Homework 5

Tim Vigers

29 March 2019

```r
# Read in data
hw5 <- read.csv("/Users/timvigers/Documents/School/UC Denver/Biostatistics/Biostatistica
l Methods 2/Homeworks/Homework 5/hw5.txt",sep = "")
```

Model 1: Change-Score Model

```r
mod1 <- lm(delta_FEV1 ~ 1,data = hw5)
sm <- summary(mod1)
sm
```

```
##
## Call:
## lm(formula = delta_FEV1 ~ 1, data = hw5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33282 -0.08407 -0.02482  0.09893  0.28718
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.07582    0.01884   4.024 0.000198 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1332 on 49 degrees of freedom
```

```r
# MSE
mean((sm$residuals^2))
```

```
## [1] 0.01739387
```

a. Model equation

$$Y_{post_i} - Y_{pre_i} = \beta_0 + \epsilon = 0.076 + \epsilon \sim \mathrm{N}(0, 0.017)$$

b. Interpretation

The difference between pre- and post-bronchodilator FEV1 is significantly different from 0 (p = 0.000198).

c. Simple test

A paired t-test gives the same result:

```
t.test(hw5$pre_FEV1,hw5$post_FEV1,paired = T)
```

```
##
##  Paired t-test
##
## data:  hw5$pre_FEV1 and hw5$post_FEV1
## t = -4.0242, df = 49, p-value = 0.0001976
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.11368207 -0.03795793
## sample estimates:
## mean of the differences
##                -0.07582
```

Model 2: Baseline-as-Covariate Model

```
mod2 <- lm(post_FEV1 ~ pre_FEV1, data = hw5)
sm <- summary(mod2)
sm
```

```
##
## Call:
## lm(formula = post_FEV1 ~ pre_FEV1, data = hw5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32549 -0.08890 -0.02863  0.09946  0.28735
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.09612    0.04028   2.386    0.021 *
## pre_FEV1     0.98768    0.02156  45.808   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1341 on 48 degrees of freedom
## Multiple R-squared:  0.9776, Adjusted R-squared:  0.9772
## F-statistic:  2098 on 1 and 48 DF,  p-value: < 2.2e-16
```

```
# MSE
mean((sm$residuals^2))
```

```
## [1] 0.01727635
```

a. Model equation

$$Y_{post_i} = \alpha_0 + \alpha_1 Y_{pre_i} + \epsilon = 0.096 + 0.988 Y_{pre_i} + \epsilon \sim \mathrm{N}(0, 0.017)$$

b. Interpretation

Pre-bronchodilator FEV1 is significantly associated with post-bronchodilator FEV1 (p < 2e-16).

Model 3: Hybrid Model

```
mod3 <- lm(delta_FEV1 ~ pre_FEV1, data = hw5)
sm <- summary(mod3)
sm
```

```
##
## Call:
## lm(formula = delta_FEV1 ~ pre_FEV1, data = hw5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32549 -0.08890 -0.02863  0.09946  0.28735
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.09612    0.04028   2.386    0.021 *
## pre_FEV1    -0.01232    0.02156  -0.571    0.570
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1341 on 48 degrees of freedom
## Multiple R-squared:  0.006756,   Adjusted R-squared:  -0.01394
## F-statistic: 0.3265 on 1 and 48 DF,  p-value: 0.5704
```

```
# MSE
mean((sm$residuals^2))
```

```
## [1] 0.01727635
```

a. Model equation

$$Y_{post_i} - Y_{pre_i} = \gamma_0 + \gamma_1 Y_{pre_i} + \epsilon =$$
$$0.096 - 0.012 Y_{pre_i} + \epsilon \sim \mathrm{N}(0, 0.017)$$

b.

Change in FEV1 is not significantly associated with pre-bronchodilator FEV1 (p = 0.570)

c.

The difference between pre- and post-bronchodilator FEV1 is significantly different from 0 when controlling for baseline FEV1 (p = 0.021).

Show algebraically that Model 1 is nested within Model 2

$$\text{Model 1: } Y_{post_i} - Y_{pre_i} = \beta_0 + \epsilon$$
$$\text{Model 2: } Y_{post_i} = \alpha_0 + \alpha_1 Y_{pre_i} + \epsilon$$
$$\text{Therefore: Model 1} = \text{Model 2} - \alpha_1 Y_{pre_i}$$

This shows that model 2 is similar to model 1, but allows for variation in baseline FEV1. So the change-score model assumes within subject correlation, while the baseline-as-covariate model tests whether there is an association between the two time points.

Show algebraically that Models 2 and 3 are equivalent

$$\text{Model 2: } Y_{post_i} = \alpha_0 + \alpha_1 Y_{pre_i} + \epsilon$$
$$\text{Model 3: } Y_{post_i} - Y_{pre_i} = \gamma_0 + \gamma_1 Y_{pre_i} + \epsilon$$
$$\alpha_0 + \alpha_1 Y_{pre_i} + \epsilon = \gamma_0 + \gamma_1 Y_{pre_i} + Y_{pre_i} + \epsilon$$
$$\alpha_0 + \alpha_1 Y_{pre_i} + \epsilon = \gamma_0 + Y_{pre_i}(\gamma_1 + 1) + \epsilon$$

Therefore the two models are the same if $\alpha_1 = \gamma_1 + 1$ (because $\alpha_0 = \gamma_0$).

Model 4: Long Format

a.

The intercept in this model is the average pre-bronchodilator FEV1.

b.

$\hat{\beta}_1$ in this model is equivalent to $\hat{\beta}_0$ in model 1, which was the average difference between the timepoints. This makes sense because in model 4, $\hat{\beta}_1$ is the average change in FEV1 when going from the first time point to the second, so they are giving you the same information.

c.

The standard error for $\hat{\beta}_1$ in model 4 is higher, because the equation for this model essentially has an additional error term. Instead of adding ϵ_i where i = subjects 1,…,n, you’re adding ϵ_{ij} where indexes time one or time 2. This raises the model MSE and increases the SE calculation for the covariates.

In other words, model 1 accounts for within-subject correlation which allows for tighter covariate estimates. Model 4 doesn’t account for this correlation, which results in a higher standard error.

d. Simple test

A regular (un-paired t-test) would produce the same results in this case.

```
t.test(hw5$pre_FEV1,hw5$post_FEV1)
```

```
## 
##  Welch Two Sample t-test
## 
## data:  hw5$pre_FEV1 and hw5$post_FEV1
## t = -0.42675, df = 98, p-value = 0.6705
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.428395  0.276755
## sample estimates:
## mean of x mean of y 
##   1.64796   1.72378
```