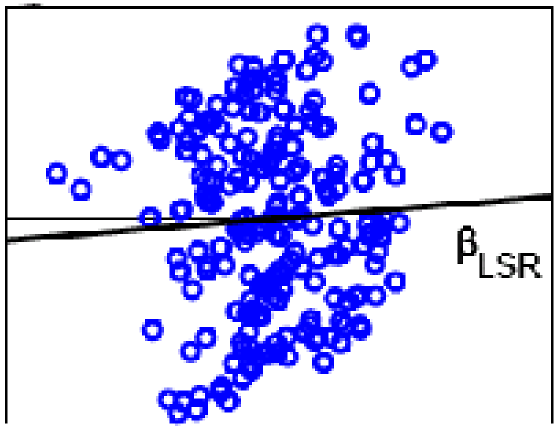
Lecture 3 BIO 206

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

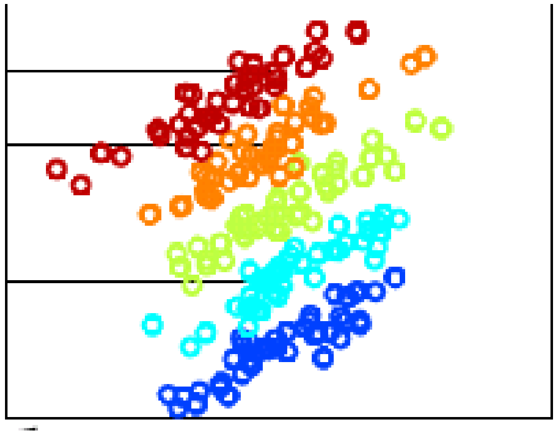
Random effects

- Look at top plot: it seems there is no significant relationship between y and x (slope b ~0)



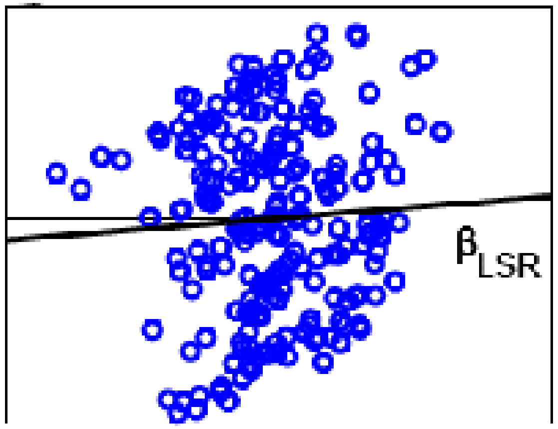
- But bottom plot shows you that points come from 5 different villages (colours)

- Now it seems there is a positive relationship between y and x within each village

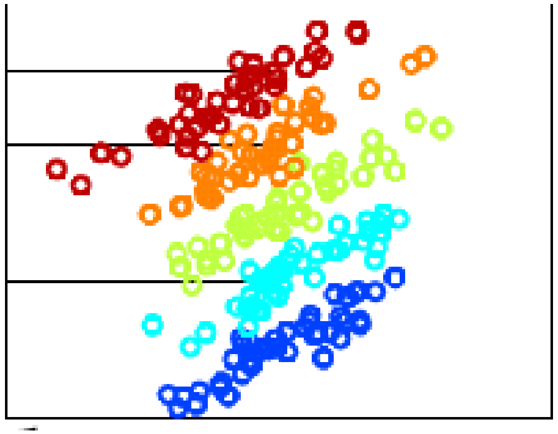


Pseudoreplication

- Pseudoreplication occurs when your data points are not independent from each other
  - They have a common or belong in the same group (=village)
  - This common feature causes them to have correlated deviations from estimates
  - =residuals are not independent
  - =fewer degrees of freedom

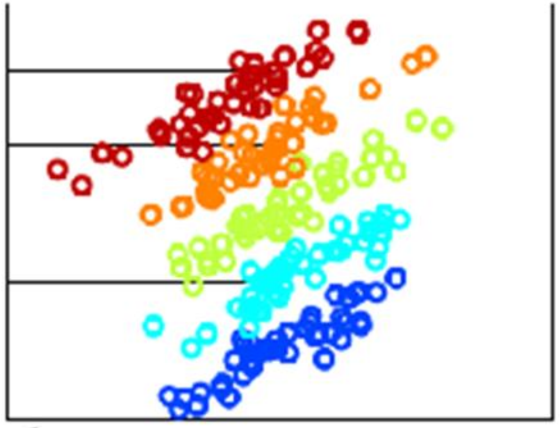


- The function of *random effects* is to structure the error term, to reveal the relationship between y (outcome) and x (a fixed effect)



