

## Supplementary File 3: *Tutorial for ASReml-R*

### Tutorial 1 (ASReml-R) - Estimating the heritability of birth weight

This tutorial will demonstrate how to run a univariate animal model using the software ASReml-R and example data files provided.

#### 1) Scenario:

In a population of gryphons there is strong positive selection on birth weight with heavier born individuals having, on average higher fitness. To find out whether increased birth weight will evolve in response to the selection, and if so how quickly, we want to estimate the heritability of birth weight.

#### 2) Data files:

Open `gryphonped.txt` and `gryphon.txt` in your text editor. The structure and contents of these files is fairly self-explanatory. The pedigree file `gryphonped.txt` contains three columns containing unique IDs that corresponding to each animal, its father, and its mother. Note that this is a multigenerational pedigree, with the earliest generation (for which parentage information is necessarily missing) at the beginning of the file. For later born individuals maternal identities are all known but paternity information is incomplete (a common situation in real world applications).

The phenotype data, as well as additional factors and covariates that we may wish to include in our model are contained in `gryphon.txt`. Columns correspond to individual identity (ANIMAL), maternal identity (MOTHER), year of birth (BYEAR), sex (SEX, where 1 is female and 2 is male), birth weight (BWT), and tarsus length (TARSUS). Each row of the data file contains a record for a different offspring individual. Note that all individuals included in the data file must be included as offspring in the pedigree file.

#### 3) Running the model:

To run an analysis copy the code below into the R console. Note that any content on a line after # is ignored and we have taken advantage of this to annotate the example below and highlight some key features. Note also that the file pathway (i.e. `C://...`) will have to be changed to direct R to the location of the tutorial data files on your computer.

```
#load the asreml library into R
library(asreml)

#read in the data file, telling R that "NA" is the symbol for missing values and #that
the first line of the file contains the column headers.
gryphon<-read.table("C://Documents and Settings//My Documents//gryphon.txt", header=T,
na.strings="NA");

#read in the pedigree file, telling R that "NA" is the symbol for missing values
#and that the first line of the file contains the column headers.
gryphonped<-read.table("C://Documents and Settings//My Documents//gryphonped.txt",
header=T, na.strings="NA");

#create the inverse of the relationship matrix from the pedigree file
ainv<-asreml.Ainverse(gryphonped)$ginv

#ensure that all variables are correctly assigned as numeric or factors.

gryphon$ANIMAL<-as.factor(gryphon$ANIMAL)
gryphon$MOTHER<-as.factor(gryphon$MOTHER)
gryphon$BYEAR<-as.factor(gryphon$BYEAR)
gryphon$SEX<-as.factor(gryphon$SEX)
gryphon$BWT<-as.numeric(gryphon$BWT)
gryphon$TARSUS<-as.numeric(gryphon$TARSUS)
```

```
#create modell, see numbered comments below for description of each line

modell<-asreml(fixed=BWT~ 1                                #1
              , random= ~ped(ANIMAL,var=T,init=1)        #2
              , data=gryphon                             #3
              ,ginverse=list(ANIMAL=ainv)                #4
              , na.method.X="omit", na.method.Y="omit")   #5

#1: BWT is the response variable and the only fixed effect is the mean(denoted as1)
#2: fit random effect of ANIMAL Va with an arbitrary starting value of 1
#3: use data file gryphon
#4: connect the individual in the data file to the pedigree
#5: omit any rows where the response or predictor variables are missing
```

This fits a simple univariate animal model with a single fixed effect (the mean) and a single random effect (the additive genetic effect). Run the following command to view the variance parameter estimates reported.

```
summary(modell)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	0.8868433	3.395215	0.6350655	5.346245	Positive
R!variance	1	3.828428	0.5185933	7.382332	Positive

We fitted a single random effect so have partitioned the phenotypic variance into two components. The ped(ANIMAL) variance component is  $V_A$  and is estimated as 3.40. Given that the ratio of  $V_A$  to its standard error (z.ratio) is considerably larger than 2 (i.e. the parameter estimate is more than 2 SEs from zero) this looks likely to be highly significant. The “R!variance” component refers to the residual variance  $V_R$ .

#### 4) Estimating heritability:

We can calculate the  $h^2$  of birth weight from the components above since  $h^2 = V_A/V_P = V_A/(V_A+V_R)$ . Thus according to this model,  $h^2 = 3.395215/(3.395215+3.828428) = 0.47$

#### 5) Adding fixed effects:

To add fixed effects to a univariate model simply modify the model statement. For example we might know (or suspect) that birth weight is a sexually dimorphic trait and therefore fit a model

