Repeated Records

LRS

CGIL

July-Aug 2012

Repeated Records

1. Cows, each lactation another measure of yield
2. Sows, each parity another observation of litter size
3. Race horses, each race another measure of speed
4. Dog trials, each event another measure of ability (tracking, retrieving)
5. Deer, each year another measure of antler yield
6. Animals, each disease occurrence another measure of immune response
7. No repeated measures on growth, happens only once

PE Effects

- Each time a trait on an individual is observed, assumed genetic correlation between each measure equals 1.
- Environment can affect animal performance, each measurement or observation on the same animal can have a constant PE effect associated with it.

Model

$$\mathbf{y} = \mathbf{Xb} + \begin{pmatrix} \mathbf{0} & \mathbf{Z} \end{pmatrix} \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_r \end{pmatrix} + \mathbf{Zp} + \mathbf{e},$$

where

$$\begin{aligned} \mathbf{b} &= \text{vector of fixed effects,} \\ \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_r \end{pmatrix} &= \begin{pmatrix} \text{animals without records} \\ \text{animals with records} \end{pmatrix}, \\ \mathbf{p} &= \text{vector of PE effects of length equal to } \mathbf{a}_r \text{ , and} \\ \mathbf{e} &= \text{vector of residual effects.} \end{aligned}$$

Prior Distributions

$$\begin{aligned} \mathbf{a} \mid \mathbf{A}, \sigma_a^2 &\sim N(\mathbf{0}, \mathbf{A}\sigma_a^2) \\ \mathbf{p} \mid \mathbf{I}, \sigma_p^2 &\sim N(\mathbf{0}, \mathbf{I}\sigma_p^2) \\ \mathbf{e} &\sim N(\mathbf{0}, \mathbf{I}\sigma_e^2) \\ \mathbf{G} &= \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_p^2 \end{pmatrix}. \end{aligned}$$

$$r = \frac{\sigma_a^2 + \sigma_p^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2}$$

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2}$$

Simulation

Animal	Sire	Dam	Year 1	Year 2	Year 3
7	1	2	√	√	√
8	3	4	√	√	
9	5	6	√		√
10	1	4		√	√
11	3	6			√
12	1	2		√	

Additive Genetic Values

Animal	Parent Ave.	*RND*	$(36 * b_i)^{.5}$	TBV
1	0.0	-2.5038	6.0	-15.0228
2	0.0	-.3490	6.0	-2.0940
3	0.0	-.2265	6.0	-1.3590
4	0.0	-.3938	6.0	-2.3628
5	0.0	1.4786	6.0	8.8716
6	0.0	2.3750	6.0	14.2500
7	-8.5584	-.8166	4.2426	-12.0229
8	-1.8609	1.0993	4.2426	2.8030
9	11.5608	1.5388	4.2426	18.0893
10	-8.6928	.0936	4.2426	-8.2957
11	6.4455	1.3805	4.2426	12.3024
12	-8.5584	-1.2754	4.2426	-13.9694

PE Values

Animal	TBV	PE
1	-15.02	2.97
2	-2.09	-9.04
3	-1.36	4.44
4	-2.36	-4.16
5	8.87	-5.68
6	14.25	6.85
7	-12.02	1.38
8	2.80	7.02
9	18.09	5.94
10	-8.30	-5.03
11	12.30	-1.06
12	-13.97	-2.69

Records

$$y_{ijk} \;=\; t_i + a_j + p_j + e_{ijk}$$

where t_i is a year effect. Let $t_1 = 53$, $t_2 = 59$, and $t_3 = 65$.

$$\sigma_p = 4, \quad \sigma_e = 6.9282$$

Residual values are generated for each observation as $RND * \sigma_e$. Add the pieces together and round to the nearest whole number.

Records

Animal 7

- In year 1,

$$y_{17k} = 53 + (-12.02) + (1.38) + (-3.36) \;=\; 39$$

- In year 2,

$$y_{27k} = 59 + (-12.02) + (1.38) + (2.64) \;=\; 51$$

- In year 3,

$$y_{37k} = 65 + (-12.02) + (1.38) + (7.64) \;=\; 62$$

Animal 12

- In year 2,

$$y_{2127k} = 59 + (-13.97) + (-2.69) + (3.66) \;=\; 46$$

Data

Animal	TBV	PE	Year 1 y_{1jk}	Year 2 y_{2jk}	Year 3 y_{3jk}
7	-12.02	1.38	39	51	62
8	2.80	7.02	48	72	
9	18.09	5.94	71		96
10	-8.30	-5.03		56	47
11	12.30	-1.06			86
12	-13.97	-2.69		46	

$$k_a = \sigma_e^2/\sigma_a^2 = 1.33333, \quad k_p = \sigma_e^2/\sigma_p^2 = 3$$

MME

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{Z} \\ \mathbf{0} & \mathbf{A}^{00}k_a & \mathbf{A}^{0r}k_a & \mathbf{0} \\ \mathbf{Z}'\mathbf{X} & \mathbf{A}^{r0}k_a & \mathbf{Z}'\mathbf{Z}+\mathbf{A}^{rr}k_a & \mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{0} & \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z}+\mathbf{I}k_p \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}}_0 \\ \hat{\mathbf{a}}_r \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{0} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}$$

For this example, order 21.

