LECTURE 15

Mixed-effects models I: random intercepts and slopes and linear regression

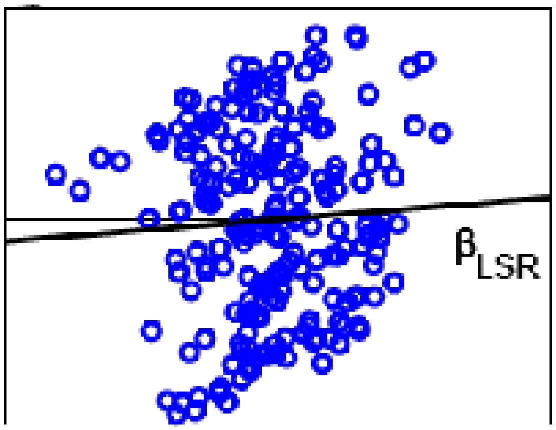
Random effects

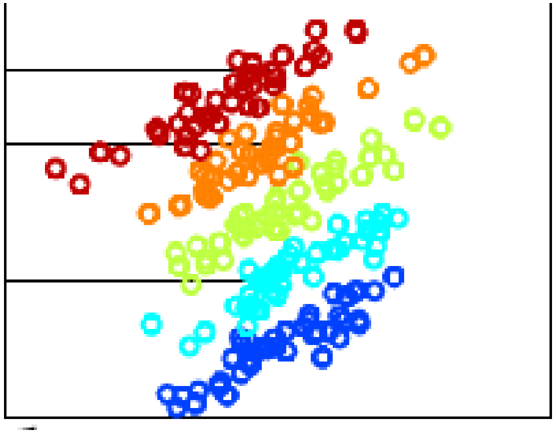
- All models we used so far are based on *fixed effects* (our usual predictors or independent variables)

- But models also come with an error term
- For example, a linear model is
- $y = a + bx + e$

(where e is the error in prediction
- Error *e* is what causes spread of points around the regression line

- The function of *random effects* is to structure the error term, based on information we have about the sample
  - in the example, information on subdivision of sample into villages (colours) suggests that error around regression line is mostly due to *variation in intercept* between villages

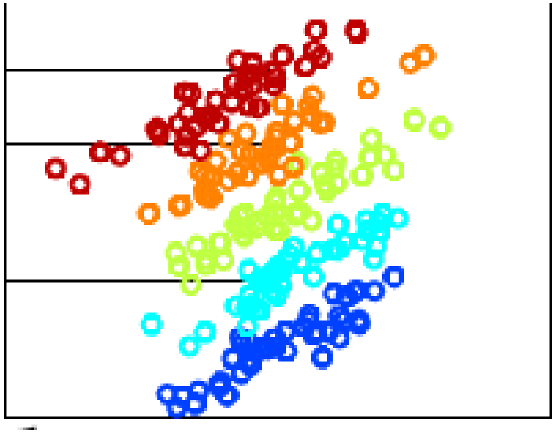




Mixed-effects models

- Mixed-effects models include both *fixed* and *random* factors

- Fixed factors affect only *mean* of response variable
  - our usual predictors
  - they are ‘informative’: we have an interpretation for levels of fixed factors
  - ‘fixed’ because they exhaust the population. i.e. all levels are covered
  - examples: sex (everybody is either male or female), age in a body growth study (ages are covering the interval of interest); body size in a brain size study



- Random factors affect only *variance* of response variable
  - add individual variation to predicted response
    - they can add variation to intercepts or slopes
  - levels are ‘uninformative’; they are only a sample of possibly levels
  - Examples: subject, location

- (Note: distinction is not clear-cut
  - location for example: I can establish that there are only 4 locations and use then as fixed factor levels)