Random effects

- Simple linear model only includes fixed factors (the x):
  - Fixed effect is the same for all cases



$$y = a + (b_x)x1 + (b_{village})x2$$

- Mixed effects models include fixed and random effects
  - Random effect adds error term to intercept or slope depending on which random group (=village) they belong to:

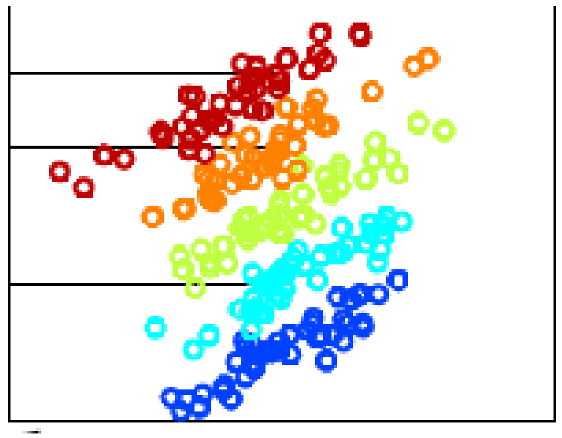
$$y = (a + a_{Random\_by\_village}) + (b + b_{Random\_by\_village})x$$

- Therefore it is very different to add village as a fixed or random effect

Mixed-effects models

- Fixed factors
  - our usual predictors (sex, age)
  - they are ‘informative’: we have an interpretation for levels of fixed factors
  - examples: sex (everybody is either male or female), age in a body growth study (ages are covering the interval of interest); body size

- Random factors
  - add variation to intercepts or slopes
  - levels are ‘uninformative’; they are only a sample of possibly levels
  - Examples: ID, location, experimenter



- Note: distinction is not clear-cut; location for example
  - can be a random effect (if I have no idea why they would differ, and don’t want to interpret differences but only control for possible effect)
  - or if I can name and study the locations, I may use them as fixed factor levels
  - If there are too many factors (500 instead of 5 locations/IDs etc), not practical to define them as fixed factors!

Sources of pseudoreplication (random factors)

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

- Spatial pseudoreplication: several measurements from the same location, group, origin etc
  - measurements will be spatially correlated
  - random effect: location, household ID

Fixed effects	Random effects
Drug administered or not	Genotype
Insecticide sprayed or not	Brood
Nutrient added or not	Block within a field
One country versus another	Split plot within a plot
Male or female	History of development
Upland or lowland	Household
Wet versus dry	Individuals with repeated measures
Light versus shade	Family
One age versus another	Parent

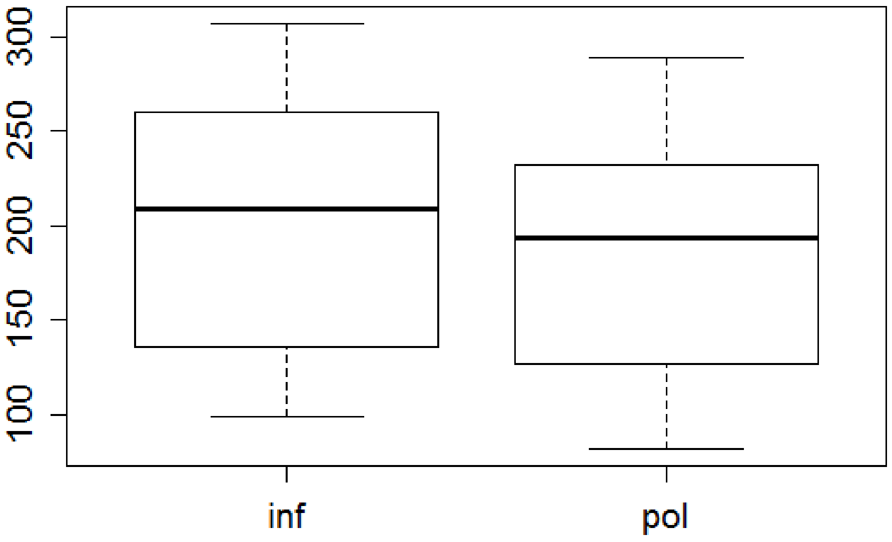
Example 1

- See R code

Example: pitch

- A study of voice pitch vs. politeness (*The R book*)
  - file *pitch*
  - question: does voice pitch differ when people are polite vs. informal?

