Best Linear Unbiased Prediction - BLUP
- A TOOL FOR GENETIC EVALUATION

To maximize selection efficiency we want to rank animals based on a

selection criterion/index/EBV

which should be **accurate**

unbiased

Maximize Accuracy by

- **including as much information as possible**
  - **all possible relatives**
  - **correlated traits**

- **using proper index weights**

➡ **this is what selection index does! (=BLP)**

But what about unbiasedness?

Unbiased EBV's is a matter of fair comparisons

Possible problems with fairness:

- **Some animals produce on better herds (better pastures) than others**
- **Animals are measured at different ages**

- **The contemporaries of different animals may have different genetic mean**
- **Some sires have better mates**
- **There is culling and selection**

Correction for fixed effects

Problem 1:

- **Some animals produce on better herds (better pastures) than others**

<u>Solution 1:</u>

- **Phenotypic observations are taken as deviations of a mean (e.g. herd mean)**

Problem 2:

- **Some animals are measured at an older age**

<u>Solution 2:</u>

Phenotypic observations are corrected for the mean of the appropriate age

Genetic level confounded with herds

Problem 3:

The contemporaries of some animals may have higher genetic mean than of others

Example

progeny means from 4 sires in 2 herds

sire	*1*	*2*	*3*	*4*	5: link sire
herd 1	325	275	-	-	*325*
herd2	-	-	325	275	*375*

Genetic level confounded with herds

Problem 3:

The contemporaries of some animals may have higher genetic mean than of others

Example

progeny means from 4 sires in 2 herds

sire	*1*	*2*	*3*	*4*	5: link sire
herd 1	**325**	**275**	**-**	**-**	*325*
herd2	**-**	**-**	**325**	**275**	*375*