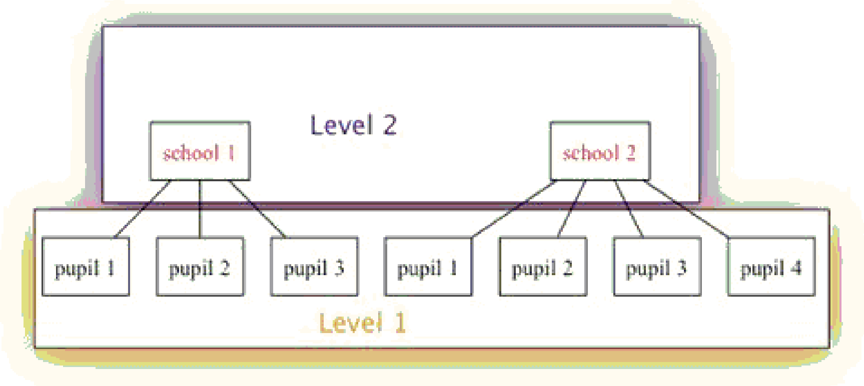
Pseudoreplication

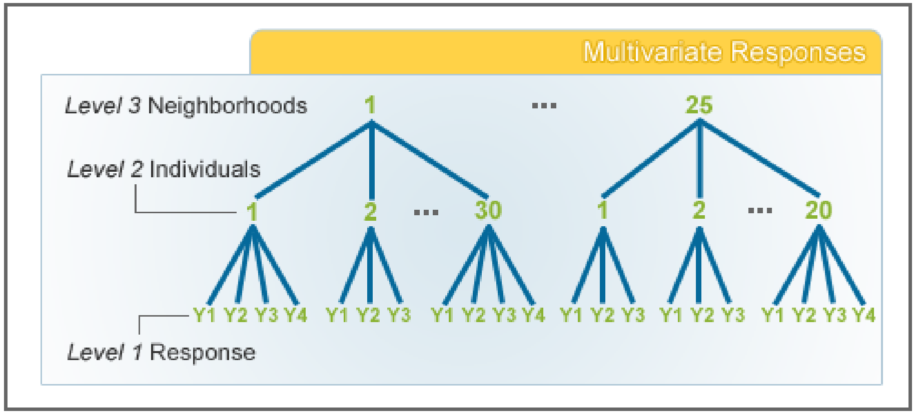
- Random effects deal with *pseudoreplication* by accounting for particularities of sample subgroups

- Pseudoreplication occurs when you analyse the data as if you had more degrees of freedom than you really have
  - pseudoreplicates have non-independent errors

- Temporal pseudoreplication involves repeated measurements from the same individual
  - repeated measures will be temporally correlated with one another
  - random effect: ID

- Spatial pseudoreplication involves several measurements from the same location, or origin
  - measurements will be spatially correlated
  - random effect: location, mother ID
  - special case: when locations are nested we have hierarchical structuring, and mixed-effects models become *multilevel modelling* (see next lecture)





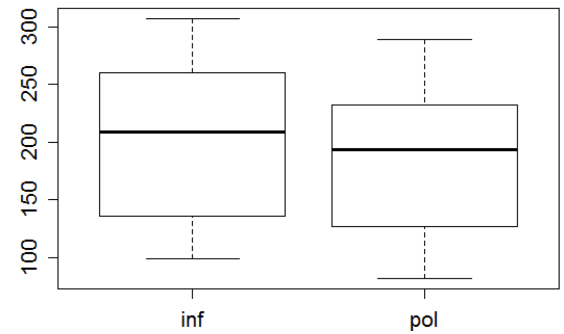
Example: random vs. fixed effects

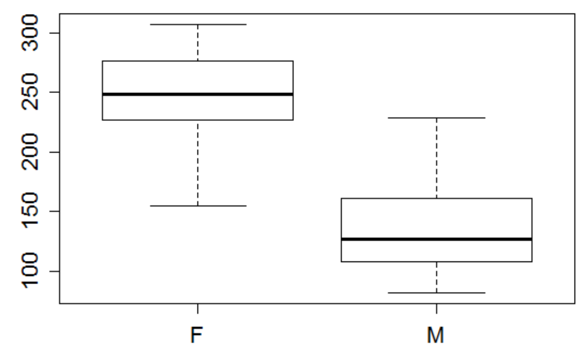
- A study of voice pitch vs. politeness
  - File *pitch*
  - Main question: does voice pitch change when people are polite vs. informal?

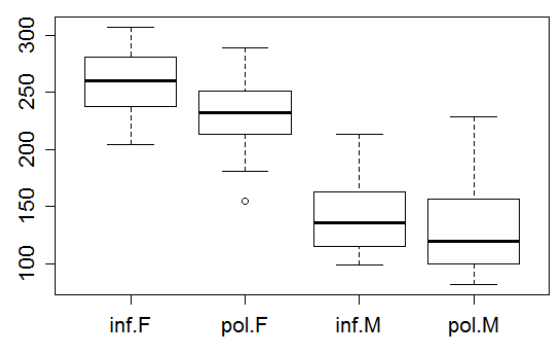
- If we just use a t-test of Wilcoxon test, difference is not significant

- `> t.test(pitch$frequency ~ pitch$attitude)`
- t = 1.2726, df = 80.938, p-value = 0.2068
- 95 percent confidence interval:
- -10.27285 46.73684
- mean in group inf mean in group pol
- 202.5881   184.3561

- `> wilcox.test(pitch$frequency ~ pitch$attitude)`
- W = 1015, p-value = 0.1621
- alternative hypothesis: true location shift is not equal to 0

