Lecture 6:
Prediction with mixed effects models

This lecture

1. Making predictions for new values (fixed and random effects included)
2. Making confidence intervals around predictions

(all in the context of normal response data)

Predictions with linear models (lm)

- Predictions of the model for the data used to fit the model is provided by
  > `fitted(model)`
- Predictions for new data are provided by
  > `predict(model, newdata=newdata)`
- Can get confidence intervals with
  > `predict(model, newdata=newdata,`
  `interval='confidence', level=0.95)`

Name of new data set

Model name

Type of interval

Width of interval

Code 6.1

lm model predictions

Predictions with lmer

predict() function an option BUT no confidence intervals given

```
> predict(mod, newdata=mynewdata,
    re.form=~0)
```

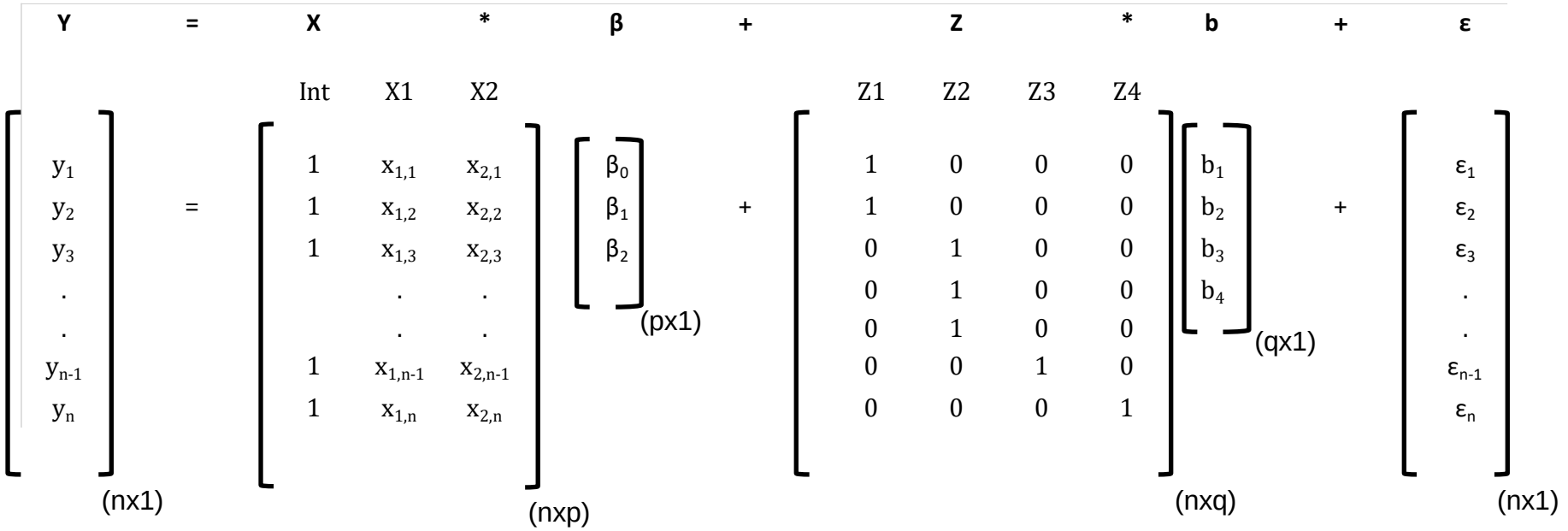
New predictor values

Which random effects do you want?
NULL = everything in model
~0 = no random effects, only fixed
~group = specify the groups you want

Code 6.2 part 1

lmer Model predictions

Recall: $y_i = X_i\beta + Z_i b_i + \varepsilon_i$ expanded



Matrix calculations (‘by hand’)

- Fixed effects part: $Y = X\beta$
- In R code this as
  `Y = X %*% beta`
- Can get beta from the fixed effects
- Need to create X – use `model.matrix()` command

model.matrix()

- model.matrix makes a model matrix given a formula and a data set
- Only need the right hand side of the formula **e.g.** `model.matrix(~x1+x2+ x3*x4, data=dat)`
- Automatically expands factors
- Make sure the formula in model matrix is **EXACTLY** the same as in your model

Making predicted values

- Once you have your model (mod) and model matrix (mat), do
  `pred <- mat %*% fixef(mod)`
- Ignores random effects. To add them, add the appropriate BLUP
- e.g. to add the data from block ‘3’ do
  `pred <- mat %*% fixef(mod) + ranef(mod)$block['3',]`
- This works for simple mixed effects models…

Code 6.2 part 2
lmer Model predictions

Confidence intervals

- This is tricky .
- Need to extract the variance-covariance (VCV) matrix for the fixed effects with `vcov(mod)`
- Then calculate the standard errors for the predictions as:
  $se = (X * VCV * X^t)^{1/2}$
- Do this in R with

```
pred.se <- sqrt(diag(mat %*% vcv %*% t(mat)))
```

Model matrix

Confidence interval for a new shadehouse

- It is more uncertain when you predict in a new shadehouse.
- So have to add extra variation – get this from the variance among groups.
- E.g., if there is a random effect for shadehouse, do

```
pred.se.new <- sqrt(diag(
      mat %*% vcv %*% t(mat) +
      VarCorr(mod)$shadehouse[1]))
```

Getting the right predictions and standard errors

	Mean	Variance
Known group	fixed effect + group effect	Residual variance
Unknown group	Fixed effect	Residual + random variance

Code 6.3
Confidence intervals

Exercise 6.1

- Use the biodepth data (biodepth.csv) to estimate the relationship between biomass and log(diversity)
- Use lme4's inbuilt functions to estimate
  1. The expected biomasses at an unknown site
  2. The expected biomasses at the Swiss site
- Now repeat this by hand!
- Can you get the same answers?

Exercise 6.2

- Calculate the **standard errors** for the predictions you just made
- Now calculate the confidence intervals.
- Remember that new block = more uncertainty