Data Visualization in R (5 op)

Lecture 6, 12.10.2020

Jussi Palomäki
jussi.palomaki@helsinki.fi

Content

- Plotting from model predictions
  - ANOVA and ANCOVA models
  - Linear regression models
  - Logistic regression models

ANOVA and ANCOVA

- In an ANOVA, the estimated marginal means (EMMs) == actual means, but not so for standard errors:

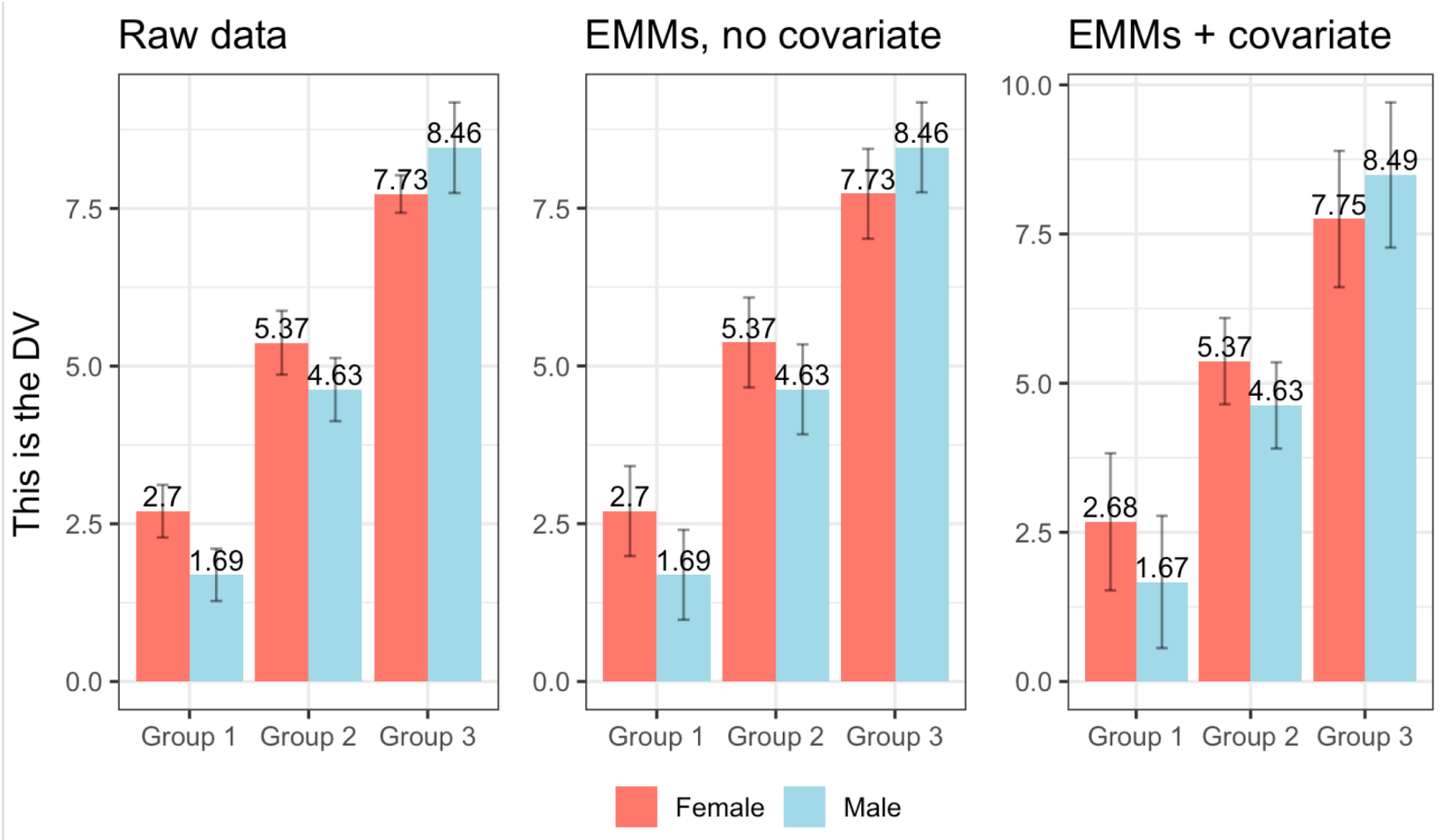
```
# A tibble: 6 x 6
# Groups:   group [3]
  group   gender mean_y  se_y lower upper
  <fct>   <fct>   <dbl> <dbl> <dbl> <dbl>
1 Group 1 Female   2.70 0.213  2.28  3.12
2 Group 1 Male     1.69 0.211  1.28  2.10
3 Group 2 Female   5.37 0.258  4.87  5.88
4 Group 2 Male     4.63 0.255  4.13  5.13
5 Group 3 Female   7.73 0.152  7.43  8.03
6 Group 3 Male     8.46 0.366  7.74  9.18
> y_emmeans
    group gender  emmean        SE df  lower.CL upper.CL
1 Group 1 Female 2.700173 0.3554439 54 1.9875504 3.412795
2 Group 2 Female 5.370822 0.3554439 54 4.6582001 6.083444
3 Group 3 Female 7.727571 0.3554439 54 7.0149488 8.440193
4 Group 1   Male 1.690750 0.3554439 54 0.9781274 2.403372
5 Group 2   Male 4.628614 0.3554439 54 3.9159922 5.341236
6 Group 3   Male 8.462937 0.3554439 54 7.7503152 9.175559
```