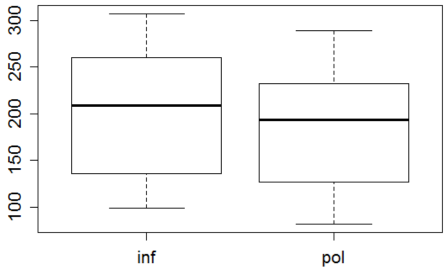




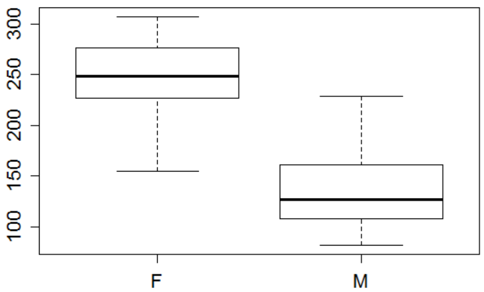


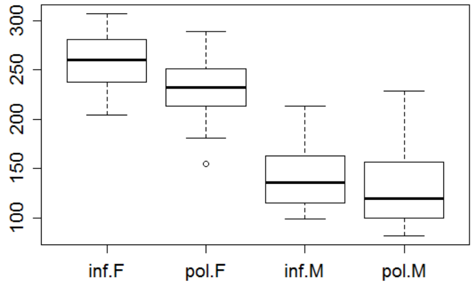
Example: random vs. fixed effects

- However, sample includes men and women and this should affect results



- So we want a model with two fixed effects:
- *pitch ~ attitude + gender + e*

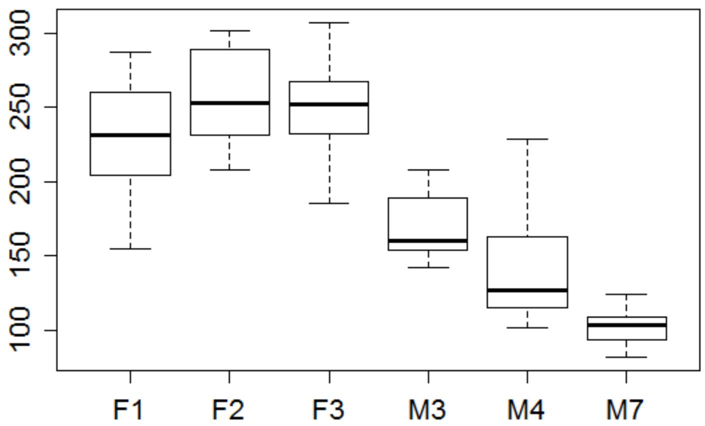




Example: random vs. fixed effects

- However, each subject is assessed 7 times and asked to provide a polite and an informal reply

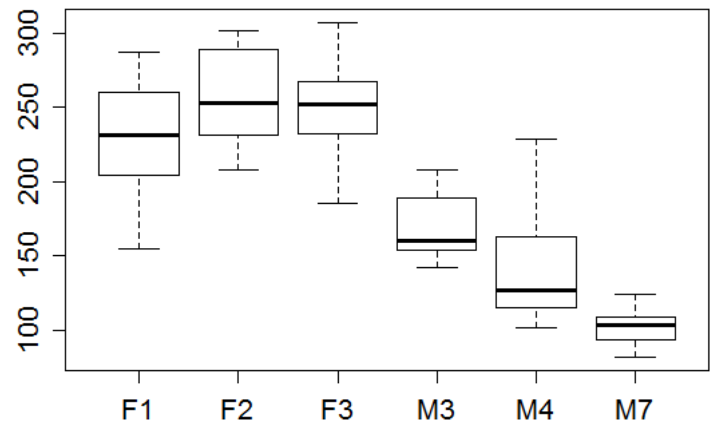
- Now we have pseudoreplication: each subject is providing 14 answers
  - particularities of each subject may add variation to pitch results
    - some may have naturally lower pitch

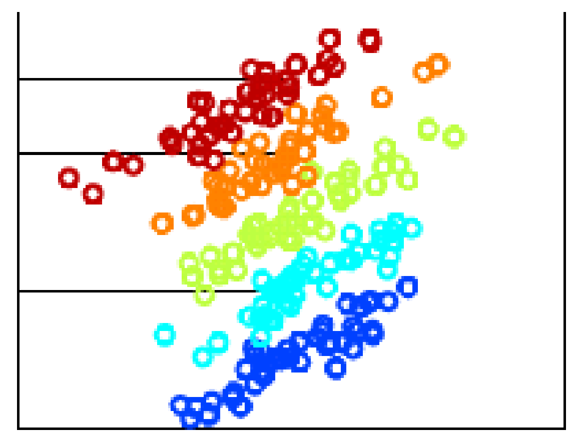


- We can introduce subject as a random effect in the model:
- *pitch ~ attitude + gender+ subject + e*

Random Intercept: subject

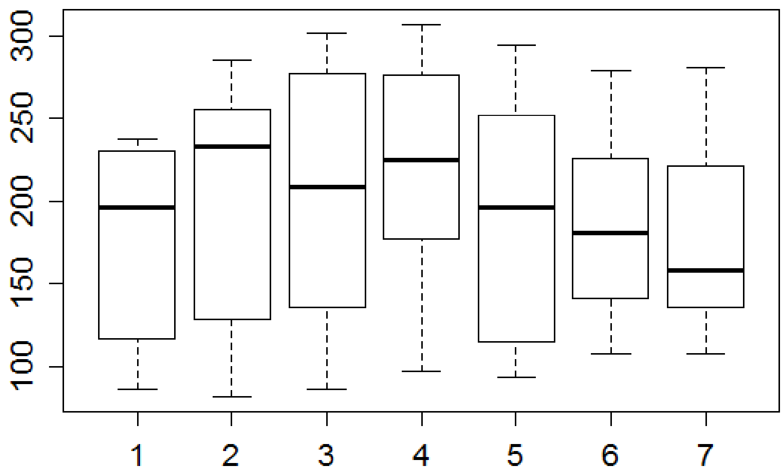
- We have an idea of how random subject peculiarities affect pitch: we assume each has a unique pitch, i.e. individuals differ by intercept
  - some people naturally have higher, some lower pitch
  - this pitch ‘height’ is an individual intercept
  - (notice we can talk of an intercept even though we are not running a regression)