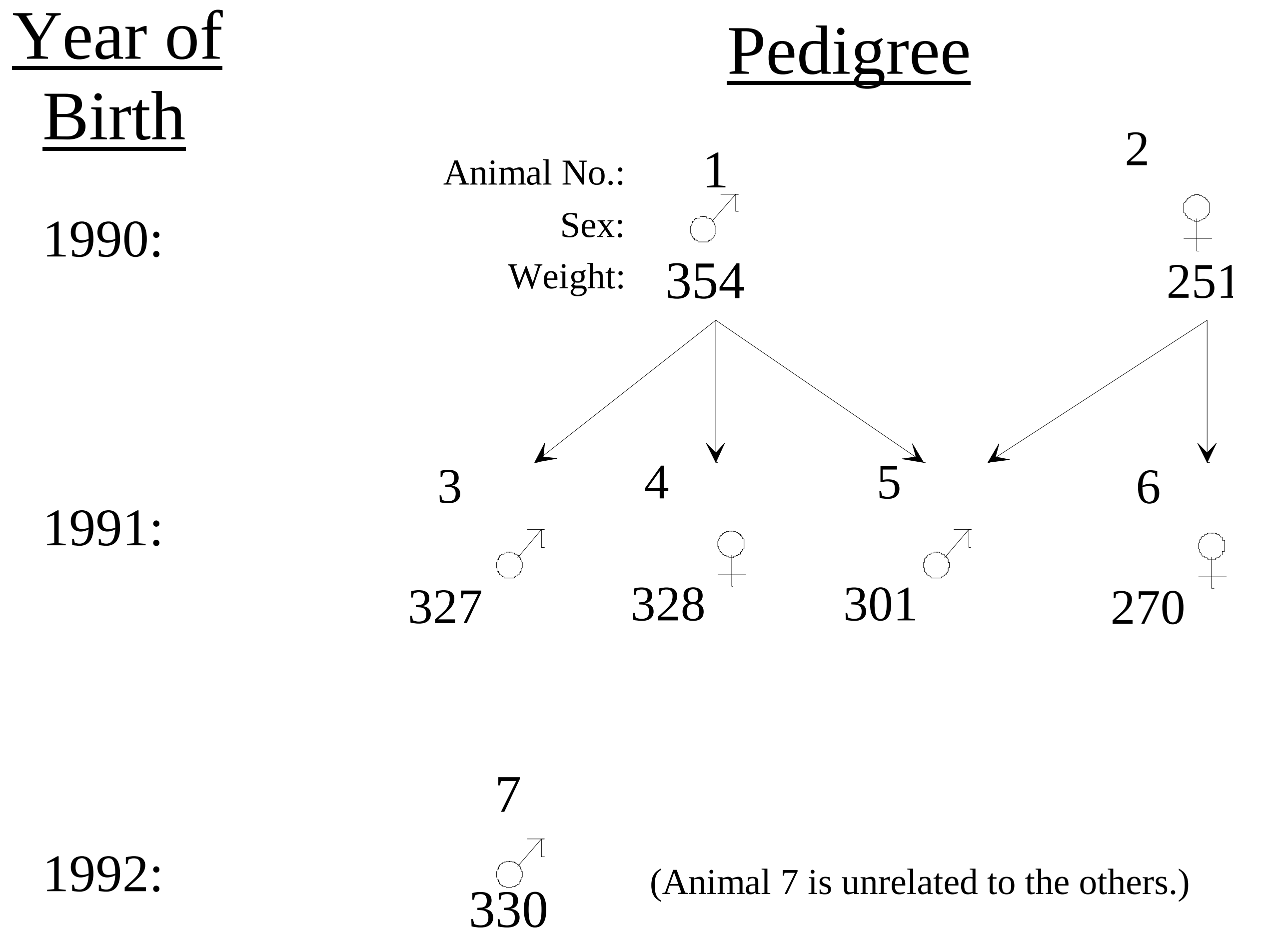
Conclusion

Need links between herds (reference sires)

Need a simultaneous evaluation of all herd and sire effects

The power of linear models

example:



$$Y = X b + e$$

$$\begin{pmatrix} 354 \\ 251 \\ 327 \\ 328 \\ 301 \\ 270 \\ 330 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_{mean} \\ b_{1990} \\ b_{1991} \\ b_{1992} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{pmatrix}$$

X is dependent

X'X can not be inverted

solutions:

$$\hat{b} = X'X^{-}X'Y$$

Can only estimate 3 parameters from 3 means

$$X'X = \begin{pmatrix} 7 & 2 & 4 & 1 \\ 2 & 2 & 0 & 0 \\ 4 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \text{ and } X'y = \begin{pmatrix} 2161 \\ 605 \\ 1226 \\ 330 \end{pmatrix}$$

Need restriction to solution

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

e.g. put effect of 1992 to zero

$$\hat{b} = (X'X)^{-1} \quad X'Y \quad = \quad \text{solutions}$$

$$\hat{b} = \begin{pmatrix} 7 & 2 & 4 \\ 2 & 2 & 0 \\ 4 & 0 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 2161 \\ 605 \\ 1226 \end{pmatrix} = \begin{pmatrix} 330 \\ -27.5 \\ -23.5 \end{pmatrix}$$

There are more solutions possible

General mean zero	**First year zero**	**Last year zero**	**Sum of years to zero**
μ = 0	**μ = 302.5**	**μ = 330**	**μ = 313**
1990 = 302.5	**1990 = 0**	**1990 = -27.5**	**1990 = -10.5**
1991 = 306.5	**1991 = +4**	**1991 = -23.5**	**1991 = -6.5**
1992 = 330	**1992 = +27.5**	**1992 = 0**	**1992 = 17**

→ estimable functions are unchanged

- expected value of an observation
- difference between years

fitting mean and year effect

$$\hat{b} = \begin{pmatrix} 330 \\ -27.5 \\ -23.5 \end{pmatrix} \rightarrow 4.0$$

the mean of 1992
the effect of year 1990 (relative to 1992)
the effect of year 1991 (relative to 1992)

fitting mean, year effect and sex

X $\hat{b}$ meaning

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 285.7 \\ -5.3 \\ -1.3 \\ 44.3 \end{pmatrix} \rightarrow 4.0$$

the mean of females in 1992
the effect of year 1990 (relative to 1992)
the effect of year 1991 (relative to 1992)
the effect of males (relative to females)

Note: year 1992 appears not so good after all!

Conclusion

- Linear models are a powerful, and relatively simple way to correct for different fixed effects in unbalanced designs

- Will use same principle to correct breeding values for different fixed effects