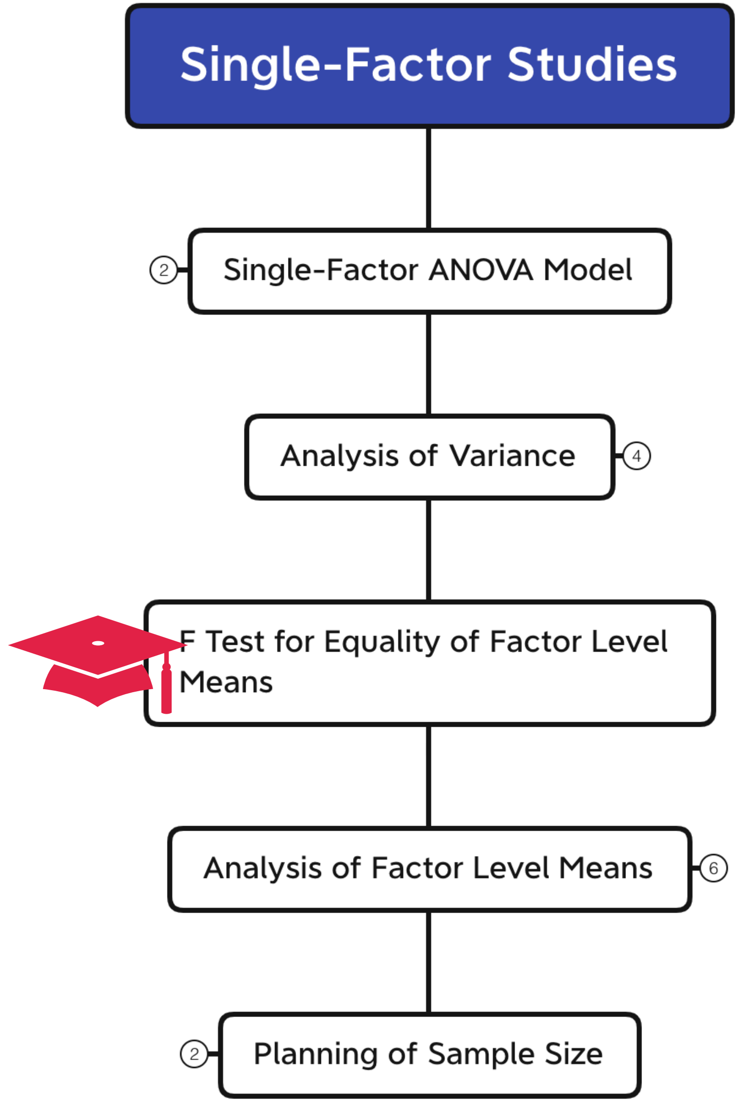
Lecture 2: Single-Factor Studies

STA 106: Analysis of Variance

Suggested reading: ALSM Chapter 16 & 17

Xiner Zhou
Department of Statistics, University of California, Davis


Single-Factor Studies
Single-Factor ANOVA Model
2
Analysis of Variance
4
F Test for Equality of Factor Level Means
Analysis of Factor Level Means
6
Planning of Sample Size
2

F Test for Equality of Factor Level Means

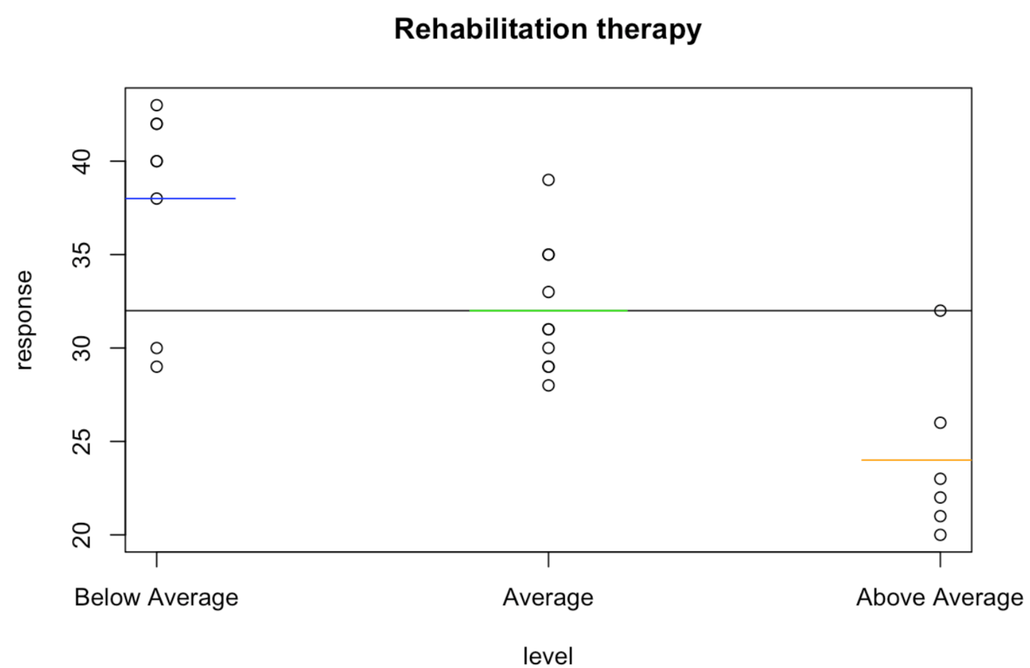
Substantive Research Questions of Interest:

Whether the factor levels or treatments differ in terms of response?
If the factor levels differ in terms of response, in what way do they differ or how do they differ?

These research questions lead to statistical questions usually performed in two steps, correspondingly.

Statistical Questions of Interest:

Whether the factor level means $\mu_i's$ are all equal or not?



F Test for Equality of Factor Level Means

To test whether or not the factor level means are the same:

$$H_0 : \mu_1 = u_2 = \ldots = \mu_r$$

H_a : not all μ_r are equal

Test statistic: $F^* = \dfrac{MSTR}{MSE}$

Large value of F^* support H_a

Small value, when $F^* \approx 1$ support H_0

—> We reject H_0 for large value of F^*, i.e. $F^* \geq c$

How do we decide what is "large" and what is "small"?

		Decision	
		H_0	H_a
Truth	H_0	✓	Type I error
	H_a	Type II error	✓

F Test for Equality of Factor Level Means

We want to control type I error at significance level α (usually .05):

If H_0 is true: $\mu_1 = \ldots = \mu_r$

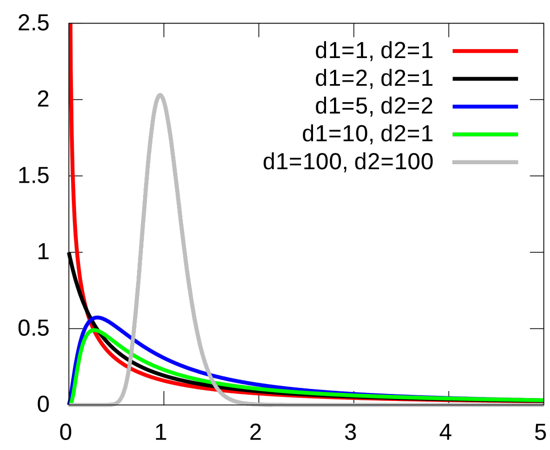
$$\frac{n_T - r}{\sigma^2} MSE = \frac{SSE}{\sigma^2} \sim \chi^2_{n_T - r}$$

$$\frac{r-1}{\sigma^2} MSTR = \frac{SSTR}{\sigma^2} \sim \chi^2_{r-1}\left(\frac{1}{\sigma^2}\left(\mu_i - \overline{\mu}.\right)^2\right) = \chi^2_{r-1}$$

MSE and MSTR are independent random variables

Re-write the test statistic:

$$\longrightarrow \quad F^* = \frac{MSTR}{MSE} = \frac{\dfrac{\frac{(r-1)MSTR}{\sigma^2}}{r-1}}{\dfrac{\frac{(n_T - r)MSE}{\sigma^2}}{n_T - r}} \sim \frac{\dfrac{\chi^2_{df=r-1}}{r-1}}{\dfrac{\chi^2_{df=n_T-r}}{n_T - r}} \sim F(r-1, n_T - r)$$

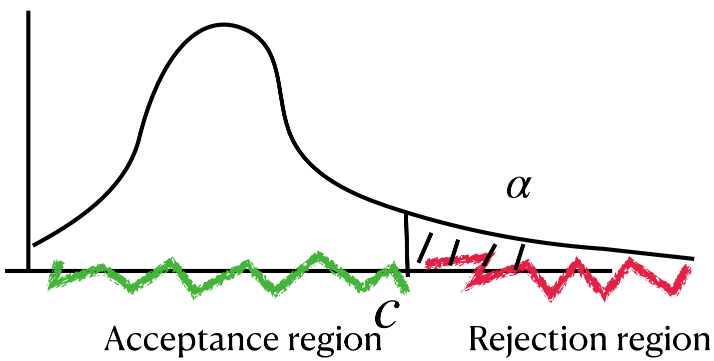


F Test for Equality of Factor Level Means

The probability of making Type I error = $P(H_0 \text{ is true, but we reject } H_0)$

$$= P(F(r-1, n_T - r) \geq c)$$

We want to control the risk of Type I error to be α



Critical value $c = F_{1-\alpha}(r-1, n_T - r)$

$(1-\alpha)100$ percentile of the F distribution

Decision rule:

If $F^* \leq F_{1-\alpha}(r-1, n_T - r)$, then conclude H_0

If $F^* > F_{1-\alpha}(r-1, n_T - r)$, then conclude H_a

Example

Is the mean number of days required for successful rehabilitation the same for the three fitness groups?
Control the significance level at .01

To test whether or not the factor level means are the same:

$$H_0 : \mu_1 = u_2 = \ldots = \mu_3$$

$$H_a : \text{not all } \mu_i \text{ are equal}$$

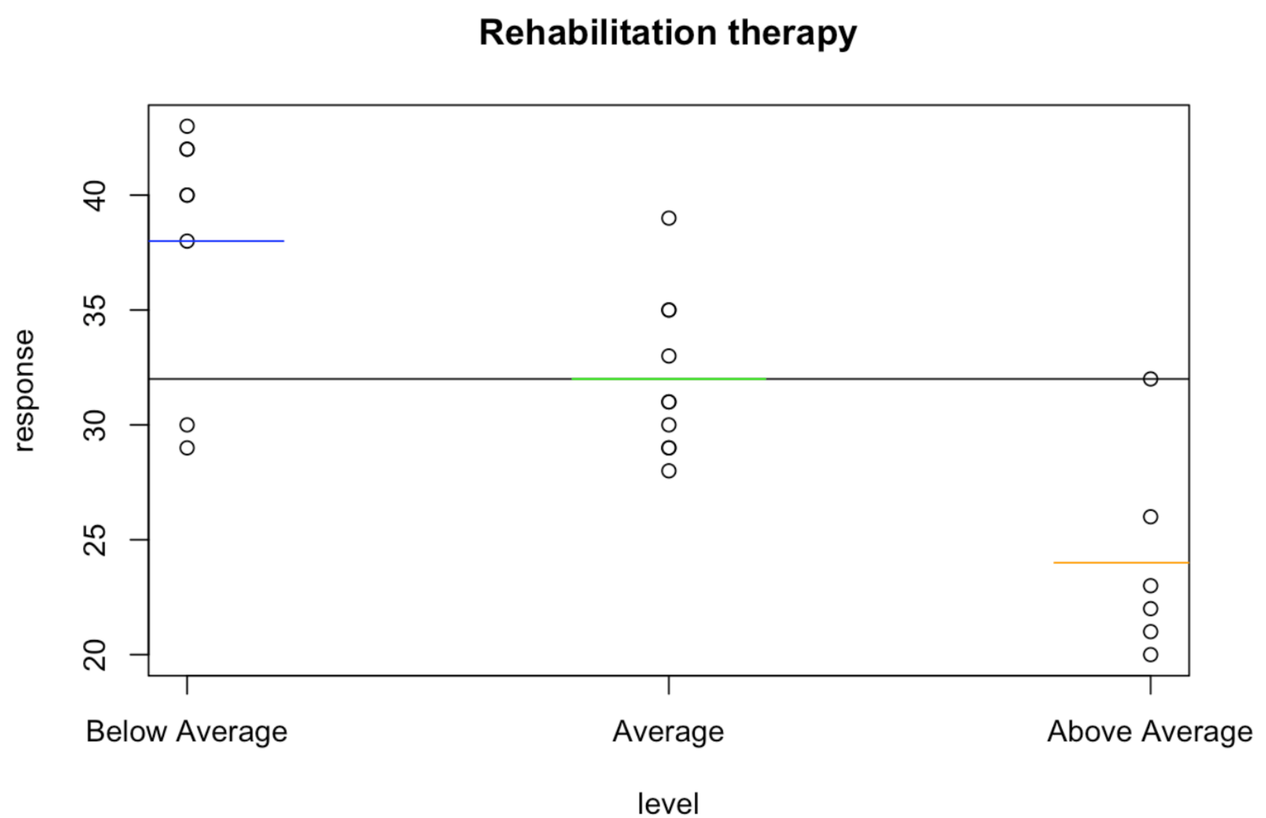
$$F^* = \frac{MSTR}{MSE} = \frac{336}{19.8} = 16.96$$

$$F\left(0.99; r-1, n_T - r\right) = 5.78$$

Since $F^* > F\left(0.99; r-1, n_T - r\right)$, we reject H_0 and conclude that, the mean number of days required for successful rehabilitation is different for the three fitness groups, which suggests that there exists some relation, further detailed analysis of the nature of the relation is required.

Example

What appears to be the nature of the relationship between physical fitness status and time required for physical therapy?



The sample means of days required for successful rehabilitation is the longest when prior fitness is below average, and it's the shortest when prior fitness is above average.

It suggests that the better prior fitness level is, the quicker the rehabilitation can be.