```
modell2<-asreml(fixed=BWT~ 1+SEX
              , random= ~ped(ANIMAL,var=T,init=1)
              , data=gryphon
              ,ginverse=list(ANIMAL=ainv)
              , na.method.X="omit", na.method.Y="omit")
```

Now we can look at the fixed effects parameters and assess their significance with a conditional Wald F-test, using the code below.

```
summary(model2)$coef.fixed
wald.asreml(model2, ssType="conditional", denDF="numeric")
```

	solution	std error	z ratio
SEX_1	0	NA	NA
SEX_2	2.206995	0.1619955	13.6238
(Intercept)	6.058647	0.1718411	35.25727

  

\$Wald	Df	denDF	F.inc	F.con	Margin	Pr
(Intercept)	1	251	3491	3491		0
SEX	1	831	185.6	185.6	A	0

  

\$stratumVariances	df	Variance	ped(ANIMAL, var = T)
ped(ANIMAL, var = T)	0	752.2849	99.7151

The probability ('Pr') of zero in the Wald test above shows that SEX is a highly significant fixed effect, and from the fixed effects we can see that the average male (sex 2) is 2.2kg ( $\pm 0.16$ SE) heavier than the average female (sex 1).

However, when we look at the variance components in the model including SEX as a fixed effect, we see that they have changed slightly from the previous model.

```
summary(model2)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	1.041438	3.058928	0.5247986	5.828765	Positive
R!variance	1	2.937215	0.4161095	7.058754	Positive

In fact since SEX effects were previously contributing to the residual variance of the model our estimate of  $V_R$  (denoted "Variance" in the output) is now slightly lower than before. This has an important consequence for estimating heritability since if we calculate  $V_P$  as  $V_A + V_R$  then as we include fixed effects we will soak up more residual variance driving  $V_P$ . Assuming that  $V_A$  is more or less unaffected by the fixed effects fitted then as  $V_P$  goes down we expect our estimate of  $h^2$  will go up.

Here  $h^2$  has increased slightly from 0.47 to 0.51. Which is the better estimate? It depends on what your question is. The first is an estimate of the proportion of variance in birth weight explained by additive effects, the latter is an estimate of the proportion of variance in birth weight *after conditioning on sex* that is explained by additive effects.

## 6) Adding random effects:

This is done by simply modifying the model statement in the same way. For instance fitting

```
model3<-asreml(fixed=BWT~ 1+SEX
, random= ~ped(ANIMAL,var=T,init=1)+BYEAR
, data=gryphon
, ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

## Results in an additional variance component of birth year

```
summary(model3)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	1.1721829	2.7068911	0.4422088	6.121297	Positive
BYEAR	0.3837875	0.8862704	0.2695933	3.287435	Positive
R!variance	1	2.3092738	0.345101	6.691588	Positive

Here the variance in BWT explained by birth year is 0.886 and, based on the z.ratio appears to be significant. Thus we would conclude that year to year variation (e.g., in climate, resource abundance) contributes to  $V_P$ . Note that although  $V_A$  has changed somewhat, most of what is now partitioned as a birth year effect was previously partitioned as  $V_R$ . Thus what we have really done here is to partition environmental effects into those arising from year to year differences versus everything else, and we do not really expect much change in  $h^2$  (since now  $h^2 = V_A / (V_A + V_{BY} + V_R)$ ).

