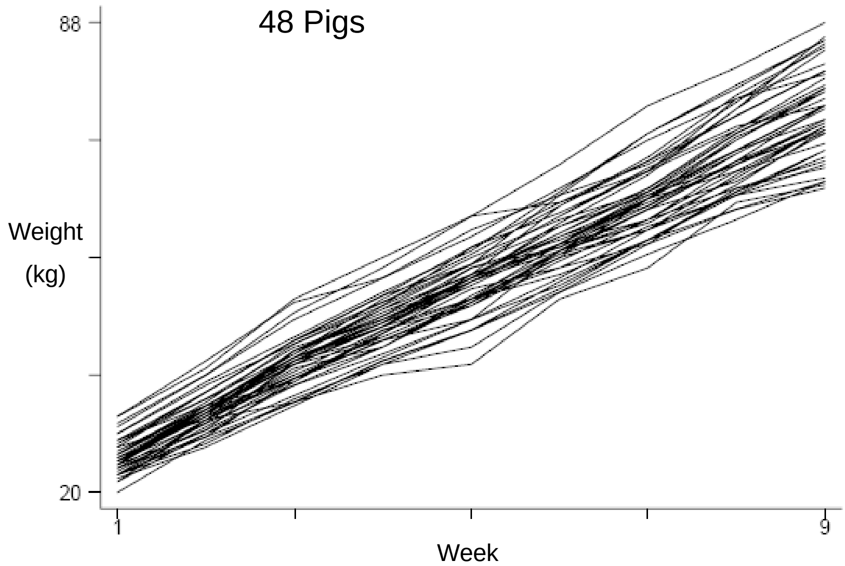
Lecture 3
Linear random intercept models

Example: Weight of Guinea Pigs

- Body weights of 48 pigs in 9 successive weeks of follow-up (Table 3.1 DLZ)
- The response is measures at n different times, or under n different conditions. In the guinea pigs example the time of measurement is referred to as a "within-units" factor. For the pigs n=9
- Although the pigs example considers a single treatment factor, it is straightforward to extend the situation to one where the groups are formed as the results of a factorial design (for example, if the pigs were separated into males and female and then allocated to the diet groups)

Pig Data



A) Linear model with random intercept

$$Y_{ij} = U_i + \beta_0 + \beta_1 t_j + \varepsilon_{ij}$$

$$U_i \sim N(0, \tau^2)$$ Variance between

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$ Variance within

$$\rho = \frac{\tau^2}{\tau^2 + \sigma^2}$$ Intraclass correlation coefficient

Pigs data model 1 – OLS fit

```
. regress weight time

      Source |       SS       df       MS              Number of obs =     432
-------------+------------------------------           F(  1,   430) = 5757.41
       Model |  111060.882     1  111060.882           Prob > F      =  0.0000
    Residual |  8294.72677   430  19.2900622           R-squared     =  0.9305
-------------+------------------------------           Adj R-squared =  0.9303
       Total |  119355.609   431  276.927167           Root MSE      =   4.392

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.209896   .0818409    75.88   0.000     6.049038    6.370754
       _cons |   19.35561   .4605447    42.03   0.000     18.45041    20.26081
------------------------------------------------------------------------------
```

OLS results

Pigs data model 1 – IND fit

```
. xtreg weight time, pa i(Id) corr(ind)

GEE population-averaged model                   Number of obs      =       432
Group variable:                         Id      Number of groups   =        48
Link:                             identity      Obs per group: min =         9
Family:                           Gaussian                     avg =       9.0
Correlation:                   independent                     max =         9
                                                Wald chi2(1)       =   5784.19
Scale parameter:                  19.20076      Prob > chi2        =    0.0000

Pearson chi2(432):                 8294.73      Deviance           =   8294.73
Dispersion (Pearson):             19.20076      Dispersion         =  19.20076

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.209896   .0816513    76.05   0.000     6.049862    6.369929
       _cons |   19.35561   .4594773    42.13   0.000     18.45505    20.25617
------------------------------------------------------------------------------
```

Independence correlation model results

Example: Weight of Pigs

For this type of repeated measures study we recognize two sources of random variation

1. Between: There is heterogeneity between pigs, due for example to natural biological (genetic?) variation
2. Within: There is random variation in the measurement process for a particular unit at any given time. For example, on any given day a particular guinea pig may yield different weight measurements due to differences in scale (equipment) and/or small fluctuations in weight during a day