- If we just apply a t-test or Wilcoxon test, difference is not significant



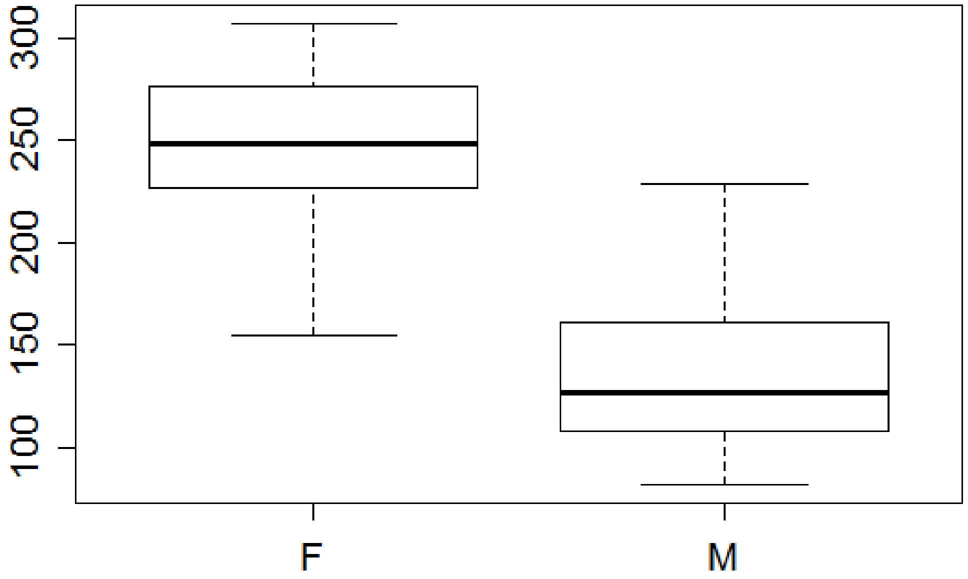

```
> t.test(pitch$frequency ~ pitch$attitude)
t = 1.2726, df = 80.938, p-value = 0.2068
95 percent confidence interval:
 -10.27285  46.73684
mean in group inf mean in group pol
         202.5881          184.3561


> 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

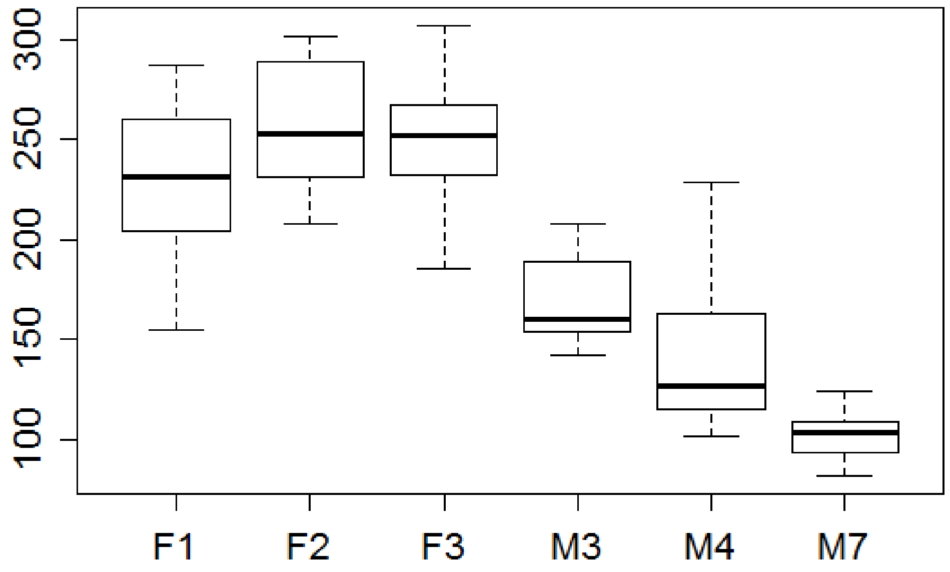
- So we want a model with two fixed effects:
- *pitch ~ attitude + gender + e*



Example: random vs. fixed effects

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

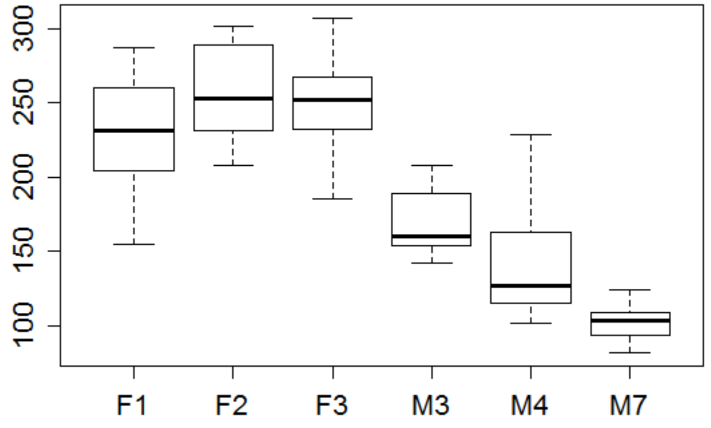
- Now we have pseudoreplication: each subject is providing 14 answers
  - particularities of each of the six subjects may add variation to pitch results
    - some may have naturally lower pitch

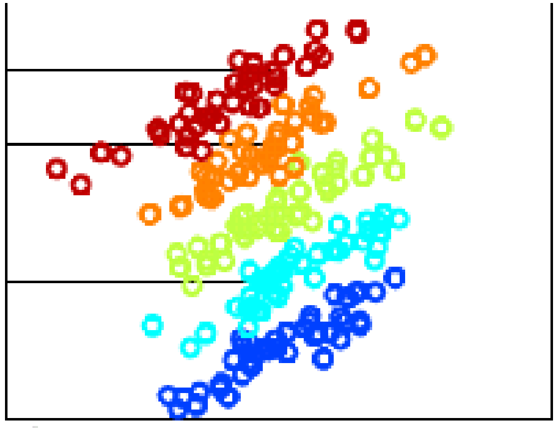


- We can introduce subject as a random effect in the model:
- *pitch ~ attitude + gender+ subject + e*

Random Intercept: subject

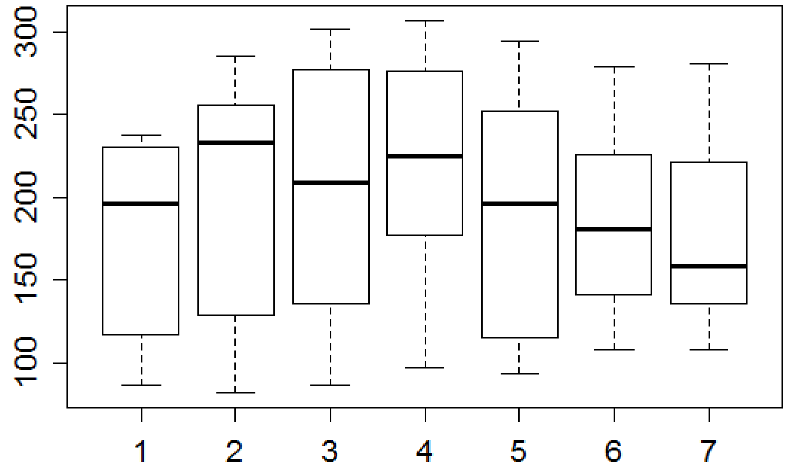
- We assume each has a unique (random) pitch, i.e. individuals differ by intercept
  - some people naturally have higher pitch
  - this pitch ‘height’ (or ‘level’) is an individual intercept
  - (notice we have an intercept even though we are not running a regression)





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
  - Mixed effects models in R: package *lme4* (function *lmer*)

- Syntax: fitting random intercepts by subject requires:
  - (1 | subject)

- In our first model, let's have only *attitude* (forget about *gender* for now) as fixed factor, plus random intercepts by
  - subject
  - 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
```

- *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 variation is 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
```

- *Fixed effects:*
- Model fits one single coefficient for *attitude* taking random intercepts into account

- 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

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
- Notice that fixed effect (the coefficient for *attitude*) is the same for all
- Male subject 7 (M7) has a particularly low pitch

Attitude + gender