However, we get a somewhat different result if we also add a random effect of MOTHER to test for maternal effects:

```
model4<-asreml(fixed=BWT~ 1+SEX
, random= ~ped(ANIMAL,var=T,init=1)+BYEAR+MOTHER
, data=gryphon
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

Gives estimated variance components of

```
summary(model4)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	1.4110046	2.2985231	0.4962709	4.63159	Positive
BYEAR	0.5414535	0.8820265	0.2632475	3.35056	Positive
MOTHER	0.686595	1.1184617	0.2386229	4.687151	Positive
R!variance	1	1.6289976	0.3714178	4.38589	Positive

Here partitioning of significant maternal variance has resulted in a further decrease in  $V_R$  but also a decrease in  $V_A$ . The latter is because maternal effects of the sort we simulated (fixed differences between mothers) will have the consequence of increasing similarity among maternal siblings. Consequently they can look very much like additive genetic effects and if present, but unmodelled, represent a type of “common environment effect” that can - and will- cause upward bias in  $V_A$  and so  $h^2$ .

### 7) Testing significance of random effects:

A final point to note in this tutorial is that while the z ratio (COMP/SE) reported in the primary results file is a good indicator of likely statistical significance, the approximate standard errors are not recommended for formal hypothesis testing. A better approach is to use likelihood ratio tests.

For example, to test the significance of maternal effects here we could compare models with and without the inclusion of maternal identity as a random effect and compare the final log-likelihoods of these models.

```
model4$loglik
```

shows that the model including maternal identity has a log-likelihood of -1094.81, and

```
model3$loglik
```

shows that the model excluding maternal identity has a log-likelihood of -1113.77.

A test statistic equal to twice the absolute difference in these log-likelihoods is assumed to be distributed as Chi square with one degree of freedom. So in this case we would conclude that the maternal effects are highly significant since:

```
2*(model4$loglik-model3$loglik)
```

equals 37.91, and

```
1-pchisq(2*(model4$loglik-model3$loglik),1)
```

corresponds to  $P < 0.0001$ . We would therefore conclude that the additional of maternal identity as a random effect significantly improves the model, given an increase in log-likelihood of approximately 19.

## Tutorial 2 (ASReml-R) – A bivariate animal model

This tutorial will demonstrate how to run a multivariate animal model using the software ASReml-R and example data files provided.

### 1) Scenario:

Since natural selection rarely acts on single traits, to understand how birth weight might evolve in our population of gryphons, we may also want to think about possible covariance with other traits. If tarsus length at fledging is also under positive selection what implications does this have for birth weight and vice versa? If the two traits are positively genetically correlated then this will facilitate evolution of larger size (since response of one trait will induce a positively correlated response in the other). If there is negative genetic covariance then this could act as an evolutionary constraint.

Using multivariate models allows the estimation of parameters relating to each trait alone (i.e.  $V_A$ ,  $h^2$  etc), but also yields estimates of covariance components between traits. These include the (additive) genetic covariance  $COV_A$  which is often rescaled to give the genetic correlation  $r_G$ . However, covariance can also arise through other random effects (e.g. maternal covariance) and these sources can be explicitly modelled in a bivariate analysis.

### 2) Data files:

Pedigree and phenotypic data files are the same as those used in tutorial 1 (i.e. `gryphonped.txt` and `gryphon.txt` respectively).

### 3) Running the model:

For running multivariate analyses in ASReml-R the code is slightly more complex than for the univariate case. This is because ASReml-R allows us to make different assumptions about the way in which traits might be related and so we need to explicitly code a model of the (co)variance structure we want to fit. To run a bivariate analysis copy the following commands into the R console. The parts describing data structure and source are almost identical to the command file used in tutorial 1 but we have now removed extraneous annotation. We have also increased the default maximum number of iterations (`maxiter`) which can help to achieve convergence for more complicated models. Note that the file pathway (i.e. `C://...`) will have to be changed to direct R towards where you have stored the tutorial data files.

```
library(asreml)
gryphon<-read.table("C://Documents and Settings//My Documents//gryphon.txt", header=T,
na.strings="NA");
gryphonped<-read.table("C://Documents and Settings//My Documents//gryphonped.txt",
header=T, na.strings="NA");
ainv<-asreml.Ainverse(gryphonped)$ginv

gryphon$ANIMAL<-as.factor(gryphon$ANIMAL)
gryphon$MOTHER<-as.factor(gryphon$MOTHER)
gryphon$BYEAR<-as.factor(gryphon$BYEAR)
gryphon$SEX<-as.factor(gryphon$SEX)
gryphon$BWT<-as.numeric(gryphon$BWT)
gryphon$TARSUS<-as.numeric(gryphon$TARSUS)

#create modela, see numbered comments below for description of each line
modela<-asreml(cbind(BWT,TARSUS)~trait #1
,random=~us(trait,init=c(1,0.1,1)):ped(ANIMAL) #2
,rcov=~units(us(trait,init=c(0.1,0.1,0.1)) #3
,data=gryphon #4
,ginverse=list(ANIMAL=ainv) #5
,maxiter=20) #6
```

```
#1: Bivariate model of BWT & TARSUS, fitting the mean of each as fixed effects
#2: Random G structure defined as an unstructured covariance matrix (us). Initial
starting values of val=1 cov12=0.1 va2=1
#3: Residual (R) structure defined as an unstructured covar matrix, initial starting
values of val, cov12, va2.
#4: Use data file gryphon
#5: connect the individual in the data file to the pedigree
#6: maximum iterations =20
```

