

NAME:

Fall 2010 - Midterm Exam - October 20, 2010

Answers Provided

Points for each question are given in parenthesis. Total points = 110.

1. Two matrices are given below. Indicate the calculations for the following operations.

$$A = \begin{pmatrix} 4 & -1 \\ 1 & 0 \\ 2 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & -6 \\ -7 & 9 \end{pmatrix}.$$

(a) (2) Determinant of B  $5 \cdot 9 - (-7) \cdot (-6) = 3$

(b) (2) The product  $A'A = \begin{pmatrix} 21 & -10 \\ -10 & 10 \end{pmatrix}$

(c) (2) The sum  $A+B$  not conformable

(d) (2) The trace of A  $= 4 + 0 = a_{11} + a_{22} = 4$

2. Explain the purpose of the following R statements.

(a) (2) `library(MASS)` basic math library

(b) (2) `zlrsc=file.choose()` # find lrscrisps.R  
`source(zlrsc)` run lrscrisps.R for later use in R session

(c) (2) `Xhy = design(hy,0)` create a design matrix for hy

(d) (2) `fhy = factor(hy)` make hy a factor for use in lm or tapply

(e) (2) `ModelA = lm(y ~ fhy + fdiet + fage)`  
linear model function

(f) (2)  $AI = AINV(sid, did, bi)$  compute  $A^{-1}$  given a list of sires and dams, and  $b_i$  values.

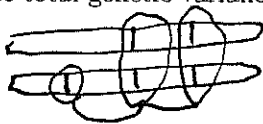
3. (4) The assumptions of the Infinitesimal Model are

- (a) there are an infinite number of loci with small and equal additive effects on genetic traits.
- (b) a randomly mating population of non-inbred individuals.

The Infinitesimal Model has been successful for over 70 years in animal agriculture. Do you believe the assumptions are valid? Explain

You could argue yes or no.  
There should be arguments for (a) and (b)

4. (4) Explain the meaning of an additive by dominance by dominance genetic variance. If two individual have an additive genetic relationship of  $\frac{1}{2}$  and a dominance genetic relationship of  $\frac{1}{4}$ , then how much of the additive by dominance by dominance variance contributes to the total genetic variance between these two relatives.



$$\frac{1}{2} \left( \frac{1}{4} \right)^2 = \frac{1}{32} \sigma_{12}^2$$

allele of 1 locus  
with pairs of alleles  
at two other loci

5. (2) Two stats can be used to compare models. They are the residual error variance and the multiple correlation coefficient, or  $R^2$ . What other information can be used to determine the fit of a model? (anova(), summary())

- F statistics on factors in the models
- residual plots ( $y - \hat{y}$ )
- likelihood values

6. (4) List two contributions of Dr. C. R. Henderson to animal breeding.

- $A^{-1}$
- $\tilde{B}LUP$ , MME
- 3 methods of variance component estimation
- Tabular method for making  $A$

7. (6) Below is a small partial section from a larger gametic relationship matrix. Fill in the missing numbers into the matrix.

	P1	P2	Q1	Q2	R1	R2	P S1	Q S2	P T1	R T2
S1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{4}$
S2	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$		1	$\frac{1}{4}$	$\frac{1}{4}$
T1	$\frac{1}{8}$	0	0	$\frac{1}{2}$	0	$\frac{3}{8}$			1	$\frac{3}{16}$
T2	$\frac{1}{4}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{8}$				1

6 numbers  
1 point each

error

8. (4) From the table in the previous problem what would be the additive and dominance relationships between animals R and S?

$$\text{add} = \frac{1}{2} \left( \frac{1}{2} + 0 + \frac{1}{8} + \frac{3}{8} \right) = \frac{1}{2} (1) = \frac{1}{2}$$

$$\text{dom} = \frac{1}{2} \cdot \frac{3}{8} + 0 \cdot \frac{1}{8} = \frac{3}{16}$$

9. (2) Why are phantom parent groups necessary in animal breeding programs?

To account for missing parents from different generations, different pathways of selection, different years of birth, and/or different countries.

10. (21) Given the following pedigrees, construct the additive relationship matrix for all animals in the table.