```
> polite2 <- 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
```

- 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 (=new 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

- 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 them through ANOVA

- *Important notes:*
- Only compared models if they have the same random effects structure
- When comparing models, fitting method is automatically changed from Restricted ML to ML

Significance

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

> anova(polite.null, polite)
Data: pitch
Models:
polite.null: frequency ~ (1 | subject) + (1 | scenario)
polite: frequency ~ attitude + (1 | subject) + (1 | scenario)
            Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
polite.null  4 826.63 836.30 -409.31   818.63
polite       5 817.04 829.13 -403.52   807.04 11.586      1  0.0006646
```

- Conclusion: model with *attitude* is significantly better then without it
  - AIC

- *attitude* significantly affects pitch (Chisq = 11.6, P < 0.0007)
- being polite significantly reduces pitch by 19.7 Hz

- Interactions are allowed too
  - ANOVA tests or *drop1* function can be used as in linear and logistic regression models

Random slopes

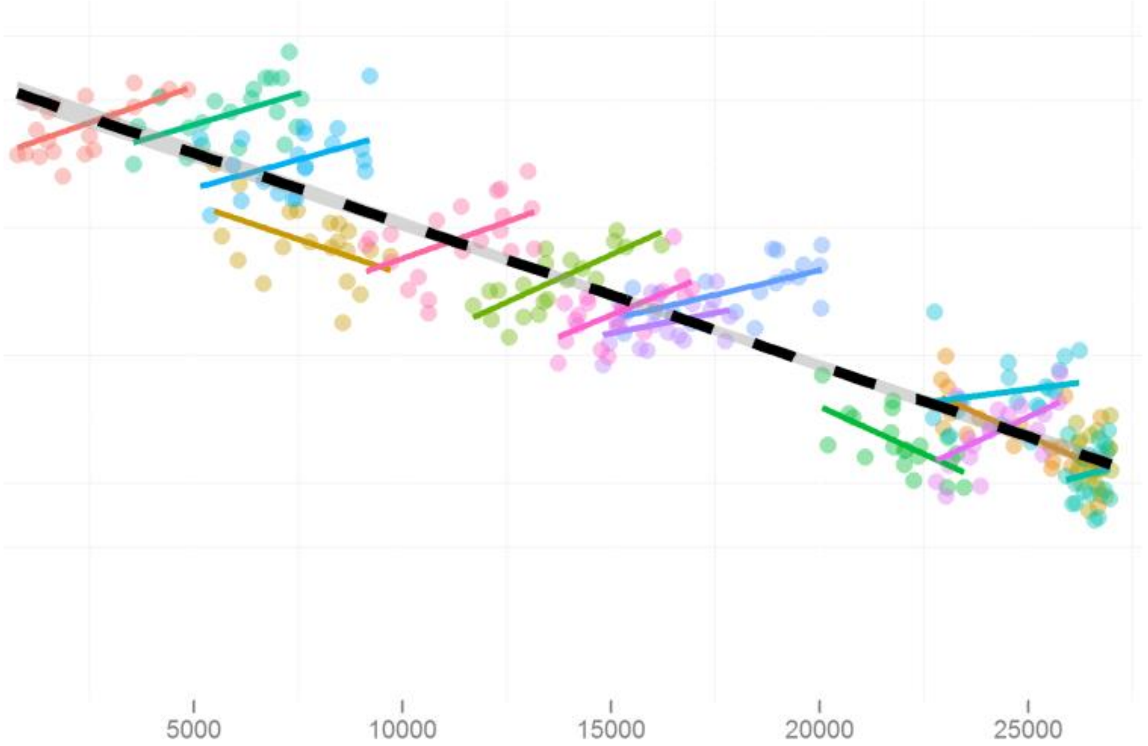
```
> coef(polite2)
$scenario
  (Intercept) attitudepol   genderM
1    243.3398   -19.72111 -108.5164
2    263.4292   -19.72111 -108.5164
3    268.2541   -19.72111 -108.5164
4    277.4757   -19.72111 -108.5164
5    254.9102   -19.72111 -108.5164
6    244.6724   -19.72111 -108.5164
7    245.8426   -19.72111 -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 randomly vary across levels, but slope (fixed effect) is the same across all levels)
- This means we're assuming that *attitude* (being polite) 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
- 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
    - “ + fixed factor” adds random slope

In our example: if we want random slopes for *attitude*, we 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)
```

