
General Biostatistics

Part 7

1

Common Statistical Methods

Comparing More than
2 Groups

2

Outline

- Comparing means among 3 or more groups
- Comparing distributions of 3 or more groups

3

Continuous Outcomes

- Comparing means among 3 or more groups
 - Analysis of variance
 - F test
 - Multiple comparisons testing

4

Analysis of Variance (ANOVA)

- Test $H_0: \mu_1 = \mu_2 = \mu_3 \dots = \mu_k$ where k is the number of groups
- H_a : at least one μ_j is different
- Technique for comparing multiple means by partitioning the total *variation* in a data set into components defined by specific sources
- Assume independent observations and equal variance within groups

5

ANOVA

- ANOVA is a global test for differences in means among groups
- If the result of the global test (F test) is statistically significant, then perform individual pairwise comparisons of groups
- Pairwise comparisons of means is performed using multiple comparison procedures

6

Rationale for ANOVA

- Partition
total variation = variation between groups + variation within groups
- F statistic = MSB/MSW where
 - MSW is the estimate of variance based on within group differences
 - MSB is the estimate of variance based on between group differences

7

F-test

- F statistic = MSB/MSW
 - large F indicates more variability among groups than within groups \Rightarrow reject H_0 ; conclude that at least one group mean is different
 - F value near 1 indicates similar variability among groups as within groups \Rightarrow cannot reject H_0 ; conclude that the group means are equal

8

Statistical Decisions

- F statistic near 1 \Rightarrow conclude that the group means are equal \Rightarrow STOP
- Large F statistic \Rightarrow conclude that at least one of the group means is different \Rightarrow perform multiple comparisons test
 - Planned comparisons: Bonferroni
 - Post-hoc comparison: LSD, HSD, Dunnett's, Duncan's test

9

Multiple Comparison Tests

- Many pairwise comparison with p groups.
- The probability of Type I error = $\alpha = 0.05$ for each separate test.
- With multiple tests, α is increased above 0.05 \Rightarrow more comparisons are likely to be significant just by chance alone.
- Multiple comparisons procedures aim to preserve overall α level.

10

Example

Group	Mean Testosterone (ug/dl)
Nonsmoker	.68700001
Former Smoker	.682
Light Smoker	.571
Heavy Smoker	.46300001
Total	.60075

11

Example

Source	SS	df	MS	F	Prob > F
Between groups	.339007497	3	.113002499	3.80	0.0183
Within groups	1.07107003	36	.029751945		
Total	1.41007753	39	.036155834		

Bartlett's test for equal variances: $\chi^2(3) = 1.1840$
Prob> $\chi^2 = 0.757$

- **Reject H_0 that mean testosterone levels are the same for all smoking groups**

12

Example

Comparison of Testosterone by group
(Bonferroni)

Row Mean	Nonsmoke	Former S	LightSm
Col Mean			

Former S	- .005		
	1.000		
Light Sm	-.116	-.111	
	0.848	0.953	
Heavy Sm	-.224	-.219	-.108
	<u>0.038</u>	<u>0.044</u>	1.000

Statistically significant differences in mean testosterone levels between

- nonsmokers and heavy smokers
- former smokers and heavy smokers

13

Categorical Outcomes

- Comparing distributions of 3 or more groups
 - Comparing dichotomous outcomes among groups
 - Comparing categorical outcomes among groups
- χ^2 statistic

14

Categorical Outcomes

- Test Ho: no association
- Compare observed versus expected counts in r x c table
- Compute χ^2 statistic = $\chi^2 = \sum_{i=1}^k \left[\frac{(O_i - E_i)^2}{E_i} \right]$
- k = number of cell in r x c table
- df = (r - 1)(c - 1)
- Reject Ho if χ^2 statistic is large

15

Example

Seat Belt Use	Extent of Injury Received				Total
	None	Minor	Major	Death	
Yes	75	160	100	15	350
No	55	175	135	25	390
Total	130	335	235	40	740

16

Example

```
. tabi 75 160 100 15 \ 55 175 135 25, chi row
```

row	1	2	3	4	Total
1	75	160	100	15	350
	21.43	45.71	28.57	4.29	100.00
2	55	175	135	25	390
	14.10	44.87	34.62	6.41	100.00
Total	130	335	235	40	740
	17.57	45.27	31.76	5.41	100.00

Pearson chi2(3) = 9.3264 Pr = 0.025

17

Example

- χ^2 statistic = 9.326 with 2 df
- p-value = 0.025
- Conclude that seat belt users and non-users do not have the same extent of injury.
- The data indicate that seat belt users have fewer deaths and major injuries.

18

Summary

- Analysis of variance provides a test of whether means from multiple independent groups are the same
 - F test
 - Multiple comparisons tests
- Chi-squared statistic provides a comparison of dichotomous (categorical) outcomes among groups
