Stat 411/511

Linear Combinations of Means

Nov 13 2015

Charlotte Wickham

DA #2

Second question involves three two group comparisons.

A separate three sentence summary for each **is not required**.

Think about a table of the results plus a sentence or two summarizing them.

Only material prior to today is required for DA #2.

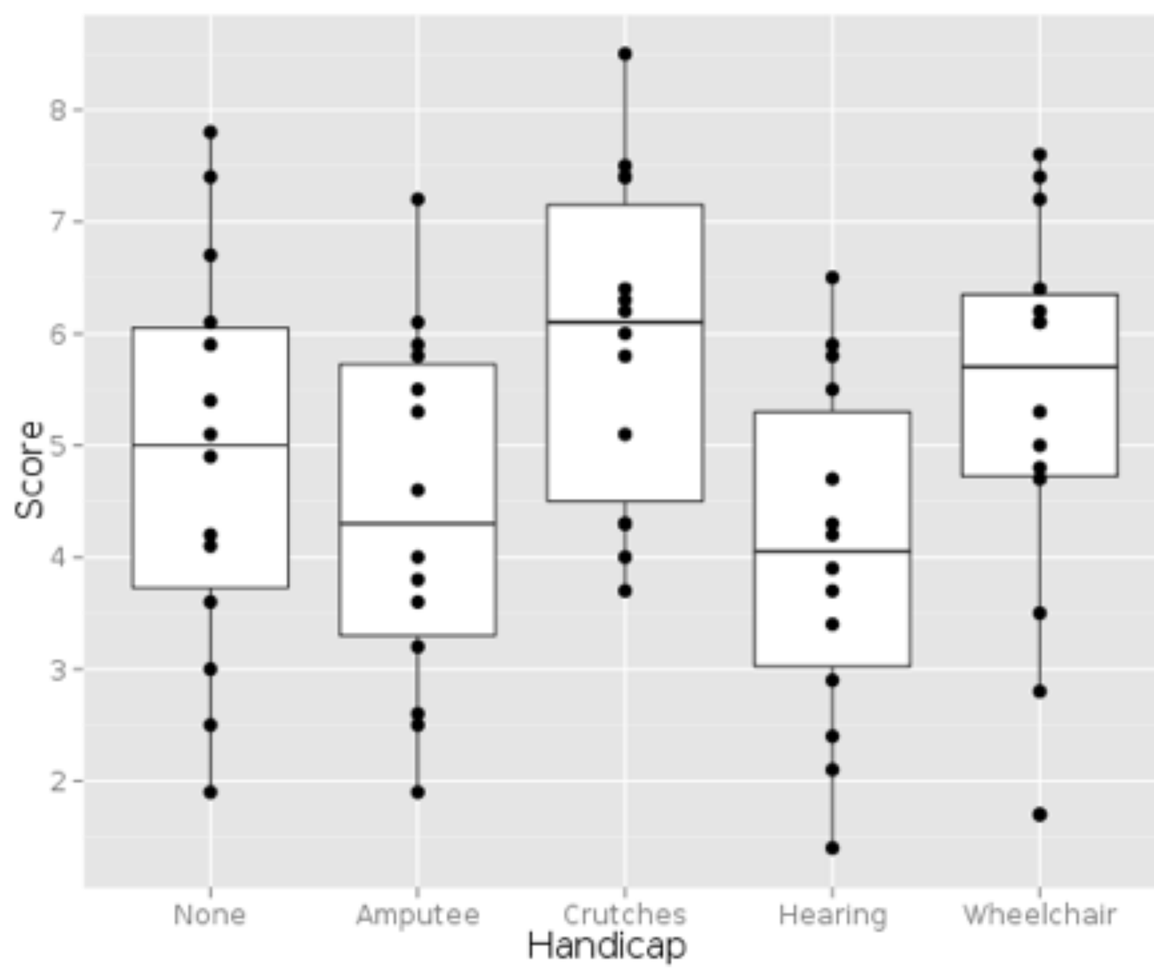
Preplanned comparisons: a few comparisons that directly answer the questions of interest. Start today

Unplanned comparisons: many comparisons are of interest, you often aren't sure which until you see the data.

