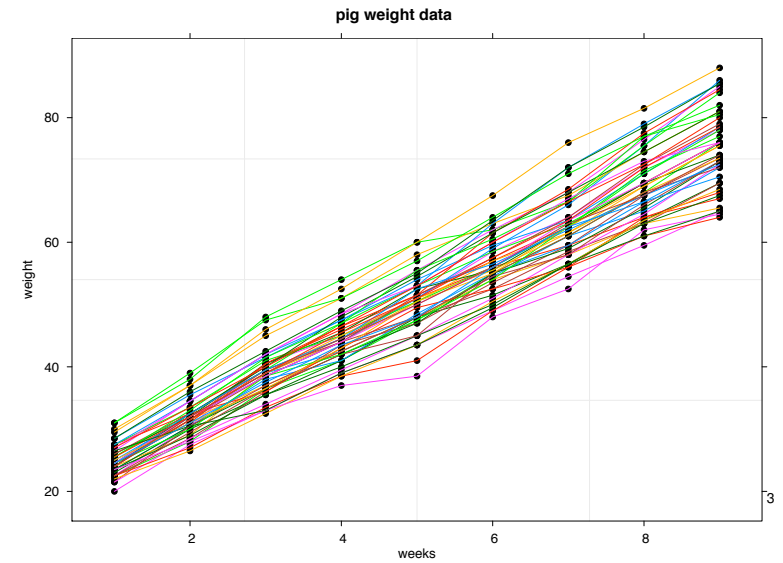


Residual Analysis and Prediction in Longitudinal Data Analysis

Matt Wand

1



Pig Weights Example

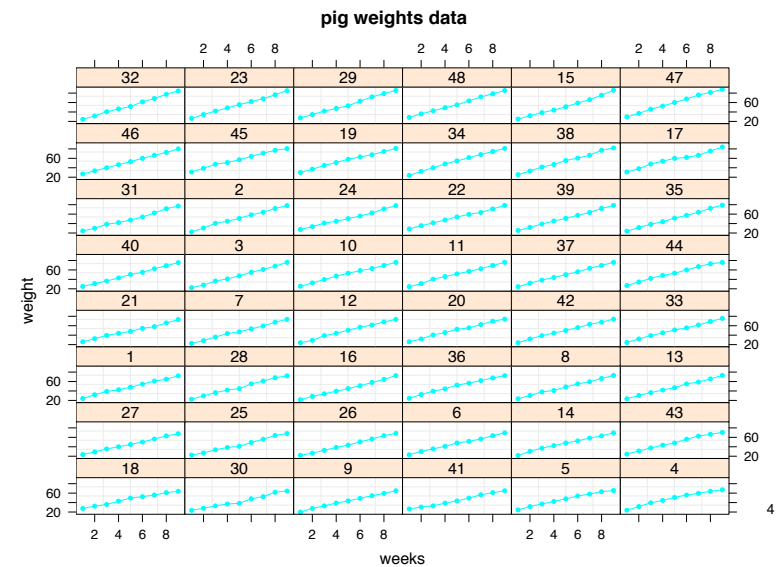
Recall the longitudinal data example involving

9 repeated measurements

on

48 (Victorian) pigs

2



4

Random Intercepts Model

$$\text{weight}_{ij} = U_i + \beta_0 + \beta_1 \text{weeks}_{ij} + \varepsilon_{ij}$$