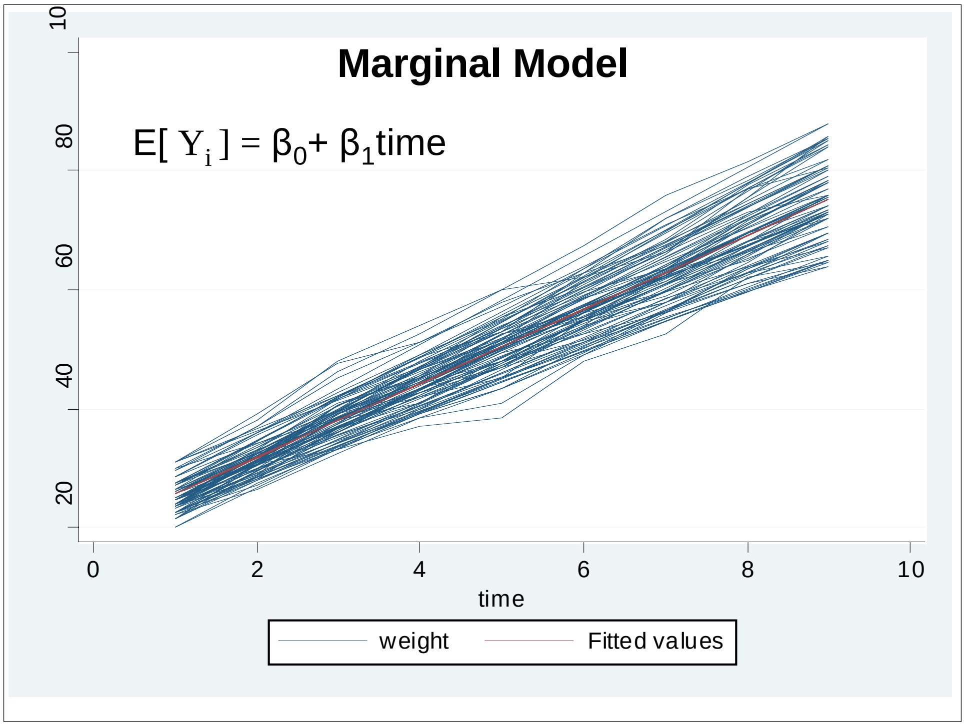
B) Marginal Model
With a Uniform correlation structure

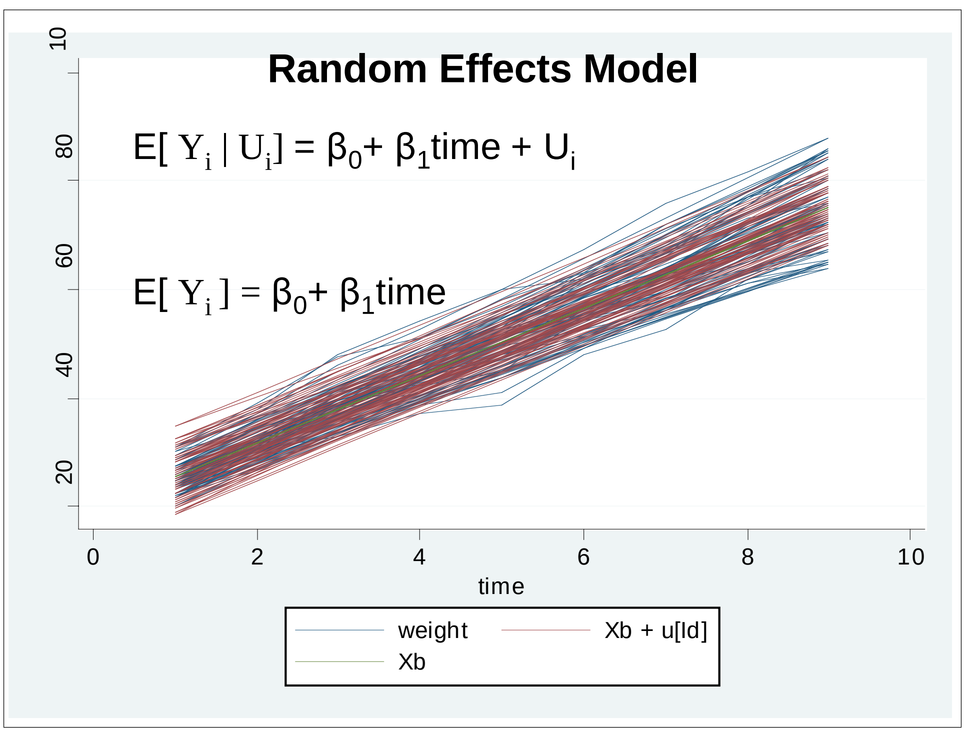
$$E[Y_{ij}] = \beta_0 + \beta_1 t_j$$

Model for the mean

$$\text{cov}(Y_{ij}) = (\tau^2 + \sigma^2)[\rho 1\,1' + (1-\rho)I]$$

Model for the covariance matrix


Marginal Model
E[Y_i] = β_0+ β_1time
20
40
60
80
10
0
2
4
6
8
10
time
weight
Fitted values