Multiple comparisons

Disability case study

	None	Amputee	Crutches	Hearing	Wheelchair	Overall
Average	4.9	4.43	5.92	4.05	5.34	4.93
SD	1.79	1.59	1.48	1.53	1.75	1.72
Sample Size	14	14	14	14	14	70



Pooled SD = 1.63, d.f. = 65

A linear combination of means

If we have I groups, a **linear combination** of the I means is:

$$\gamma = C_1\mu_1 + C_2\mu_2 + C_3\mu_3 + \ldots + C_I\mu_I$$

where the C_i are constants

Preplanned comparisons can (often) be written as linear combinations of means.

Some examples

Two group comparison: Does group 1 have the same mean as group 2?

$\mu_1 - \mu_2$ $(C_1 = 1, C_2 = -1, C_3 = \ldots = C_I = 0)$

Disability Study: Compare crutches and wheelchair to hearing and amputee.

$1/2(\mu_{crutches} + \mu_{wheelchair}) - 1/2(\mu_{hearing} + \mu_{amputee})$ difference in average of means

$(C_{crutches} = 1/2, C_{wheelchair} = 1/2, C_{hearing} = -1/2 , C_{amputee} = -1/2)$

Spock study: Does Spock's judge select fewer women than the other judges on average?

$\mu_{Spock} - 1/6 (\mu_A + \mu_B + \ldots + \mu_F)$

$(C_{Spock} = 1, C_A = -1/6, C_B = -1/6 = \ldots = C_F = -1/6)$

Your turn

Find the linear combination for:

Disability study: Does the average of the mean scores of the **mobility** disabilities equal the mean of the hearing disability?

Estimate

To estimate the linear combination of means we substitute in the sample averages:

$$g = C_1\bar{Y}_1 + C_2\bar{Y}_2 + C_3\bar{Y}_3 + \ldots + C_I\,\bar{Y}_I$$

Disability Study: Compare crutches and wheelchair to hearing and amputee.

Estimate of $1/2(\mu_{crutches} + \mu_{wheelchair}) - 1/2(\mu_{hearing} + \mu_{amputee})$

$= 0.5\ \bar{Y}_{crutches} + 0.5\ \bar{Y}_{wheelchair} - 0.5\ \bar{Y}_{hearing} - 0.5\ \bar{Y}_{amputee}$

$= 0.5 * 5.92 + 0.5 * 5.34 - 0.5 * 4.05 - 0.5 * 4.43$

$= 1.39$

Standard Error

The **standard error** on the estimate on the linear combination of means is:

$$\mathrm{SE}_g = s_p \sqrt{\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \ldots + \frac{C_I^2}{n_I}}$$

d.f. = n - I

Disability Study: Compare crutches and wheelchair to hearing and amputee.

Standard error on estimate, g,

$= s_p \text{ sqrt}(0.5^2/14 + 0.5^2/14 + (-0.5)^2/14 + (-0.5)^2/14)$

= 1.63 sqrt(0.25/14 + 0.25/14 + 0.25/14 + 0.25 /14)

= 0.436

Your turn

Find the estimate and standard error of the linear combination:

$\mu_{Hearing} - 1/3\ (\mu_{Amputee} + \mu_{Crutches} + \mu_{Wheelchair})$

	None	Amputee	Crutches	Hearing	Wheelchair	Overall
Average	4.9	4.43	5.92	4.05	5.34	4.93
SD	1.79	1.59	1.48	1.53	1.75	1.72
Sample Size	14	14	14	14	14	70

Pooled SD = 1.63,
d.f. = 65

Testing a hypothesis

$$\text{t-ratio} = (g - \gamma) / SE_g$$

Has Student's t-distribution with n - I degrees of freedom

Disability Study: Compare crutches and wheelchair to hearing and amputee.

Null: The average mean score for crutches and wheelchair minus the average mean score for hearing and amputee is zero, $\gamma = 0$.

$\text{t-statistic} = (g - 0) / SE_g$

$= 1.39 / 0.436$

$= 3.189$

under the null, g/SE_g has a Student's t-distribution with n - I d.f.