Random Intercept: scenario

- But there is another source of non-independence: the 7 scenarios where subjects provide an informal and a polite reply
  - ‘asking for a favour’
  - ‘apologising for being late’
  - ‘telling someone off for being late’
  - scenario 3 may lower pitch more than others
  - scenarios would have particular intercepts too



- Therefore we want a model with:
- *pitch ~ attitude + gender + subject + scenario + e*

- This model accounts for by-subject and by-scenario variation in pitch

Mixed-effects models in R

- We need a random intercept model

- There are two packages and functions
  - *nlme* (function *lme*)
    - being gradually dropped
  - *lme4* (function *lmer*)
    - the package we'll be using

- Syntax: In *lmer*, fitting random intercepts by subject requires:
- (1 | subject)

- In our first model, let's have only *attitude* as fixed factor, plus random intercepts by subject and by scenario

```
> polite <- lmer(frequency ~ attitude +
(1|subject) + (1|scenario), data=pitch)
```

Random effects

```
> polite <- lmer(frequency ~ attitude
        + (1|subject) + (1|scenario), data=pitch)
> summary(polite)
Linear mixed model fit by REML
Data: pitch
   AIC   BIC logLik deviance REMLdev
 803.5 815.5 -396.7    807.1   793.5
Random effects:
 Groups   Name        Variance Std.Dev.
 scenario (Intercept)  218.98  14.798
 subject  (Intercept) 4014.54  63.360
 Residual              646.02  25.417
Number of obs: 83, groups: scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  202.588     26.750   7.573
attitudepol  -19.695      5.585  -3.527

Correlation of Fixed Effects:
            (Intr)
attitudepol -0.103
```

- *Random effects:*
- Variance column shows how much variation (error) is due to
  - between-scenario differences,
  - between-subject differences,
  - other residual factors (for example, within-subject variation; maybe subject 2 was tired during scenario 7)

- This is variation not explained by the fixed factors
  - equivalent to variation around a regression line; we are partitioning that random noise into categories 'subject', 'scenario', 'others'

- Most random effects are caused by between-subject variation in intercepts

Fixed effects

```
> polite <- lmer(frequency ~ attitude
        + (1|subject) + (1|scenario), data=pitch)
> summary(polite)
Linear mixed model fit by REML
Data: pitch
   AIC   BIC logLik deviance REMLdev
 803.5 815.5 -396.7    807.1   793.5
Random effects:
 Groups   Name        Variance Std.Dev.
 scenario (Intercept)  218.98  14.798
 subject  (Intercept) 4014.54  63.360
 Residual              646.02  25.417
Number of obs: 83, groups: scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  202.588     26.750   7.573
attitudepol  -19.695      5.585  -3.527

Correlation of Fixed Effects:
            (Intr)
attitudepol -0.103
```

- *Fixed effects:*
- Model fits one single coefficient for *attitude* taking random intercepts into account effects
- Intercept is an average pitch when *attitude*=baseline=informal
- Coefficient for *attitude*: -19.695
  - when *attitude*=pol, pitch drops by 19.695 Hz
- Note: *lmer* does no provide P values
  - we'll deal with significance later, but as a rule, just check for t-value:
  - t =-3.527 <-1.96, hence P < 0.05
  - also notice that model is fit by REML (restricted maximum likelihood); we'll need to change this to calculate significance

Intercept estimates

```
> coef(polite)
$scenario
  (Intercept) attitudepol
1    189.0772   -19.69454
2    209.1573   -19.69454
3    213.9799   -19.69454
4    223.1972   -19.69454
5    200.6423   -19.69454
6    190.4839   -19.69454
7    191.5789   -19.69454

$subject
   (Intercept) attitudepol
F1    241.4417   -19.69454
F2    267.2980   -19.69454
F3    259.9317   -19.69454
M3    179.0927   -19.69454
M4    154.7216   -19.69454
M7    113.0429   -19.69454
```

- Mixed model calculates one intercept for each scenario and for each subject
- Male subject 7 (M7) has a particularly low pitch
- Notice that fixed effect (the coefficient for *attitude*) is the same for all

Attitude + gender

```
> polite <- lmer(frequency ~ attitude + gender
        + (1|subject) + (1|scenario),
data=pitch)
> summary(polite2)
Linear mixed model fit by REML
AIC BIC l  ogLik deviance REMLdev
 787.5 802 -387.7    795.4   775.5
Random effects:
 Groups   Name        Variance Std.Dev.
 scenario (Intercept) 219.45   14.814
 subject  (Intercept) 615.57   24.811
 Residual             645.90   25.414
Number of obs: 83, groups: scenario, 7;
subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  256.846     16.114  15.940
attitudepol  -19.721      5.584  -3.532
genderM     -108.516     21.010  -5.165

Correlation of Fixed Effects:
            (Intr) atttdp
attitudepol -0.173
genderM     -0.652  0.004
```

- Now let's add fixed effect *gender* to our model
- Now total random effects are much smaller
  - mostly due to decrease in between-subject variance
  - that's because most of between-subject effects were gender differences
- Intercept:
- Now it is pitch for inf and female (=baseline)
- Coefficient for *attitude*:
- little change after introduction of *gender*
- Coefficient for *gender*:
- pitch drops -108.516 Hz in males
  - t =-5.2 < -1.96: significant effect

