Handout 12-ANOVA

JIN, ICK HOON

2009. 07. 06.

- The statistical methodology for comparing several means is called analysis of variance, or ANOVA.
- In this case one variable is categorical: This variable forms the groups to be compared.
- The response variable is numeric.
- This methodology is the extension of comparing two means.

- “An investigator is interested in studying the average number of days rats live when fed diets that contain different amounts of fat. Three populations were studied, where rats in population 1 were fed a high-fat diet, rats in population 2 were fed a medium-fat diet, and rats in population 3 were fed a low-fat diet. The variable of interest is ‘Days lived.’” (from Graybill, Iyer and Burdick, *Applied Statistics*, 1998).

- “A state regulatory agency is studying the effects of secondhand smoke in the workplace. All companies in the state that employ more than 15 workers must file a report with the agency that describes the companys smoking policy. In particular, each company must report whether (1) smoking is allowed (no restrictions), (2) smoking is allowed only in restricted areas, or (3) smoking is banned. In order to determine the effect of secondhand smoke, the state agency needs to measure the nicotine level at the work site. It is not possible to measure the nicotine level for every company that reports to the agency, and so a simple random sample of 25 companies is selected from each category of smoking policy.’ (from Graybill, Iyer and Burdick, *Applied Statistics*, 1998).

- Each of the I population or group distributions is normal.
  - Check with a Normal Quantile Plot (or boxplot) of each group.
- These distributions have identical variances (standard deviations).
  - Check if largest standard deviation is greater 2 times smallest standard deviation.
- Each of the I samples is a random sample.
- Each of the I samples is selected independently of one another.

- **Step 1**: The null hypothesis for comparing several means is

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$$

  where I is the number of populations to be compared.

- **Step 2**: The alternative hypothesis is H_a: not all of the μ_i are equal (at least one of the means is different from the others)
- **Step 3**: State the significance level
- **Step 4**: Calculate the F-statistic

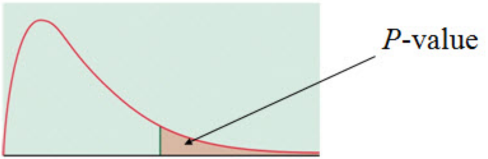
$$F = \frac{\text{Mean Squares Group}}{\text{Mean Squares Error}}$$

  This compares the variation between groups (group mean to group mean) to the variation within groups (individual values to group means).

- **Step 5**: Find the P-value
  - The P-value for an ANOVA F-test is always one-sided.
  - The P-value is

$$P(F_{n_1,n_2} > F_{Calculated})$$

where $n_1 = I - 1$ (number of groups minus 1) and $n_2 = N - 1$ (total sample size minus number of groups).



- **Step 6**: Reject or fail to reject H_0 based on the P-value.
  - If the P-value is less than or equal to α, reject H_0.
  - It the P-value is greater than α, fail to reject H_0.
- **Step 7**: State your conclusion.
  - If H_0 is rejected, "There is significant statistical evidence that **at least one** of the population means is different from another."
  - If H_0 is not rejected, "There is not significant statistical evidence that **at least one** of the population means is different from another."

Source	DF	Sum of Squares	Mean Square	F	p-Value
Group	$I-1$	$\sum n_i(\bar{x}_i - \bar{x})^2$	$\frac{SSG}{df_G} = MSG$	$\frac{MSG}{MSE} = F$	$P(F > F_{obs})$
Error	$N-I$	$\sum (n_i - 1)s_i^2$	$\frac{SSE}{df_E} = MSE$		
Total	$N-1$	$\sum (x_{ij} - \bar{x})^2$	$\frac{SST}{df_T} = MST$		

- Note: MSE is the pooled sample variance and SSG + SSE = SST.
- $R^2 = \frac{SSG}{SST}$ is the proportion of the total variation explained by the difference in means.

- “An experimenter is interested in the effect of sleep deprivation on manual dexterity. Thirty-two (N) subjects are selected and randomly divided into four (I) groups of size 8 (n_i). After differing amount of sleep deprivation, all subjects are given a series of tasks to perform, each of which requires a high amount of manual dexterity. A score from 0 (poor performance) to 10 (excellent performance) is obtained for each subject. Test at the α = 0.05 level the hypothesis that the degree of sleep deprivation has no effect on manual dexterity.” (from Milton, McTeer, and Corbet, *Introduction to Statistics*, 1997)

Information: Sample Size N=32

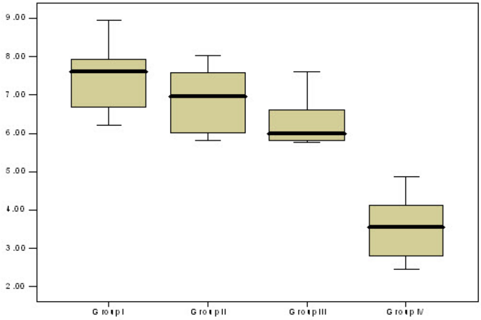
$$\text{Stddev1} = 0.89316$$
$$\text{Stddev2} = 0.86603$$
$$\text{Stddev3} = 0.64507$$
$$\text{Stddev4} = 0.85206$$