Random Effects Model
E[Y_i | U_i] = β_0+ β_1time + U_i
E[Y_i] = β_0+ β_1time
20
40
60
80
10
0
2
4
6
8
10
time
weight
Xb + u[Id]
Xb

Models A and B are equivalent

$$E[Y_{ij} \mid U_i] = U_i + \beta_0 + \beta_1 t_j$$

$$E[Y_{ij}] = E[E[Y_{ij} \mid U_i]] = \beta_0 + \beta_1 t_j$$

$$\text{cov}(Y_{ij}) = \text{cov}[E[Y_{ij} \mid U_i]] + E[\text{cov}[Y_{ij} \mid U_i]]$$

$$\text{cov}[E[Y_{ij} \mid U_i]] = \text{cov}(1U_i) = \tau^2 1 1'$$

$$E[\text{cov}[Y_{ij} \mid U_i]] = E[\sigma^2 I] = \sigma^2 I$$

$$\text{cov}(Y_{ij}) = (\tau^2 + \sigma^2)[\rho 1 1' + (1-\rho)I]$$

$$\rho = \frac{\tau^2}{\tau^2 + \sigma^2}$$

Pigs – Marginal model

```
xtreg weight time, pa i(Id) corr(exch)

Iteration 1: tolerance = 5.585e-15

GEE population-averaged model                   Number of obs      =       432
Group variable:                         Id      Number of groups   =        48
Link:                             identity      Obs per group: min =         9
Family:                           Gaussian                     avg =       9.0
Correlation:                  exchangeable                     max =         9
                                                Wald chi2(1)       =  25337.48
Scale parameter:                  19.20076      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.209896   .0390124   159.18   0.000     6.133433    6.286359
       _cons |   19.35561   .5974055    32.40   0.000     18.18472    20.52651
------------------------------------------------------------------------------
```

"Population Average", Marginal Model with Exchangeable Correlation structure results

Pigs – RE model

```
xtreg weight time, re i(Id) mle

Random-effects ML regression                    Number of obs      =       432
Group variable (i): Id                          Number of groups   =        48

Random effects u_i ~ Gaussian                   Obs per group: min =         9
                                                               avg =       9.0
                                                               max =         9

                                                LR chi2(1)         =   1624.57
Log likelihood  = -1014.9268                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.209896   .0390124   159.18   0.000     6.133433    6.286359
       _cons |   19.35561   .5974055    32.40   0.000     18.18472    20.52651
-------------+----------------------------------------------------------------
    /sigma_u |    3.84935   .4058114                      3.130767    4.732863
    /sigma_e |   2.093625   .0755471                       1.95067    2.247056
         rho |    .771714   .0393959                      .6876303    .8413114
------------------------------------------------------------------------------
```

"Population Average", Marginal Model with Exchangeable Correlation structure results

Pigs data model 1 – GEE fit

```
. xtgee weight time, i(Id) corr(exch)

Iteration 1: tolerance = 5.585e-15

GEE population-averaged model                   Number of obs      =       432
Group variable:                         Id      Number of groups   =        48
Link:                             identity      Obs per group: min =         9
Family:                           Gaussian                     avg =       9.0
Correlation:                  exchangeable                     max =         9
                                                Wald chi2(1)       =  25337.48
Scale parameter:                  19.20076      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.209896   .0390124   159.18   0.000     6.133433    6.286359
       _cons |   19.35561   .5974055    32.40   0.000     18.18472    20.52651
------------------------------------------------------------------------------
```

GEE fit – Marginal Model with Exchangeable Correlation structure results

Pigs data model 1 – GEE fit

```
. xtgee weight time, i(Id) corr(exch)
. xtcorr
```

Estimated within-Id correlation matrix R:

	c1	c2	c3	c4	c5	c6	c7	c8	c9
r1	1.0000								
r2	0.7717	1.0000							
r3	0.7717	0.7717	1.0000						
r4	0.7717	0.7717	0.7717	1.0000					
r5	0.7717	0.7717	0.7717	0.7717	1.0000				
r6	0.7717	0.7717	0.7717	0.7717	0.7717	1.0000			
r7	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	1.0000		
r8	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	1.0000	
r9	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	0.7717	1.0000

