Baseline & FU

SDC
http://bendixcarstensen.com/SDC/LEAD
August 2015
1.1

Compiled Wednesday 19th August, 2015, 11:24
from: /home/bendix/sdc/proj/LEAD/Vickers-ex

Bendix Carstensen Steno Diabetes Center, Gentofte, Denmark
& Department of Biostatistics, University of Copenhagen
bxc@steno.dk
http://BendixCarstensen.com

Contents

1 Acupuncture example from Vickers & Altman

Here we read the data from acupuncture example in the BMJ article by Vickers and Altman [**?**] — data has kindly been put at my disposal by DGA.

```
> library( Epi )
> library( foreign )
> acp <- read.dta( # "c:/bendix/artikler/rep-meas/data/sportsmen.dta" )[,-4]
+                  "./data/sportsmen.dta" )[,-4]
> names( acp ) <- c("bl","fu","gr")
> acp$gr <- factor( acp$gr, labels=c("Placebo","Acupuncture") )
> str( acp )
'data.frame':        54 obs. of  3 variables:
 $ bl: num  59 53 46 38 52 63 30 73 44 48 ...
 $ fu: num  81 53 83 51 81 86 42 74 45 54 ...
 $ gr: Factor w/ 2 levels "Placebo","Acupuncture": 1 1 1 1 1 1 1 1 1 1 ...
> head( acp )
  bl fu      gr
1 59 81 Placebo
2 53 53 Placebo
3 46 83 Placebo
4 38 51 Placebo
5 52 81 Placebo
6 63 86 Placebo
```

1.1 Conditional model

This is a model for the follow-up measurement on the i^{th} person, y_{1i}, as a function of the treatment (g = acupuncture, placebo), and the baseline value for the same person, y_{0i}:

$$y_{1i} = M + D_g + By_{0i}$$

The model implied that the changes from baseline to follow-up, $y_{1i} - y_{0i}$ in the two groups are:

$$y_{1i} - y_{0i} = M + D_g + By_{0i} - y_{0i} = M + D_g + (B-1)y_{0i}$$

SO the change is different between the groups; it is $D_1 - D - 0$ larger in group 1 than in group 0 while the effect of the baseline on the difference is the *same* in the two groups, namely $B - 1$.

Practically the model is fitted very simply:

```
> mc <- lm( fu ~ bl + gr, data=acp )
> summary( mc )
Call:
lm(formula = fu ~ bl + gr, data = acp)

Residuals:
    Min      1Q  Median      3Q     Max
-28.549  -9.258  -1.104  13.059  29.753

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    23.9973     9.1092   2.634  0.01125
bl              0.7102     0.1602   4.432 5.25e-05
grAcupuncture  12.7057     4.2857   2.965  0.00467
```

```
Residual standard error: 14.98 on 49 degrees of freedom
  (2 observations deleted due to missingness)
Multiple R-squared:   0.43,        Adjusted R-squared:  0.4067
F-statistic: 18.48 on 2 and 49 DF,  p-value: 1.046e-06
```

From the model with conditioning on baseline we see that the treatment effect is 12.7 points, and from the formulae above we have (since we assume no confounderes present) that the mean change in the placebo group is $M + (B-1)y_{i1} = 23.997 - 0.290 \times y_i$ and in the acupuncture group $M + (B-1)y_{i1} + D_g = 23.997 - 0.290 \times y_i + 12.706$. In order to report any of these two sensibly, we could stick in the mean of the baseline measurements:

```
> ( mb <- mean( acp$bl ) )
[1] 57.04259
```

so we get the estimated change from baseline to follow-up for a person with a baseline value equal to the mean baseline in the study under the two treatments:

```
> ( cf <- coef( mc ) )
  (Intercept)            bl grAcupuncture
   23.9973054     0.7102148    12.7057205
> (cf-c(0,1,1)) %*% cbind( c(1,mb,0),  c(1,mb,1) )
         [,1]     [,2]
[1,] 7.467206 19.17293
```

However there is nothing particularly sacred about the value 57.04 (mean baseline), we could also consider the expected changes for persons with baseline score of say 40 and 75:

```
> (cf-c(0,1,1)) %*% cbind( c(1,40,0),  c(1,40,1) )
        [,1]     [,2]
[1,] 12.4059 24.11162
> (cf-c(0,1,1)) %*% cbind( c(1,75,0),  c(1,75,1) )
         [,1]     [,2]
[1,] 2.263416 13.96914
```

as we see dramatically different changes, but with *differences* between changes equal to 12.71 in both cases.