ANOVA and ANCOVA

- EMMs of an ANCOVA model give different SEs *and* means:

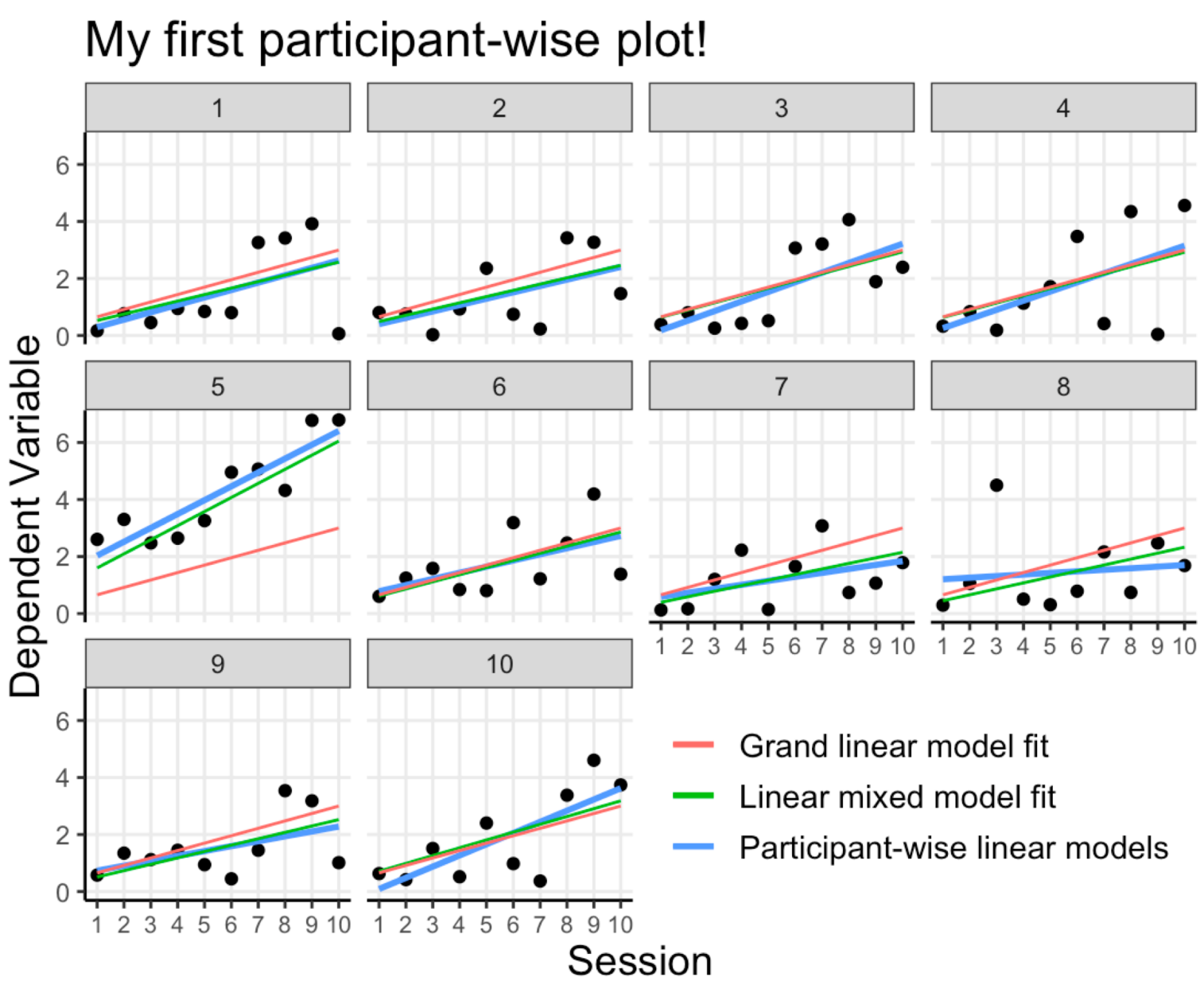
```
# A tibble: 6 x 6
# Groups:   group [3]
  group   gender mean_y  se_y lower upper
  <fct>   <fct>   <dbl> <dbl> <dbl> <dbl>
1 Group 1 Female   2.70 0.213  2.28  3.12
2 Group 1 Male     1.69 0.211  1.28  2.10
3 Group 2 Female   5.37 0.258  4.87  5.88
4 Group 2 Male     4.63 0.255  4.13  5.13
5 Group 3 Female   7.73 0.152  7.43  8.03
6 Group 3 Male     8.46 0.366  7.74  9.18
> y_emmeans
    group gender  emmean        SE df  lower.CL upper.CL
1 Group 1 Female 2.676459 0.5724137 53 1.5283422 3.824575
2 Group 2 Female 5.369035 0.3603428 53 4.6462792 6.091791
3 Group 3 Female 7.751108 0.5698289 53 6.6081760 8.894040
4 Group 1   Male 1.668411 0.5524966 53 0.5602431 2.776579
5 Group 2   Male 4.626899 0.3602205 53 3.9043878 5.349409
6 Group 3   Male 8.488955 0.6067913 53 7.2718861 9.706024
```