Parts of MME

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} 158 \\ 225 \\ 291 \end{pmatrix},$$

$$\mathbf{X}'\mathbf{Z} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\mathbf{Z}'\mathbf{Z} = diag(3 \;\; 2 \;\; 2 \;\; 2 \;\; 1 \;\; 1)$$

$$\mathbf{Z}'\mathbf{y} = \begin{pmatrix} 152 \\ 120 \\ 167 \\ 103 \\ 86 \\ 46 \end{pmatrix}$$

Solutions

$$\hat{t}_1 = 50.0858 \quad 53$$
$$\hat{t}_2 = 63.9612 \quad 59$$
$$\hat{t}_3 = 72.0582 \quad 65$$

Animal	TBV	PE	$\hat{\mathbf{a}}$	$\hat{\mathbf{p}}$
7	-12.02	1.38	-8.06	-1.66
8	2.80	7.02	1.01	0.79
9	18.09	5.94	11.14	4.51
10	-8.30	-5.03	-8.76	-3.10
11	12.30	-1.06	6.93	1.75
12	-13.97	-2.69	-8.78	-2.30

Variance Estimation

- Solve for each effect in MME, one at a time. Add GS "noise" to each.
- Residual variance, inverted Chi-square.
- Genetic variance, inverted Chi-square.
- PE variance, inverted Chi-square.
- Save samples, also for h^2 and r.

Animal Quadratic Form

$$
\begin{aligned}
\mathbf{a}'\mathbf{A}^{-1}\mathbf{a} &= \mathbf{a}'\mathbf{T}'^{-1}\mathbf{D}^{-2}\mathbf{T}\mathbf{a} \\
\mathbf{T}\mathbf{a}_i &= a_i - 0.5(a_{is} + a_{id}) \\
&= m_i, \mathrm{M}endelian \\
\mathbf{a}'\mathbf{A}^{-1}\mathbf{a} &= \sum_{i=1}^{q} m_i^2 d^{ii} \\
&= \sum (BY1) + \sum (BY2) + \ldots + \sum (BYp) \\
&= \sum (with) + \sum (without)
\end{aligned}
$$

Variance estimation may work better if only animals with records are used, not all animals. Try it out.

MPPA

Most Probable Producing Ability

MPPA = EBV + PE, can be used to rank animals before they make their next record, and to cull the unprofitable animals.

Selection

- Repeated records occur over time.
- Animals may be culled based on poor performance. Animals have different numbers of records.
- First records should always be present when analyzing later records, otherwise bias can occur.

Permanent?

- Repeated records model assumed for many years that the permanent environmental effect of an animal was constant and common to all records of that animal.
- Environmental effects can occur continuously during an animal's life, and therefore, should be considered as cumulative rather than permanent.
- Environmental effects on first records is due to everything that happened up until the first record was made.
- Environmental effects on second records is due to environmental effects up to first record PLUS environmental effects from first to second record.

Cumulative PE

Record 1 = G + PE 1 + TE 1

Record 2 = G + PE 1 + PE 2 + TE 2

Record 3 = G + PE 1 + PE 2 + PE 3 + TE 3

Model Differences

$$\mathbf{Z}_p\mathbf{p} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} p_{21} \\ p_{22} \\ p_{41} \\ p_{42} \\ p_{61} \\ p_{71} \\ p_{72} \\ p_{73} \end{pmatrix}$$

Model Differences

$$\mathbf{Z}'_p\mathbf{Z}_p = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Model Differences

The covariance matrix of PE effects for an animal with three records:

$$\begin{pmatrix} \sigma^2_{p1} & \sigma^2_{p1} & \sigma^2_{p1} \\ \sigma^2_{p1} & (\sigma^2_{p1} + \sigma^2_{p2}) & (\sigma^2_{p1} + \sigma^2_{p2}) \\ \sigma^2_{p1} & (\sigma^2_{p1} + \sigma^2_{p2}) & (\sigma^2_{p1} + \sigma^2_{p2} + \sigma^2_{p3}) \end{pmatrix}$$

Three PE effects for animal with 3 records.