Animal	Sire	Dam	Year	Record
A				
B				
C	A		1	45
D	B	C	1	20
E	B	D	2	31
F	E	D	2	38

	A	B	C	D	E	F
A	1	0	1/2	1/4	1/8	3/16
B	0	1	0	1/2	3/4	5/8
C			1	1/2	1/4	3/8
D				1	3/4	1/2
E					1/4	1
F						3/8

symmetric

11. (2) Using the data from the previous question, construct a design matrix for an overall mean and year effects, as fixed effects, and a Z matrix for the random animal additive genetic effects.

$$\begin{aligned}
 y &= \mu + \text{year} + \text{animal} + e \\
 \begin{pmatrix} 45 \\ 20 \\ 31 \\ 38 \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_A \\ a_B \\ a_C \\ a_D \\ a_E \\ a_F \end{pmatrix} + e
 \end{aligned}$$

12. (6) From the inbreeding coefficients in the previous question, compute the  $b_i$  values for animals A to F.

( $b_i = 0.5 - 0.25(F_{sire} + F_{dam})$  for both parents known,

$b_i = 0.75 - 0.25(F_{parent})$  for only one parent known,

$b_i = 1$  for parents unknown.)

Animal	Sire	Dam	Inbreeding	$b_i$
A			0	1
B			0	$\frac{1}{4}$
C	A		0	$\frac{3}{4}$
D	B	C	0	$\frac{1}{2}$
E	B	D	$\frac{1}{4}$	$\frac{1}{2}$
F	E	D	$\frac{3}{8}$	$\frac{7}{16}$

from A on previous page

application of formulas

13. (21) Using the  $b_i$  values construct  $A^{-1}$  using Henderson's rules. Leave all of your results as fractions - no decimal numbers please.

	animal	sire	dam
animal	x	-.5x	-.5x
sire	-.5x	.25x	.25x
dam	-.5x	.25x	.25x

	A	B	C	D	E	F
A	$1\frac{1}{3}$		$-\frac{2}{3}$			
B		2	$\frac{1}{2}$	$-\frac{1}{2}$	-1	
C			$\frac{11}{6}$	-1		
D				$\frac{43}{14}$	$-\frac{3}{7}$	$-\frac{8}{7}$
E					$2\frac{4}{7}$	$-\frac{8}{7}$
F						$\frac{16}{7}$

symmetric

20 pts - just one or two mistakes

15 pts - more than 2, but method was okay

10 pts - seemed to have no idea what to do.

14. (4) Write out the Mixed Model Equations (in matrix notation) and explain the variables in the equations.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$G = \text{Var}(u)$   
 $R = \text{Var}(e)$   
 $y = \text{observations}$

$X = \text{design matrix for fixed effects}$ ,  $Z = \text{design matrix for random effects}$   
 $\hat{b}, \hat{u} = \text{solutions to the equations}$

15. (4) Below are the relative EBVs of 4 male alpacas for wool quality, the reliabilities of the EBVs, and the cost to buy that male. Which of these males, if any, would you purchase, assuming you are the owner of many female alpacas, and why?

Male	EBV	Reliability	Cost
John	122	.75	\$8,710
Paul	112	.95	\$6,160
George	114	.82	\$7,270
Ringo	124	.62	\$9,105

Everything accepted.

- high reliability Paul, George
- high EBV. John, Ringo, risk OK
- low relationships to females already in your herd.
- low cost Paul, George

16. (6) Skunks in the south end Guelph area have had DNA collected to determine parentage of most individuals. Volunteers have valiantly collected backblast distances (the number of meters a skunk can spray). For 2010, there are data on over 25 males and 34 females for backblast distance. Assume that parentage is known. Some individuals may have been measured more than once. Write a linear model to analyze the data and include all 3 parts of the model. Include the factors that you believe may be important for this trait. No skunks were harmed in this research, but a few people were sprayed.

$$\text{BBD} = \underbrace{\text{sex} + \text{age} + \text{wt} + \text{time of day} + \text{season}}_{\text{fixed}} + \underbrace{\text{skunk add. genetic}}_{\text{random}} + e$$

$$h^2 = .3, \text{ maybe}$$

- no escaped pet skunks
- skunks are equally nourished
- same technique to get skunks to spray
- people collecting data trained to use same protocol
- skunks have not sprayed within last 12 hours.
- no PE effects if skunks measured more than once.

5 = good model

4 = average model

3 = poor model, things missing from model