We see that the changes are different in the different scenarios, but that the treatemnt effect — the difference between the changes — is the same troughout, namely 12.70.

Thus we can say that the treatment group increase the score 12.70 *more* than the placebo group, but exactly *how much* the increase is, depends on the baseline value.

1.2 Graphical illustration

We can illustrate this in figure 1, where the thin vertical line is drawn at the mean baseline (for *all* persons), and the *mean* (expected) change for a person with baseline equal to the overall baseline mean is the distance from the intersect with the identity to either the red or blue line depending on the treatment group.

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( acp, plot( bl, fu, pch=16, col=c("blue","red")[gr],
+                  xlim=c(20,100), ylim=c(20,100),
+                  xlab="Baseline score", ylab="Follow-up score" ) )
> abline( 0, 1 )
> text( rep(100,2), c(25,30), levels(acp$g), font=2, col=c("blue","red"), adj=1 )
```

```
> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( acp, plot( bl, fu, pch=16, col=c("blue","red")[gr],
+                  xlim=c(20,100), ylim=c(20,100),
+                  xlab="Baseline score", ylab="Follow-up score" ) )
> cf <- coef( mc )
> abline( 0, 1 )
> abline( v=mb )
> abline( v=c(40,75), lty="26" )
> abline( cf[1]      , cf[2], lwd=3, col="blue" )
> abline( cf[1]+cf[3], cf[2], lwd=3, col="red" )
> text( rep(100,2), c(25,30), levels(acp$g), font=2, col=c("blue","red"), adj=1 )

> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( acp, plot( bl, fu, pch=16, col=c("blue","red")[gr],
+                  xlim=c(20,100), ylim=c(20,100),
+                  xlab="Baseline score", ylab="Follow-up score" ) )
> ( fu <- with( acp, tapply( fu   , gr, mean ) ) )
    Placebo Acupuncture
    62.2963     79.6000
> diff( fu )
Acupuncture
    17.3037
> abline( 0, 1 )
> abline( v=mb )
> abline( v=c(40,75), lty="26" )
> abline( h=fu[1], lwd=2, lty=2, col="blue" )
> abline( h=fu[2], lwd=2, lty=2, col="red" )
> text( rep(100,2), c(25,30), levels(acp$g), font=2, col=c("blue","red"), adj=1 )

> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( acp, plot( bl, fu, pch=16, col=c("blue","red")[gr],
+                  xlim=c(20,100), ylim=c(20,100),
+                  xlab="Baseline score", ylab="Follow-up score" ) )
> ( df <- with( acp, tapply( fu-bl, gr, mean ) ) )
    Placebo Acupuncture
    8.37037    19.20000
> diff( df )
Acupuncture
   10.82963
> abline( 0, 1 )
> abline( v=mb )
> abline( v=c(40,75), lty="26" )
> abline( df[1], 1, lwd=2, lty=2, col="blue" )
> abline( df[2], 1, lwd=2, lty=2, col="red" )
> text( rep(100,2), c(25,30), levels(acp$g), font=2, col=c("blue","red"), adj=1 )