Significance

- There is a serious debate over P-values and statistical significance in mixed-effects models
  - mostly due difficulties in defining degrees of freedom
  - the older package *nlme* and function *lme* provide P-values but are being replaced by *lme4/lmer*

- One popular approach is to calculate P-values by comparing log-likelihood values of models through ANOVA
  - equivalent to comparing AIC values (=-log-likelihood+penalty factor)

- Example: to test for significance of *attitude*, we run
  - model without attitude (null model)
  - model with *attitude* (full model)
  - compare then through ANOVA

- *Important notes:*
- Only compared models if they have the same random effects structure
- When comparing models, we need to modify fitting method from Restricted ML to ML
  - *REML=F*

Significance

```
> polite.null <- lmer(frequency ~ gender + (1|subject) + (1|scenario),
+                     data=pitch, REML=F)
> polite.full <- lmer(frequency ~ attitude + gender
+                 + (1|subject) + (1|scenario),
 data=pitch, REML=F)
> anova(polite.null, polite.full)
            Df    AIC    BIC    logLik  Chisq Chi Df Pr(>Chisq)
polite.null  5  816.72 828.81 -403.36
polite.att   6  807.10 821.61 -397.55 11.618      1  0.0006532 ***
```

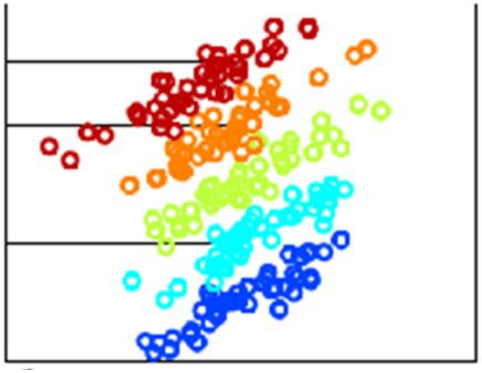
- Conclusion: model with *attitude* is significantly better
- Attitude significantly affects pitch (Chisq = 11.6, P < 0.0007); politeness significantly reduces pitch by 19.7 Hz
- Fixed effects interactions are allowed too
  - ANOVA tests or *drop1* function can be used as in linear and logistic regression models

Random slopes

```
> coef(polite2)
$scenario
   (Intercept) attitudepol   genderM
1     243.3398   -19.72111 -108.5164
2     263.4292   -19.72111 -108.5164
3     268.2541   -19.72111 -108.5164
4     277.4757   -19.72111 -108.5164
5     254.9102   -19.72111 -108.5164
6     244.6724   -19.72111 -108.5164
7     245.8426   -19.72111 -108.5164

$subject
    (Intercept) attitudepol   genderM
F1     242.9367   -19.72111 -108.5164
F2     267.2668   -19.72111 -108.5164
F3     260.3353   -19.72111 -108.5164
M3     285.2322   -19.72111 -108.5164
M4     262.2255   -19.72111 -108.5164
M7     223.0811   -19.72111 -108.5164
```

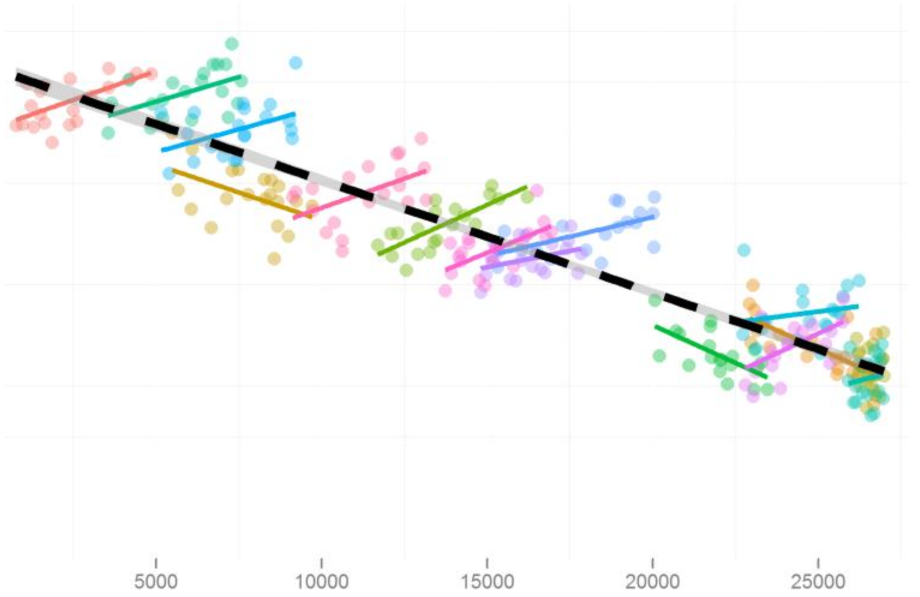
- So far we've fitted random intercept models
  - intercepts are allowed to be randomly variable across levels, but slope (the fixed effect) is the same across all groups)
- This means we're assuming that politeness lowers pitch uniformly across all individuals and scenarios



Random slopes

- But this is not necessarily true
- For example, politeness may affect M7's pitch more than F1's