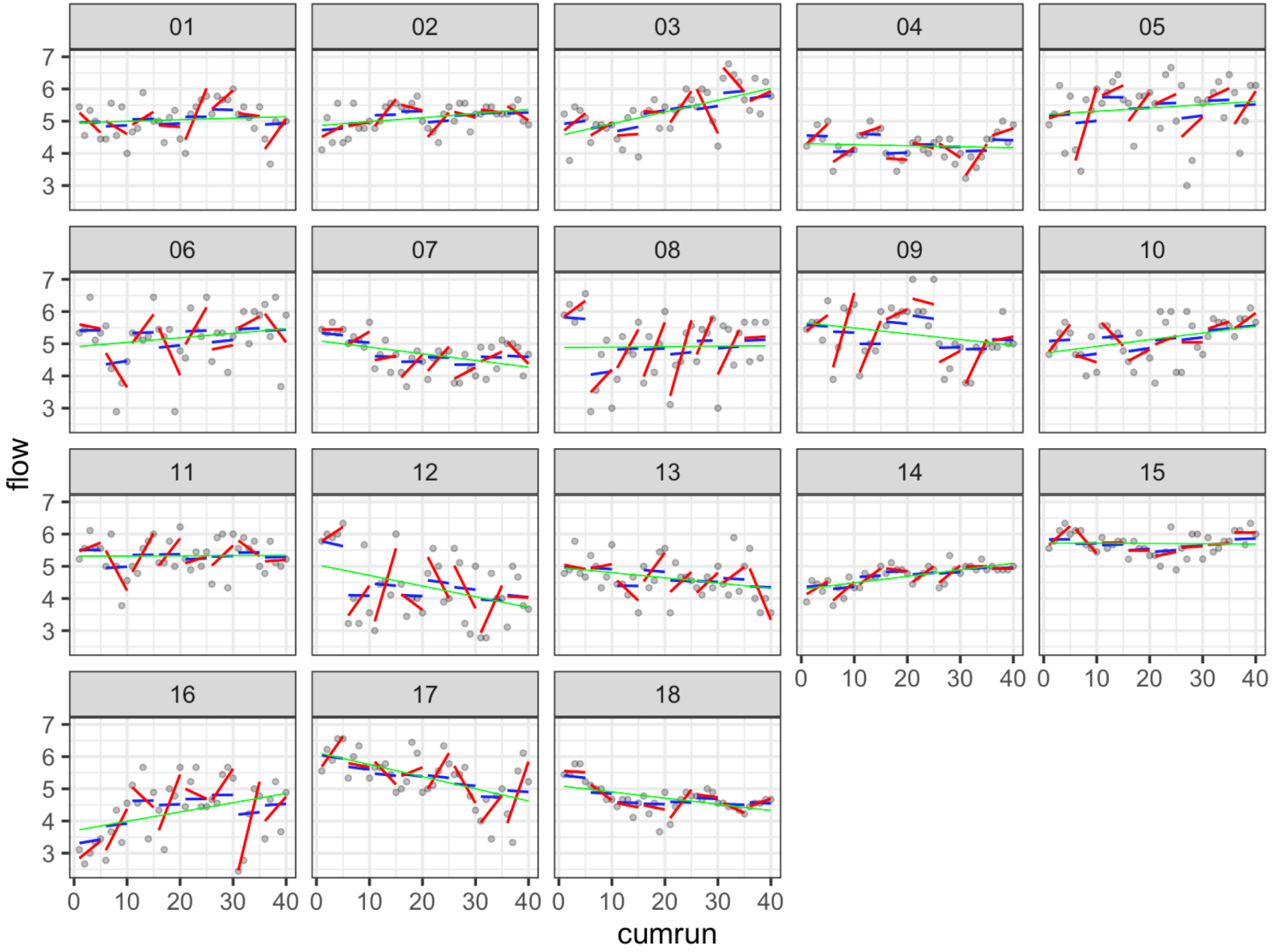
ANOVA and ANCOVA



ANOVA and ANCOVA

- Which is correct: plotting ”actual data” or ”data as modelled”?
- Which is more ”representative” of the world?
- Not a simple question!


My first participant-wise plot!
1
2
3
4
5
6
7
8
9
10
Dependent Variable
Session
Grand linear model fit
Linear mixed model fit
Participant-wise linear models


01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
flow
cumrun
0
10
20
30
40
3
4
5
6
7