> par( mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
> with( acp, plot( bl, fu, pch=16, col=c("blue","red")[gr],
+                  xlim=c(20,100), ylim=c(20,100),
+                  xlab="Baseline score", ylab="Follow-up score" ) )
> df <- with( acp, tapply( fu-bl, gr, mean ) )
> fu <- with( acp, tapply( fu   , gr, mean ) )
> abline( 0, 1 )
> abline( v=mb )
> abline( v=c(40,75), lty="26" )
> abline( cf[1]      , cf[2], lwd=3, col="blue" )
> abline( cf[1]+cf[3], cf[2], lwd=3, col="red" )
> abline( h=fu[1], lwd=2, lty=2, col="blue" )
> abline( h=fu[2], lwd=2, lty=2, col="red" )
> abline( df[1], 1, lwd=2, lty=2, col="blue" )
> abline( df[2], 1, lwd=2, lty=2, col="red" )
> text( rep(100,2), c(25,30), levels(acp$g), font=2, col=c("blue","red"), adj=1 )
```

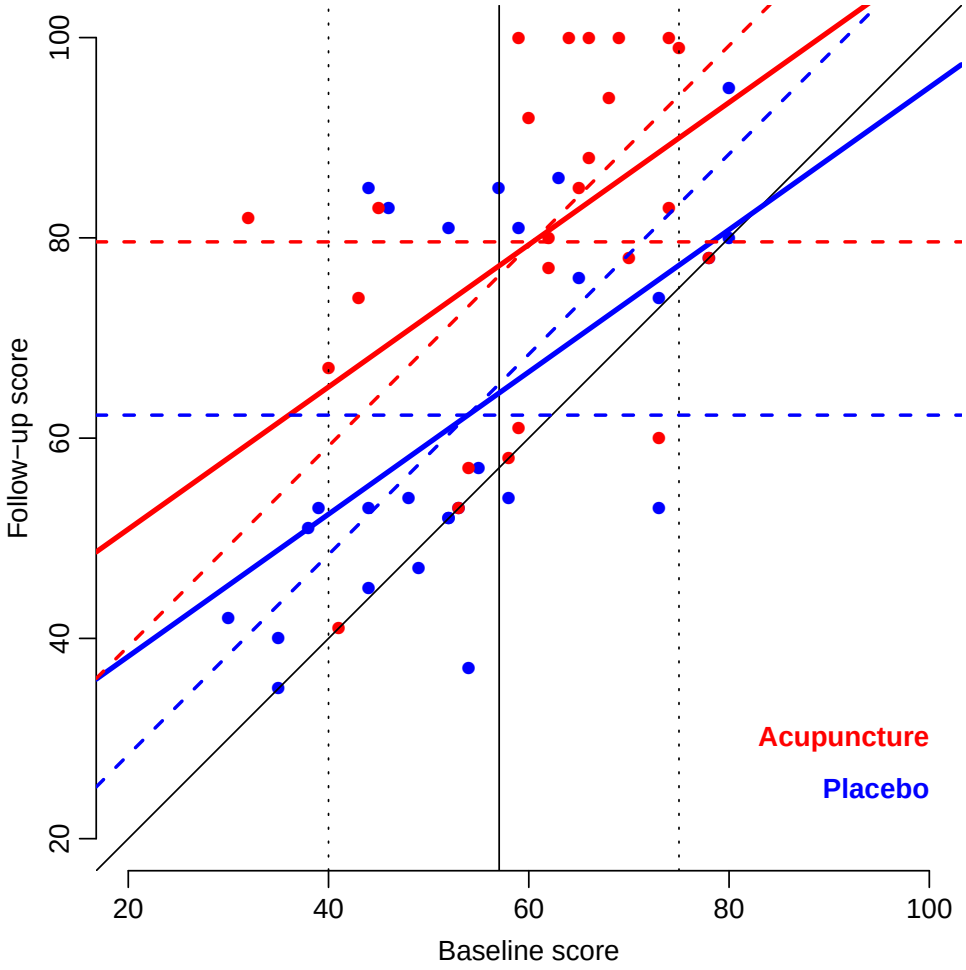

Figure 1: *Follow-up versus baseline score for acupuncture data. Regression lines are from the ANCOVA model, the horizontal dashed lines are the means of the follow-up data, and the* 45° *dashed lines correspond to the analysis of the change scores.*
The three lines of each color necessarily all pass through the point (mean(`bl`), mean(`fu`)).
The change for each person is the vertical *distance to the identity line.*

1.3 Comparing approaches

We can compare what the size of the estimates obtained from the three different approaches:

```
> # Follow-up
> fu <- with( acp, tapply( fu    , gr, mean ) )
> c( fu, diff( fu ) )
    Placebo Acupuncture Acupuncture
    62.2963     79.6000     17.3037

> mf <- lm( fu ~ gr, data=acp )
> # differences
> df <- with( acp, tapply( fu-bl, gr, mean ) )
> c( df, diff( df ) )
    Placebo Acupuncture Acupuncture
    8.37037    19.20000    10.82963
```

```
> md <- lm( fu-bl ~ gr, data=acp )
> # compare with regression
> cmp.cf <- rbind( ci.lin( mc, subset="Acu" ),
+                  ci.lin( mf, subset="Acu" ),
+                  ci.lin( md, subset="Acu" ) )
> rownames( cmp.cf ) <- c("Cond.","FU onl","Diff.")
> round( cmp.cf, 4 )
       Estimate StdErr      z      P   2.5%   97.5%
Cond.   12.7057 4.2857 2.9647 0.0030 4.3059 21.1056
FU onl  17.3037 4.8723 3.5515 0.0004 7.7542 26.8532
Diff.   10.8296 4.2516 2.5472 0.0109 2.4966 19.1627
```

1.4 A random effects model

In order to fit the random effects model we must have the data in the long format:

```
> lg <- reshape( acp, varying=1:2, v.names="score", direction="long" )
> head( lg )
         gr time score id
1.1 Placebo    1    59  1
2.1 Placebo    1    53  2
3.1 Placebo    1    46  3
4.1 Placebo    1    38  4
5.1 Placebo    1    52  5
6.1 Placebo    1    63  6
> str( lg )
'data.frame':        108 obs. of  4 variables:
 $ gr   : Factor w/ 2 levels "Placebo","Acupuncture": 1 1 1 1 1 1 1 1 1 1 ...
 $ time : int  1 1 1 1 1 1 1 1 1 1 ...
 $ score: num  59 53 46 38 52 63 30 73 44 48 ...
 $ id   : int  1 2 3 4 5 6 7 8 9 10 ...
 - attr(*, "reshapeLong")=List of 4
  ..$ varying:List of 1
  .. ..$ score: chr  "bl" "fu"
  .. ..- attr(*, "v.names")= chr "score"
  .. ..- attr(*, "times")= int  1 2
  ..$ v.names: chr "score"
  ..$ idvar  : chr "id"
  ..$ timevar: chr "time"
```

Fitting a model for the long-form data is easily done using `lmer` — we fit both the model with baseline imbalance and the model with assumed equal mean at beaseline:

```
> library( lme4 )
> mr <- lmer( score ~  gr + gr:factor(time) + (1|id), data=lg )
> round( summary( mr )$coef, 2 )
                           Estimate Std. Error t value
(Intercept)                   53.93       2.99   18.03
grAcupuncture                  6.47       4.31    1.50
grPlacebo:factor(time)2        8.37       2.95    2.84
grAcupuncture:factor(time)2   19.20       3.06    6.27
```

Thus the acupuncture group has a mean at baseline which is 6.47 larger then th eplacebo group; the change in the placebo group is 8.37, in the acupuncture group it is 19.20, the difference thus 10.83, not far from the difference we saw in the conditiona model.

If we were to compare to the parameter estimated in the conditional model it should be $(1-\rho)\delta_g + \gamma_{g2}$. Now this refers to a different parametrization:

```
> mR <- lmer( score ~  gr*factor(time) + (1|id), data=lg )
> round( summary( mR )$coef, 2 )
                           Estimate Std. Error t value
(Intercept)                   53.93       2.99   18.03
grAcupuncture                  6.47       4.31    1.50
factor(time)2                  8.37       2.95    2.84
grAcupuncture:factor(time)2   10.83       4.25    2.55
```

where we now have $\gamma_{g2} = 10.83$ and $\delta_g = 6.47$; the ρ can be derived from the the variance components:

```
> summary( mR )
Linear mixed model fit by REML ['lmerMod']
Formula: score ~ gr * factor(time) + (1 | id)
   Data: lg