- This means that random factors may affect slopes too
  - slopes of random groups may deviate from the general slope



Random slopes

- Random slope models structure errors around random levels both for intercepts and slopes

- Syntax:
- Keep fixed factors the same

- Random intercepts and slopes:
- (1 + fixed factor | random factor)
  - “1” is for random intercepts
  - your fixed factor adds random slope

- In our example: if we want random slopes for *attitude*, we must keep *attitude* as fixed factor *and* add random variation in slope by subject and by scenario
- (1 + attitude | subject) + (1 + attitude | scenario)

Random slopes

```
> pol.slope <- lmer(frequency ~ attitude + gender
+ (1+attitude|subject) + (1+attitude|scenario),
data=pitch, REML=F)
> summary(pol.slope)
Linear mixed model fit by maximum likelihood
   AIC   BIC  logLik deviance REMLdev
 814.9 839.1 -397.4    794.9   775.4
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 scenario (Intercept) 182.0743 13.4935
          attitudepol  31.2365  5.5890  0.219
 subject  (Intercept) 392.4726 19.8109
          attitudepol   1.7064  1.3063  1.000
 Residual             627.8901 25.0577
Number of obs: 83, groups: scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  257.987     13.529  19.069
attitudepol  -19.747      5.922  -3.335
genderM     -110.797     17.511  -6.327
```

- *Random intercepts:*
- Still a larger effect on pitch variance from subject than from scenario

- *Random slopes*
- Much less random variation in slopes than in intercepts
- Larger effect of *scenario* on slope of *attitude*

- *Fixed effects:*
- The fixed effects of *attitude* and *gender* are very similar

Random slopes

- > coef(pol.slope)
- $scenario
-   (Intercept) attitudepol   genderM
- 1   245.2581   -20.43858 -110.7974
- 2   263.2997   -15.94625 -110.7974
- 3   269.1402   -20.63232 -110.7974
- 4   276.8288   -16.30266 -110.7974
- 5   256.0557   -19.40611 -110.7974
- 6   246.8579   -21.94726 -110.7974
- 7   248.4671   -23.55537 -110.7974

- $subject
-   (Intercept) attitudepol   genderM
- F1   243.8045   -20.68210 -110.7974
- F2   266.7324   -19.17027 -110.7974
- F3   260.1485   -19.60439 -110.7974
- M3   285.6928   -17.92004 -110.7974
- M4   264.1937   -19.33766 -110.7974
- M7   227.3489   -21.76715 -110.7974

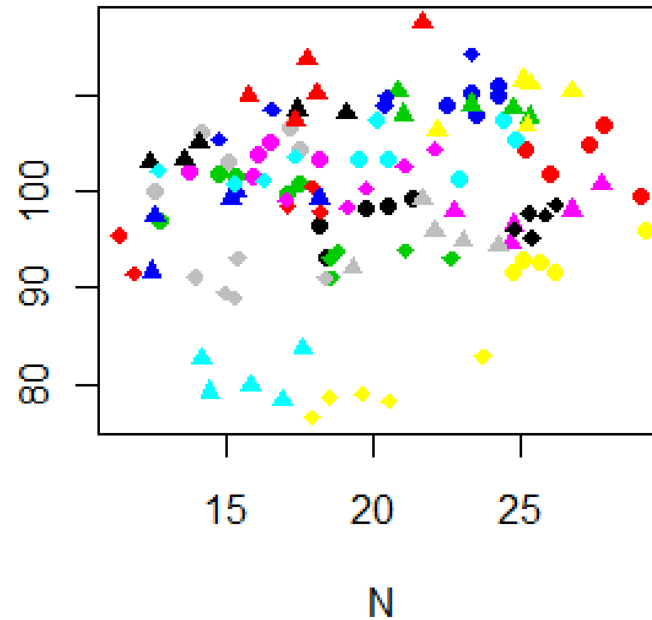
- Individual slopes by scenario and by subject are all very similar to the general estimate

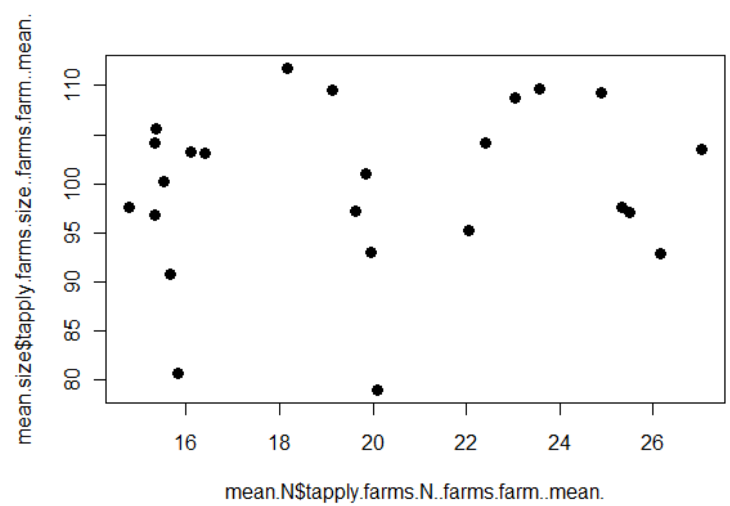
  (-19.72 Hz)