GEE fit – Marginal Model with
Exchangeable Correlation structure results

One group polynomial growth curve model

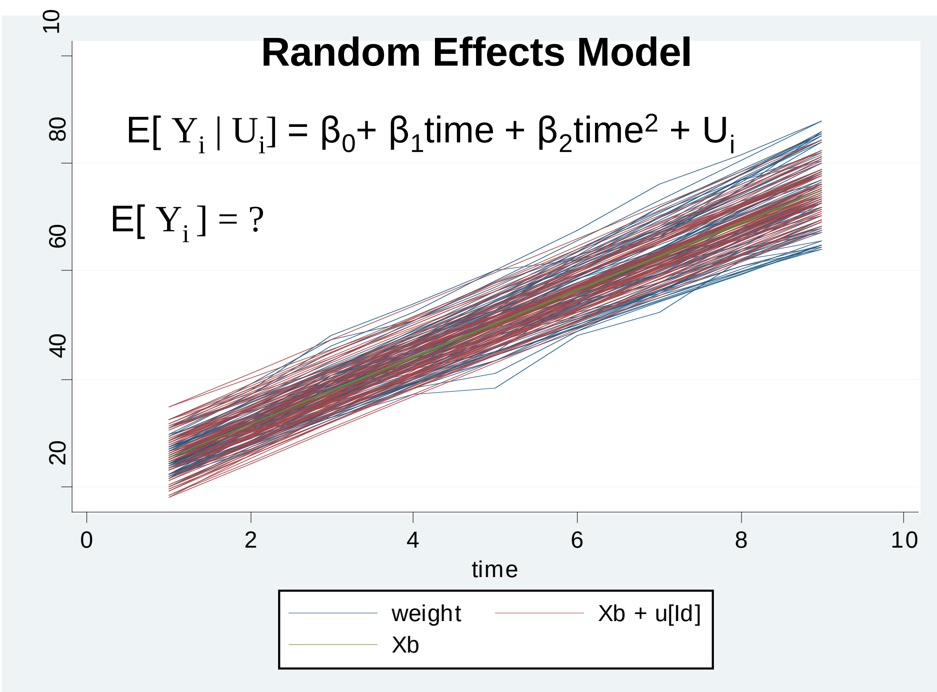
- Similarly, if you want to fit a quadratic curve
  $E[Y_{ij} \mid U_i] = U_i + \beta_0 + \beta_1 t_j + \beta_2 t_j^2$

$$E(\boldsymbol{Y}_i) = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ & \cdots & \\ 1 & t_n & t_n^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Pigs – RE model, quadratic trend

```
. gen timesq = time*time
. xtreg weight time timesq, re i(Id) mle

Random-effects ML regression                    Number of obs      =       432
Group variable (i): Id                          Number of groups   =        48

Random effects u_i ~ Gaussian                   Obs per group: min =         9
                                                               avg =       9.0
                                                               max =         9

                                                LR chi2(2)         =   1625.32
Log likelihood  = -1014.5524                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.358818   .1763799    36.05   0.000      6.01312    6.704516
      timesq |  -.0148922    .017202    -0.87   0.387    -.0486075    .0188231
       _cons |   19.08259    .675483    28.25   0.000     17.75867    20.40651
-------------+----------------------------------------------------------------
    /sigma_u |   3.849473   .4057983                      3.130909    4.732951
    /sigma_e |   2.091585   .0754733                      1.948769    2.244866
         rho |   .7720686   .0393503                      .6880712    .8415775
------------------------------------------------------------------------------
```

Exchangeable Correlation structure results



Pigs – Marg. model, quadratic trend

```
. xtgee weight time timesq, i(Id) corr(exch)

GEE population-averaged model                   Number of obs      =       432
Group variable:                         Id      Number of groups   =        48
Link:                             identity      Obs per group: min =         9
Family:                           Gaussian                     avg =       9.0
Correlation:                  exchangeable                     max =         9
                                                Wald chi2(2)       =  25387.68
Scale parameter:                  19.19317      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.358818   .1763801    36.05   0.000     6.013119    6.704517
      timesq |  -.0148922    .017202    -0.87   0.387    -.0486076    .0188231
       _cons |   19.08259   .6754833    28.25   0.000     17.75867    20.40651
------------------------------------------------------------------------------
```

Exchangeable Correlation structure results

Pigs data model 1 – GEE fit

```
. xtcorr

Estimated within-Id correlation matrix R:

        c1      c2      c3      c4      c5      c6      c7      c8      c9
r1  1.0000
r2  0.7721  1.0000
r3  0.7721  0.7721  1.0000
r4  0.7721  0.7721  0.7721  1.0000
r5  0.7721  0.7721  0.7721  0.7721  1.0000
r6  0.7721  0.7721  0.7721  0.7721  0.7721  1.0000
r7  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  1.0000
r8  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  1.0000
r9  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  0.7721  1.0000
```