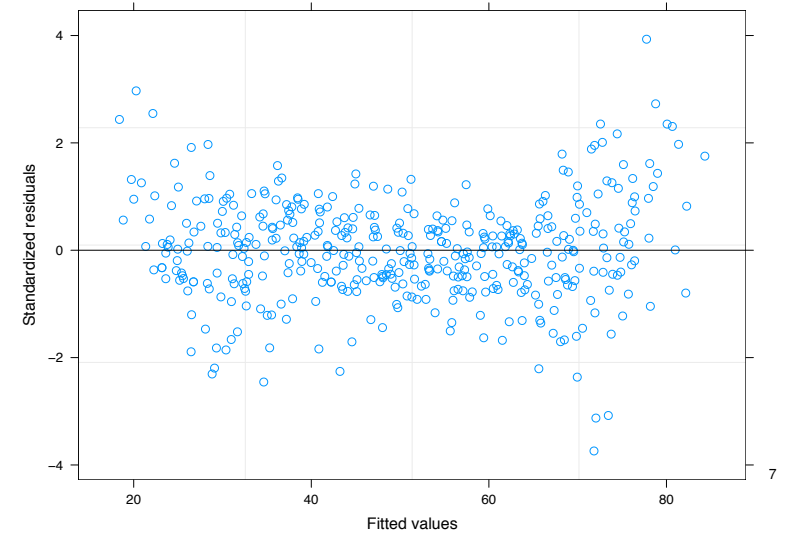
where

$$U_i \stackrel{\text{ind.}}{\sim} N(0, \sigma_U^2)$$

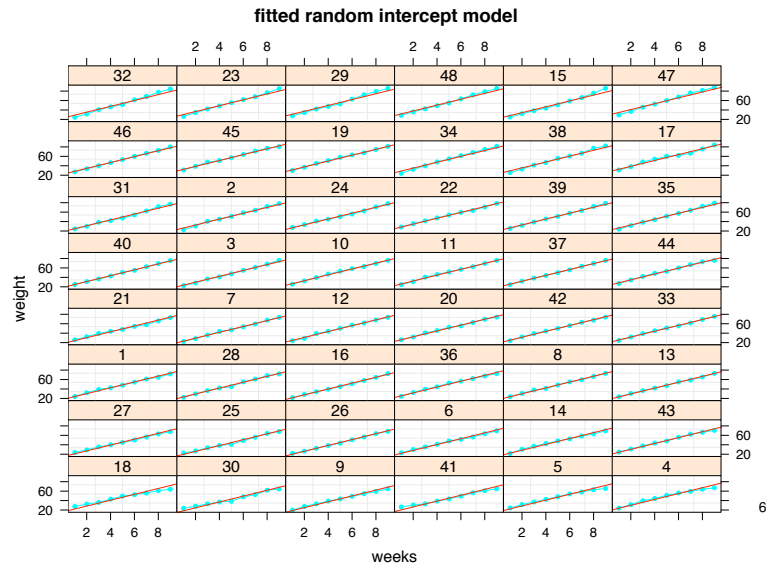
are independent of the

$$\varepsilon_{ij} \stackrel{\text{ind.}}{\sim} N(0, \sigma_\varepsilon^2).$$

5



7



6

Comments on Previous Two Slides

Close inspection of the fitted red lines shows that the parallel line assumption

inherent in the random intercepts model is too restrictive.

This is confirmed in the residual plot which shows a pronounced

bow tie pattern.

8

Random Intercepts and Slopes Model

A remedy is to allow **each pig** to have
his/her own slope...

9

Random Intercepts and Slopes Model II

$$\text{weight}_{ij} = U_i + \beta_0 + (\beta_1 + V_i) \text{weeks}_{ij} + \varepsilon_{ij}$$

where

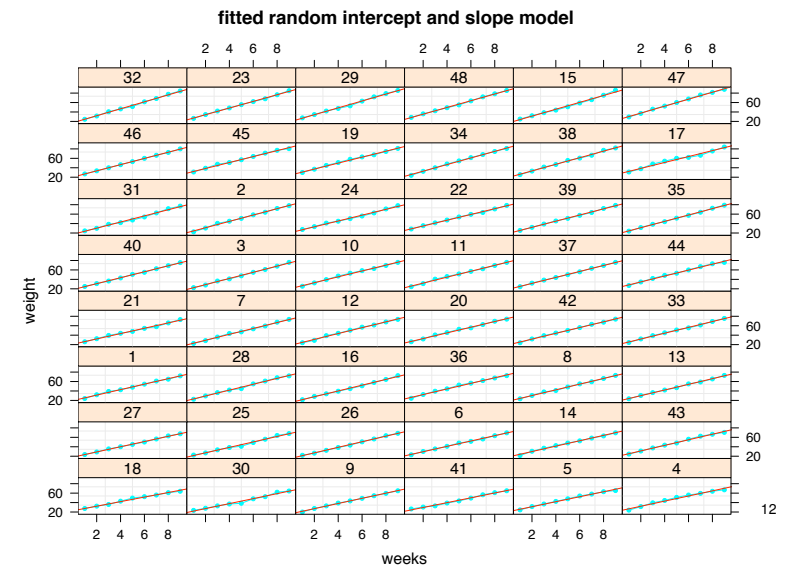
$$\begin{bmatrix} U_i \\ V_i \end{bmatrix} \stackrel{\text{ind.}}{\sim} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_U^2 & \rho_{UV}\sigma_U\sigma_V \\ \rho_{UV}\sigma_U\sigma_V & \sigma_V^2 \end{bmatrix} \right)$$

are independent of the

$$\varepsilon_{ij} \stackrel{\text{ind.}}{\sim} N(0, \sigma_\varepsilon^2).$$

10

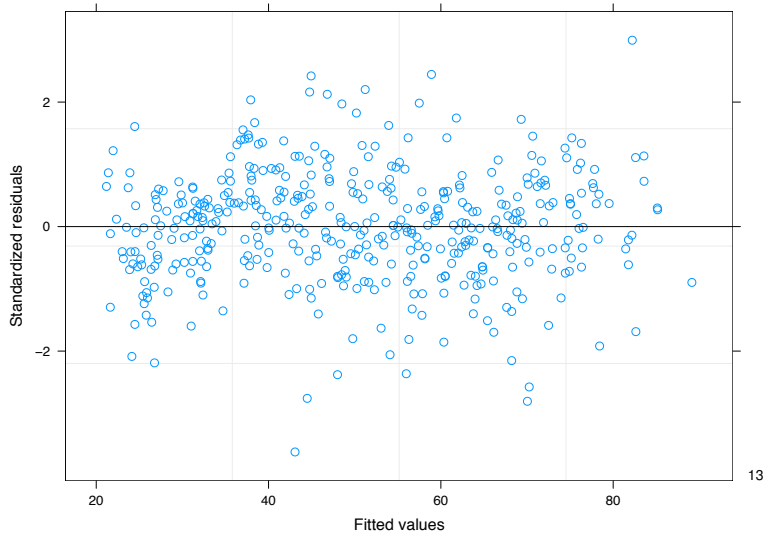
11



11

Fundamental Question

How are the **red lines** fitted?



15

Comments on Previous Two Slides

Close inspection of the fitted **red lines** shows that the lines have slightly different slopes

The **residual plot** is now showing
no systematic patterns

The second model is an improvement.

14

Partial Answer

The fitted line for the i th pig is

$$\hat{U}_i + \hat{\beta}_0 + (\hat{\beta}_1 + \hat{V}_i) \text{ weeks}$$

The **fixed effects** estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ are obtained via **maximum likelihood**.

How about the **random effects** counterparts: \hat{U}_i and \hat{V}_i ?

16

Answer

Maximum likelihood not defined for random effects.

We need to appeal to best prediction theory:

\hat{U}_i = best predictor of U_i given (weight) data

i.e.

$$\hat{U}_i = E(U_i|\mathbf{y})$$

where \mathbf{y} is the vector of weight observations.

Similarly for \hat{V}_i .

17

Whiteboard Interlude

We will now go through the mathematics of

best prediction of random effects

for general linear mixed models.

18