- However, you should include random slopes unless there is a justification for only using random intercepts

Mixed-effects linear regression

- We have so far used only categorical variables as fixed factors (*attitude* and *gender*)
- Let's apply mixed-effects models to linear regression
- Example: a regression of plant size against local measure of soil nitrogen
  - 24 farms
  - 5 measurements of different plant + Nitrogen per farm
- There is a significant regression of size on slope ($P < 0.04$)
- Problem: spatial pseudo replication
  - the 5 measurements per farm are not independent
  - plant size may differ by farm (random effects on intercept)
  - effect of Nitrogen on size may differ by farm (random effects on slope)
- Solution 1: averaging per farm
  - but this reduces sample from 120 to 24 points
  - result: no significant regression (P=0.53)





Mixed-effects linear regression

```
> linear.models <- lmList(size~N|farm,data=farms)
> summary(linear.models)
Coefficients:
   N
      Estimate Std. Error    t value    Pr(>|t|)
1    1.5153805  0.7339471  2.0646997 0.042554030
2   -0.5550273  0.6549563 -0.8474264 0.399565696
3    0.5551292  0.5171763  1.0733847 0.286682892
4    0.9212662  1.3333924  0.6909190 0.491838005
5    0.5380276  0.4231217  1.2715671 0.207619204
6    0.3845431  0.6373785  0.6033198 0.548191917
7    0.9339957  0.5574404  1.6755079 0.098172601
8    0.8220482  0.4813921  1.7076478 0.092010564
9    0.8842662  0.3554759  2.4875558 0.015179218
10   1.4676459  0.4579923  3.2045206 0.002017109
11  -0.2689370  0.4825728 -0.5572982 0.579052168
12   1.0138488  0.4251644  2.3846040 0.019733448
13   0.1324811  0.6584853  0.2011907 0.841116729
14   0.6551149  0.5085903  1.2880996 0.201835663
15   0.9809902  0.5974972  1.6418322 0.104985938
16   0.3699154  0.5391661  0.6860880 0.494861050
17   1.7555136  1.8365657  0.9558675 0.342337641
18   0.8715070  0.2944240  2.9600404 0.004162116
19   0.2043755  0.5360565  0.3812573 0.704135062
20   0.8567066  0.2910756  2.9432445 0.004368666
21   0.7830692  0.3654642  2.1426699 0.035520836
22   1.1441291  0.5138967  2.2263795 0.029118334
23   0.9536750  0.4343910  2.1954298 0.031356685
24   0.1091016  0.6047977  0.1803936 0.857350350
```

- Solution 2: one regression per farm
  - but then every regression has a sample size of 5 points
  - result: most regressions non-significant

Regression in mixed-effects model

- Solution 3: fitting one mixed-effects model, taking into account the random effects of farm on
  - standard deviation in intercept
  - standard deviation in slope

- First, let's fit a random intercept model only

```
>farm.model <-lmer(size~N+(1|farm), data=farms)
```

Regression in mixed-effects model

```
> farm.model <- lmer(size~N+(1|farm),
data=farms)
> summary(farm.model)
Linear mixed model fit by REML
Formula: size ~ N + (1 | farm)
   Data: farms
   AIC   BIC logLik deviance REMLdev
 614.4 625.5 -303.2    606.4   606.4
Random effects:
 Groups     Name        Variance Std.Dev.
 farm       (Intercept) 72.3545  8.5061
 Residual                3.7244  1.9299
Number of obs: 120, groups: factor(farm), 24

Fixed effects:
            Estimate Std. Error t value
(Intercept) 85.56754    2.54060   33.68
N            0.70875    0.09287    7.63

Correlation of Fixed Effects:
  (Intr)
N -0.727
```

- *Random effects:*
- Almost all residual variance is explained by variation in *intercepts* across farms

- *Fixed effects:*
- After controlling for farm effects on size baseline, now we find a significant and positive effect of Nitrogen on size!

Regression in mixed-effects model

```
> farm.slope <- lmer(size~  N+ (1+N|farm),data=farms)
> summary(farm.slope)
Linear mixed model fit by REML
Formula: size ~ N + (1 + N | farm)
   Data: farms
   AIC   BIC logLik deviance REMLdev
 617.8 634.5 -302.9    605.8   605.8
Random effects:
 Groups       Name        Variance   Std.Dev. Corr
 farm         (Intercept) 49.5020310 7.035768
              N            0.0056976 0.075482 1.000
 Residual                  3.6952189 1.922295
Number of obs: 120, groups: factor(farm), 24

Fixed effects:
            Estimate Std. Error t value
(Intercept) 85.82438    2.31787   37.03
N            0.69876    0.09301    7.51

Correlation of Fixed Effects:
  (Intr)
N -0.668
```