GEE fit – Marginal Model with
Exchangeable Correlation structure results

Pigs – Marginal model: AR(1)

```
xtgee weight time, i(Id) corr(AR1) t(time)

GEE population-averaged model                   Number of obs      =       432
Group and time vars:                  Id time   Number of groups   =        48
Link:                                identity   Obs per group: min =         9
Family:                              Gaussian                  avg =       9.0
Correlation:                            AR(1)                  max =         9
                                                Wald chi2(1)       =   6254.91
Scale parameter:                     19.26754   Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.272089   .0793052    79.09   0.000     6.116654    6.427524
       _cons |   18.84218   .6745715    27.93   0.000     17.52004    20.16431
------------------------------------------------------------------------------
```

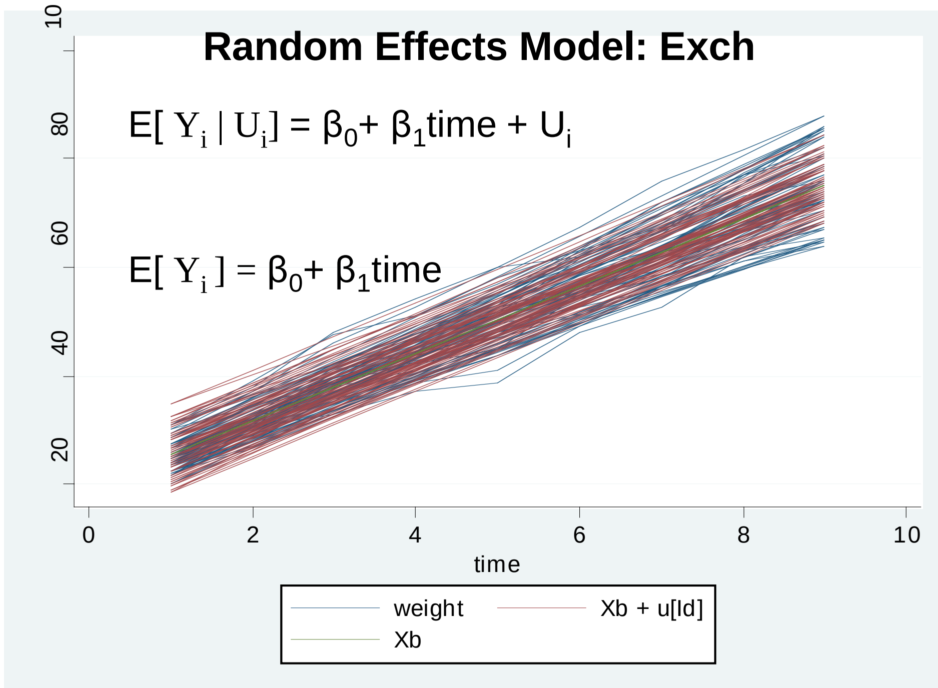
GEE-fit Marginal Model with AR1 Correlation structure

Pigs – RE model: AR(1)

```
xtregar weight time

RE GLS regression with AR(1) disturbances       Number of obs      =       432
Group variable (i): Id                          Number of groups   =        48

R-sq:  within  = 0.9851                         Obs per group: min =         9
       between = 0.0000                                        avg =       9.0
       overall = 0.9305                                        max =         9

                                                Wald chi2(2)       =  12688.55
corr(u_i, Xb)      = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   6.257651   .0555527   112.64   0.000      6.14877    6.366533
       _cons |   19.00945   .6281622    30.26   0.000     17.77827    20.24062
-------------+----------------------------------------------------------------
      rho_ar |  .73091237   (estimated autocorrelation coefficient)
     sigma_u |   3.583343
     sigma_e |  1.5590851
     rho_fov |  .84082696   (fraction of variance due to u_i)
       theta |  .60838037
------------------------------------------------------------------------------
```

GEE-fit Marginal Model with AR1 Correlation structure

Random Effects Model: Exch



$$E[\, Y_i \mid U_i] = \beta_0 + \beta_1 \text{time} + U_i$$

$$E[\, Y_i \,] = \beta_0 + \beta_1 \text{time}$$

Important Points

- Modelling the correlation in longitudinal data is important to be able to obtain correct inferences on regression coefficients β
- There are correspondences between random effect and marginal models in the linear case because the interpretation of the regression coefficients is the same as that in standard linear regression