- `> summary(pol.slope)`

```
> summary(pol.slope)
Linear mixed model fit by REML ['lmerMod']
Formula: frequency ~ attitude + gender + (1 + attitude |
subject) + (1 +
    attitude | scenario)
   Data: pitch
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 scenario (Intercept) 203.552  14.267
          attitudepol  71.068   8.430   0.00
 subject  (Intercept) 587.977  24.248
          attitudepol   1.477   1.216   1.00
 Residual             627.446  25.049
Number of obs: 83, groups:  scenario, 7; subject, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   258.14      15.86  16.272
attitudepol   -19.75       6.38  -3.096
genderM      -111.11      20.94  -5.305
```

- *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)
- You should include random slopes unless there is a justification for only using random intercepts

Mixed-effects linear regression

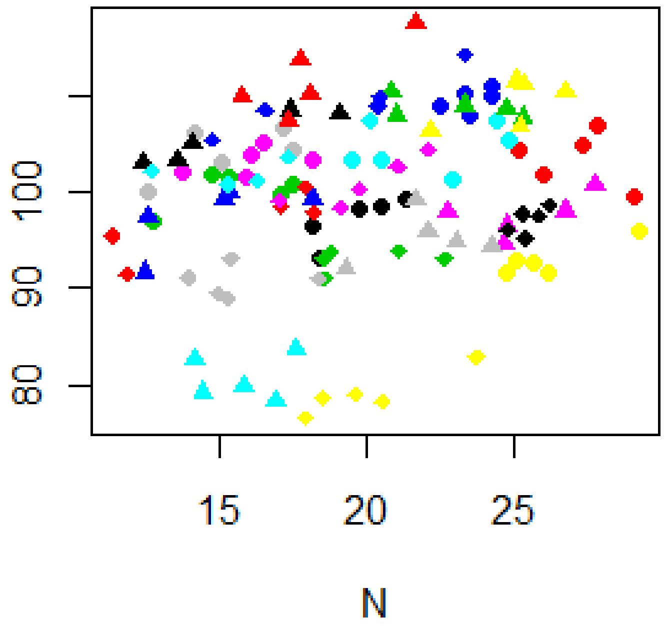
- We have so far used only categorical variables as fixed factors (*attitude* and *gender*)
- Let's apply mixed-effects models to linear regression

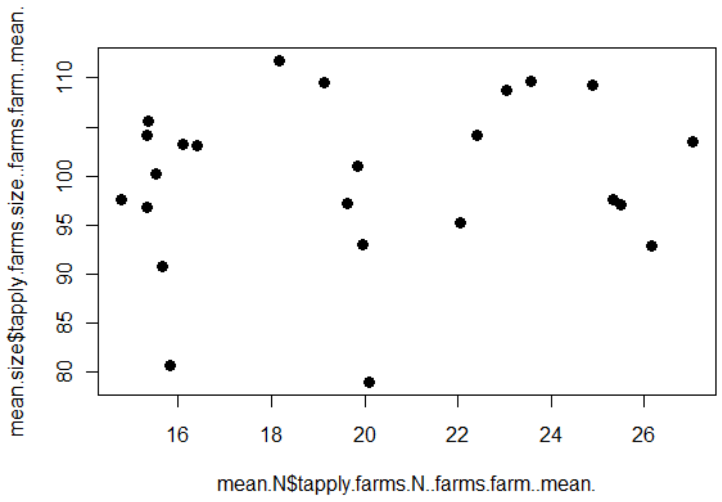
- Example: regression of plant size on a measure of soil nitrogen
  - 24 farms (farm is a factor)
  - 5 measurements of different plants and soil Nitrogen per farm

- Starting with a simple linear regression (size ~ N):
  - There is a significant effect of N on size ($P < 0.05$)

- Problem: spatial pseudo replication
  - the five 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