Group I 16 hours	Group II 20 hours	Group III 24 hours	Group IV 28 hours
8.95	7.7	5.99	3.78
8.04	5.81	6.79	3.35
7.72	6.61	6.43	2.45
6.21	6.07	5.85	4.27
6.48	8.04	5.78	4.87
7.81	5.96	7.6	3.14
7.5	7.3	5.78	3.98
6.9	7.46	6	2.47

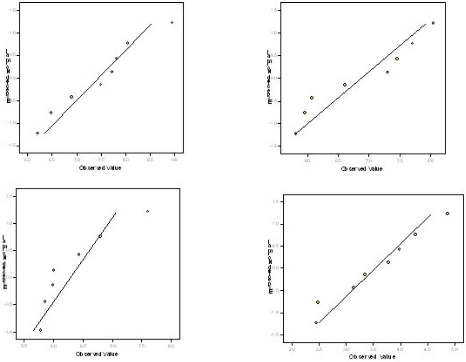
Variation within groups

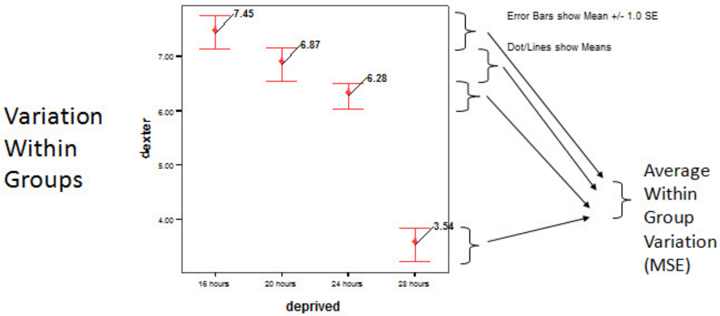
Variation between groups

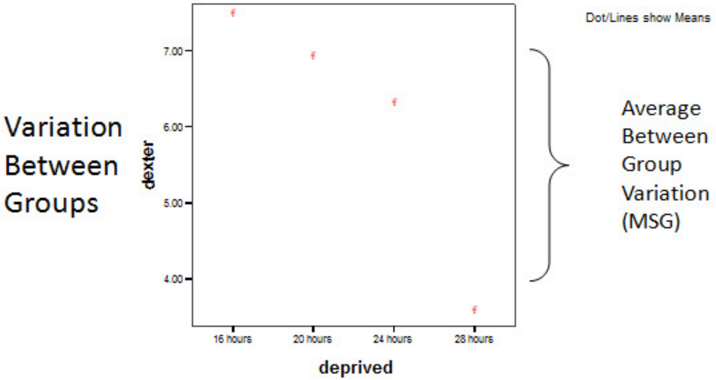
Side by Side Boxplot



Normal Quantile Plot




Variation
Within
Groups
7.45
6.87
6.28
3.54
7.00
6.00
5.00
4.00
dexter
16 hours
20 hours
24 hours
28 hours
deprived
Error Bars show Mean +/- 1.0 SE
Dot/Lines show Means
Average
Within
Group
Variation
(MSE)


Variation
Between
Groups
dexter
7.00
6.00
5.00
4.00
16 hours
20 hours
24 hours
28 hours
deprived
Dot/Lines show Means
Average
Between
Group
Variation
(MSG)

- **Step 1**: The null hypothesis is H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$.
- **Step 2**: The alternative hypothesis is H_a: not all of the μ_i are equal.
- **Step 3**: The significance level is α = 0.05.
- **Step 4**: Calculate the F-statistic

$$F = \frac{MSG}{MSE} = \frac{23.976}{0.671} = 35.73$$

MSG and MSE are found in the ANOVA table when the analysis is run on the computer:

ANOVA — MSG, MSE

DEXTER

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71.928	3	23.976	35.730	.000
Within Groups	18.789	28	.671		
Total	90.716	31			

- **Step 5**: Find the P-value

$$P(F_{n_1,n_2} > F_{Calculated}) = P(F_{n_1,n_2} > 35.73) < .0001$$

where $n_1 = I - 1 = 4 - 1 = 3$ and $n_2 = N - I = 32 - 4 = 28$

- **Step 6**: Reject or fail to reject H_0 based on the P-value.
  - Because the P-value is less than $\alpha = 0.05$, reject H_0.
- **Step 7**: State your conclusion.
  - "There is significant statistical evidence that **at least one** of the population means is different from another."

An additional test will tell us which means are different from the others.