2*(1 - pt(3.189, 65)) = 0.002

2-sided p-value

Constructing a 95% CI

$$g \pm qt(0.975, d.f.) \times SE_g$$

Disability Study: 95% CI for difference in average mean of crutches and wheelchair to hearing and amputee.

$$g \pm qt(0.975, d.f.) \times SE_g$$

$$= 1.39 \pm qt(0.975, 65)\ 0.436$$

$$= 1.39 \pm 2.00 \times 0.436$$

$$= 1.39 \pm 0.872$$

$$= (0.518, 2.262)$$

The average of mean scores for crutches and wheelchair is between 0.52 and 2.26 points higher than the average of mean score for hearing and amputee (95% confidence).

$SE_g = 0.50$
$g = -1.18$

Your turn

Test the null hypothesis that the linear combination,

$$\mu_{Hearing} - 1/3\,(\mu_{Crutches} + \mu_{Wheelchair} + \mu_{Amputee}) = 0$$

Find a 95% CI for the linear combination:

$$\mu_{Hearing} - 1/3\,(\mu_{Crutches} + \mu_{Wheelchair} + \mu_{Amputee})$$

Hint: qt(0.975, 65) = 2.00

Linear combinations in R

replicate “by hand” calculations (lab)

Or...

```
library(ggplot2)
library(Sleuth3)

install.packages("multcomp")
library(multcomp)
```

a new package

```
# first fit the individual group means model
full_model <- lm(Score ~ Handicap - 1, data  = case0601)

# set up the constants, checking the order using the full model
(groups <- names(coef(full_model)))

[1] "HandicapAmputee"    "HandicapCrutches"   "HandicapHearing"
[4] "HandicapNone"       "HandicapWheelchair"

C <- matrix(c(-1/3, -1/3, 1, 0, -1/3), nrow = 1,
  dimnames = list(1, groups))

> C
 HandicapAmputee HandicapCrutches HandicapHearing HandicapNone HandicapWheelchair
      -0.3333333       -0.3333333               1            0         -0.3333333
```

```
>  comp2b <- glht(full_model, linfct = C)
>  summary(comp2b)

     Simultaneous Tests for General Linear Hypotheses

Fit: lm(formula = Score ~ Handicap - 1, data = case0601)

Linear Hypotheses:
       Estimate Std. Error t value Pr(>|t|)
1 == 0  -1.1810     0.5039  -2.343   0.0222 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)

>  confint(comp2b)

     Simultaneous Confidence Intervals

Fit: lm(formula = Score ~ Handicap - 1, data = case0601)

Quantile = 1.9971
95% family-wise confidence level


Linear Hypotheses:
       Estimate lwr     upr
1 == 0 -1.1810  -2.1874 -0.1745
```

All pairwise differences in means for the disability study

95% confidence level

	Estimate	lwr (Lower)	upr (Upper)
Amputee - None == 0	-0.47143	-1.70405	0.76119
Crutches - None == 0	1.02143	-0.21119	2.25405
Hearing - None == 0	-0.85000	-2.08262	0.38262
Wheelchair - None == 0	0.44286	-0.78976	1.67548
Crutches - Amputee == 0	1.49286	0.26024	2.72548
Hearing - Amputee == 0	-0.37857	-1.61119	0.85405
Wheelchair - Amputee == 0	0.91429	-0.31833	2.14690
Hearing - Crutches == 0	-1.87143	-3.10405	-0.63881
Wheelchair - Crutches == 0	-0.57857	-1.81119	0.65405
Wheelchair - Hearing == 0	1.29286	0.06024	2.52548

Your turn

Imagine you have 100 tests for the difference between two group means.

In each test you reject the null hypothesis if the p-value is < 0.05.

If the null hypothesis is in fact true in all cases, how many tests would you expect to reject the null?

What is the problem?

Individual error rate: the probability of incorrectly rejecting the null hypothesis in a single test, α.

Familywise (or experimentwise) error rate: the probability of incorrectly rejecting **at least one** null hypothesis in a family of tests, α_E.

If $\alpha = 0.05$, $\alpha_E > 0.05$, and α_E gets bigger the more comparisons you make.