- 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 a mixed-effects model taking into account random effects of farm on
  - intercept
  - 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
```

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

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

- Now let's fit a random intercepts and slopes model

Random effects:

- Variation in slope is very small
- Most random variation is explained by farm intercepts

Fixed effects:

- Significance and value of N slope changed very little
- Introduction of random slopes did not add much to model

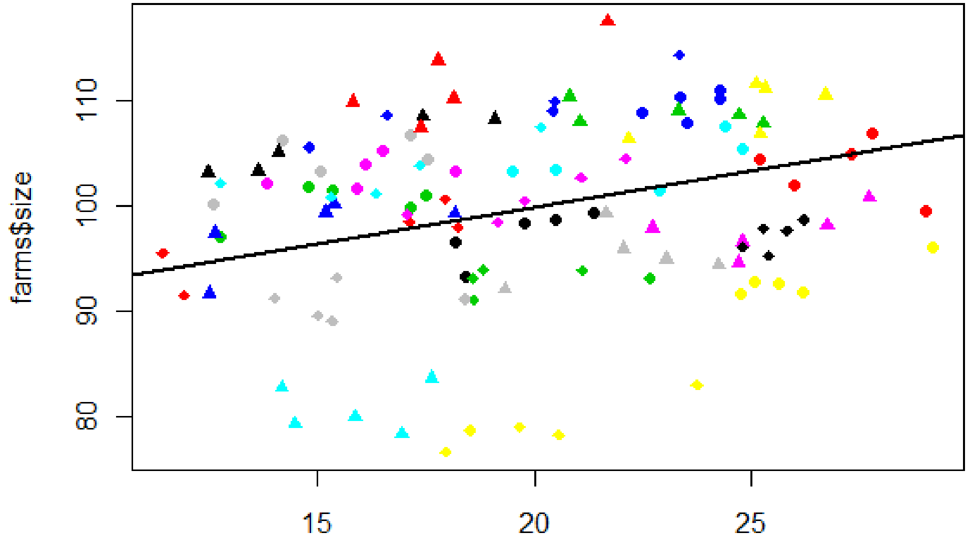
Regression in mixed-effects model

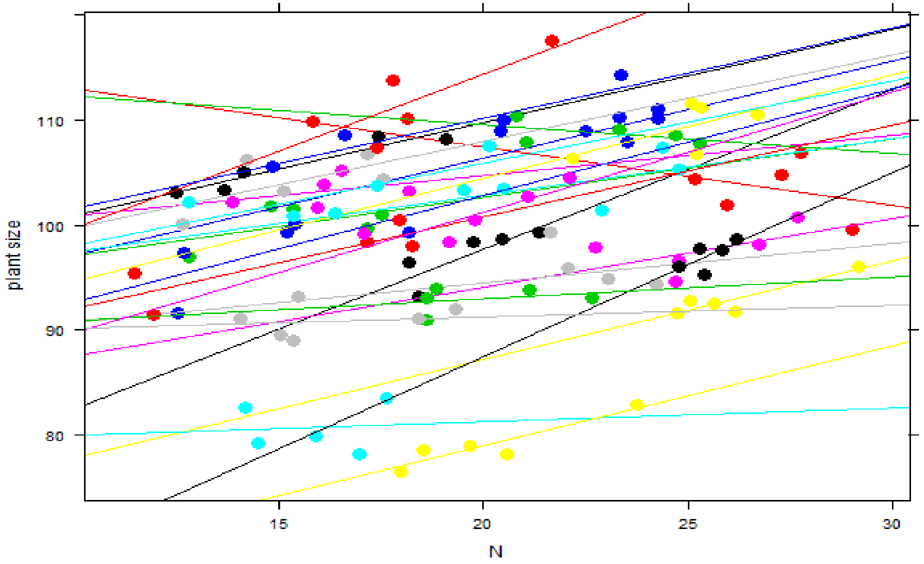
- So what is our regression then?
- Using the random slopes model, regression is given by:

$$Size = 85.82 + 0.69876(N)$$

- 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 structures

- 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 use random intercept and slope models in linear regressions, unless advised to use only random intercepts
- Random effects may cause changes in interpretation. Possible outcomes:
    - Best outcome: fixed effects weren’t significant, but become significant after controlling for random effects
    - Good outcome: fixed effects were significant and remain significant after controlling for random effects
    - Bad outcome: fixed effects are no longer significant after controlling for random effects

- Exercise: file *simp*:

1) Look at the plots of *ysimp* by *xsimp*, with colours representing 15 groups (variable *group*)

- Run a linear regression using the whole sample; is it significant? What is it?
  - Add 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?

2) 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 intercept effects? 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)

3) Finally run a random intercept and slopes linear model
  - After controlling for random intercepts and coefficients and slopes, is the model significant?