REML criterion at convergence: 830.1

Scaled residuals:
     Min       1Q   Median       3Q      Max
-1.80685 -0.56741  0.01961  0.58225  1.69548

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 124.2    11.14
 Residual             117.3    10.83
Number of obs: 104, groups:  id, 52

Fixed effects:
                           Estimate Std. Error t value
(Intercept)                  53.926      2.991  18.031
grAcupuncture                 6.474      4.313   1.501
factor(time)2                 8.370      2.948   2.839
grAcupuncture:factor(time)2  10.830      4.252   2.547

Correlation of Fixed Effects:
            (Intr) grAcpn fct()2
grAcupunctr -0.693
factor(tm)2 -0.493  0.342
grAcpnc:()2  0.342 -0.493 -0.693
> VarCorr( mR )
 Groups   Name        Std.Dev.
 id       (Intercept) 11.143
 Residual             10.832
> ( tausq <- as.numeric( VarCorr( mR )$id ) )
[1] 124.1719
> ( sigsq <- attr( VarCorr( mR ), "sc" )^2 )
[1] 117.323
> ( rho <- tausq/(tausq+sigsq) )
[1] 0.5141802
```

Hence what we need to compute is:

```
> round( summary( mR )$coef, 2 )
                           Estimate Std. Error t value
(Intercept)                   53.93       2.99   18.03
grAcupuncture                  6.47       4.31    1.50
factor(time)2                  8.37       2.95    2.84
grAcupuncture:factor(time)2   10.83       4.25    2.55
```

```
> round( cf <- fixef( mR ), 2 )
                (Intercept)                grAcupuncture                factor(time)2
                      53.93                         6.47                         8.37
grAcupuncture:factor(time)2
                      10.83
> ( 1- rho ) * cf[2] + cf[4]
grAcupuncture
     13.97486
```

— also a little bit from the 12.7 in the conditional model.

Now if we fit a random effects model where we assume equal levels at baseline, that is just $\delta_g = 0$, we must hand-code the interaction:

```
> lg <- transform( lg, G2 = Relevel( interaction( gr, time ),
+                                    list( B=1:2 ) ) )
> with( lg, ftable( gr, time, G2 ) )
                 G2  B Placebo.2 Acupuncture.2
gr          time
Placebo     1       27         0             0
            2        0        27             0
Acupuncture 1       25         0             0
            2        0         0            25
> ms <- lmer( score ~ G2 + (1|id), data=lg )
> round( summary( ms )$coef, 3 )
                Estimate Std. Error t value
(Intercept)       57.038      2.172  26.262
G2Placebo.2        6.873      2.780   2.472
G2Acupuncture.2   20.817      2.874   7.242
> # CM <- rbind( diag(3), c(0,-1,1) )
> # rownames( CM ) <- c( names( fixef(ms) ), "Acp-eff" )
> # round( ci.lin( ms, ctr.mat=CM ), 2 )
```

— and we see it makes very little difference whether we fit the baseline difference or not.

The relevant parameters to compare are those from the models that purport to be the treatment effect in each of the models; for the sake of completeness we also compute the follow-up difference and the differences in change-score:

```
> round( ci.lin( lm( fu    ~ gr, data=acp ) )[2,], 2 )
Estimate   StdErr        z        P     2.5%    97.5%
   17.30     4.87     3.55     0.00     7.75    26.85
> round( ci.lin( lm( fu-bl ~ gr, data=acp ) )[2,], 2 )
Estimate   StdErr        z        P     2.5%    97.5%
   10.83     4.25     2.55     0.01     2.50    19.16
```

Treatment effect	Estimate	s.e.
Conditional	12.71	4.29
Random effects:		
different baseline	10.83	4.25
identical baseline	13.94	3.72
Follow-up difference	10.83	4.25
Changescore difference	17.30	4.87

So we see that allowing for different baseline gives the same s.e. as the conditional model but an estimate that deviates about 0.5 s.e., whereas the random effects model with identical baseline between the groups has a slightly smaller s.e. and an estimate that deviates about a third s.e. A fair summary would be that the three approaches in this case produces pretty much the same results.

References