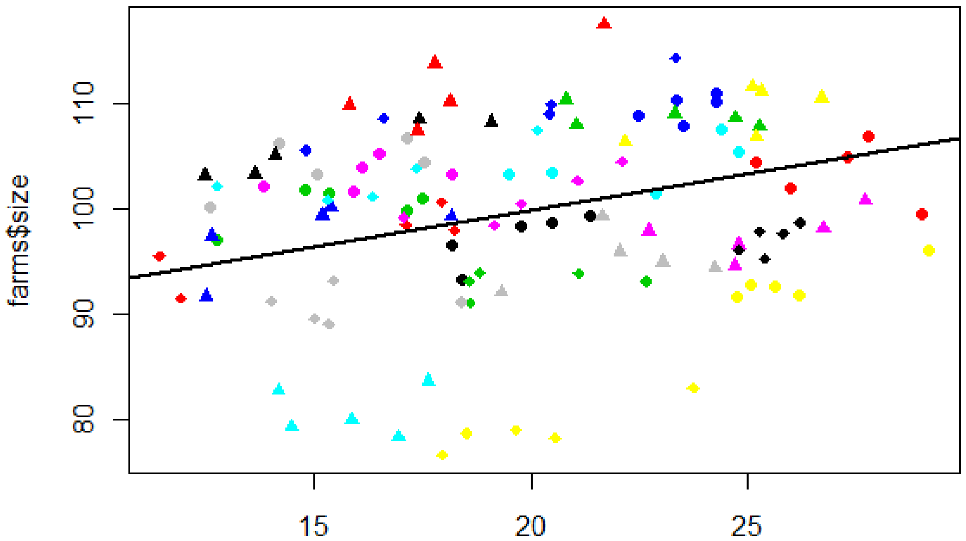
- Now let’s fit a random slopes model where both intercepts and slopes can account for residual variation

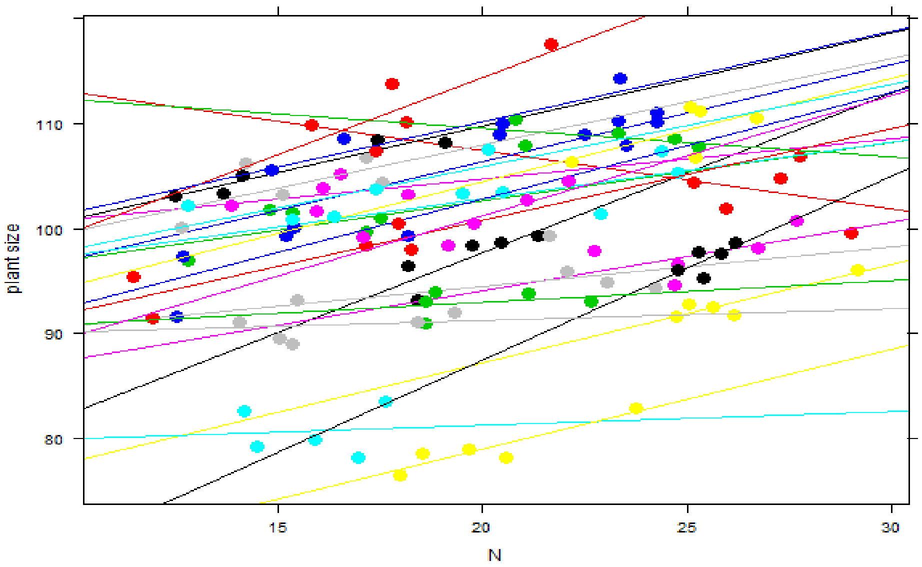
- *Random effects:*
- Variation in slope is very small
- Most random variation is explined by farm intercepts

- *Fixed effects:*
- Significance and value of Nitrogen intercept changed very little
- Introduction of random slopes did not add much to model

Regression in mixed-effects model

- So what is our regression then?
- Using the random slopes model, regression is given by the fixed effects
- *Size = 85.82 + 0.69876(N)*
- Random intercept model is not much different
  - we cannot compare log-likelihood or AIC between the random slope and random intercept models because they have different random-effect structure
- In this case, we use random slopes because it is more 'complete' i.e. it leaves no doubts as to possible slope effects





Conclusions

- We only use mixed models when we need to control for random effects; otherwise don't!
- Always se random slope models, unless advised to use only random intercepts
- Random effects may cause changes in interpretation. Possible outcomes:
  - Good outcome 1: fixed effects weren't significant, but become significant after controlling for random effects
  - Good outcome 2: fixed effects were significant and remain significant after controlling for random effects
  - Bad outcome: fixed effects were are no longer significant after controlling for random effects

Quiz

- File *simp*:
- Look at the plots of *ysimp* by *xsimp*, with colours representing 15 groups (variable *group*)

- Run a linear regression using the whole sample; is there a significant regression? What is it?
  - Add the regression line to scatterplot using abline
  - Using lmList, calculate the 15 separate regression lines. Do the slopes seem to vary across groups? What is the range of variation in slopes?
-
- Now run a random intercept linear regression of ysimp on xsimp with group as a random effect.
  - How much residual variance is explained by random effects of intercepts? Calculate it as (group intercept variance)/total variance=group intercept variance + residual)

- What is the regression coefficient for xsimp? Is it significant?
  - Compare it to the coefficient obtained through simple linear regression.
  - Plot another abline with the coefficients from the mixed model using command abline(a,b, lty=2)

- Now run a random slopes linear model
  - Afer controlling for random intercepts and coefficients and slopes, is the model significant?

- Plot points with their regression lines using command
  - xyplot (see Lecture 16, R code). Based on the linear regressions and mixed model lienar regresions, what do yo uconclude about the ralationship between xsimp and ysimp