Note that the starting values supplied here are arbitrary. If the model is difficult to fit then it can be because the starting values are too far from the best estimates. One way around this is to run single trait models first to get good starting values for the variances (but you still have to “guess” starting values for the covariances).

Run this model and run the following code to look at the variance components.

```
summary(modela)$varcomp
```

Notice that there are now six (co)variance components reported in the table.

	gamma	component	std.error	z.ratio	constraint
trait:ped(ANIMAL)!trait.BWT:BWT	3.368449	3.368449	0.6348844	5.305611	Unconstrained
trait:ped(ANIMAL)!trait.TARSUS:BWT	2.460006	2.460006	1.0734265	2.291732	Unconstrained
trait:ped(ANIMAL)!trait.TARSUS:TARSUS	12.346304	12.346304	3.0749274	4.015153	Unconstrained
R!variance	1	1	NA	NA	Fixed
R!trait.BWT:BWT	3.849875	3.849875	0.5200013	7.403586	Unconstrained
R!trait.TARSUS:BWT	3.313128	3.313128	0.9129064	3.629209	Unconstrained
R!trait.TARSUS:TARSUS	17.646017	17.646017	2.6669515	6.61655	Unconstrained

The first three terms relate to the genetic matrix and, in order are  $V_{A,BWT}$ ,  $COV_A$ ,  $V_{A,TARSUS}$ . Below is a line where the gamma component equals 1 that is always outputted from a bivariate model and can be ignored. The final three terms relate to the residual matrix and correspond to  $V_{R,BWT}$ ,  $COV_R$ ,  $V_{R,TARSUS}$ . Based on our quick and dirty check (is  $z.ratio = Comp/SE > 2$ ) all components look to be statistically significant.

We can calculate the genetic correlation as  $COV_A / ((V_{A,BWT} * V_{R,TARSUS})^{0.5})$

Thus this model gives an estimate of  $r_G = +0.3815$ .

We can also calculate the heritability of birth weight as

$$h^2_{BWT} = V_{A,BWT} / V_{P,BWT} = V_{A,BWT} / (V_{A,BWT} + V_{R,BWT})$$

$$h^2_{BWT} = 3.368449 / (3.368449 + 3.849875) = 0.4667$$

And similarly the heritability of tarsus length can be calculated as 0.4116

#### 4) Adding fixed and random effects:

Fixed and random effects can be added just as for the univariate case. Given that our full model of BWT from tutorial 1 had SEX as a fixed effect as well as random effects of BIRTH YEAR and MOTHER we could specify a bivariate formulation of this using:

```
modelb<-asreml(cbind(BWT,TARSUS)~trait+trait:SEX
,random=~us(trait,init=c(1,0.1,1)):ped(ANIMAL)
+us(trait,init=c(1,0.1,1)):BYEAR
+us(trait,init=c(1,0.1,1)):MOTHER
```

```
,rcov= ~ units:us(trait,init=c(0.1,0.1,0.1))
,data=gryphon
,ginverse=list(ANIMAL=ainv)
,maxiter=20)
```

Note that we have specified a covariance structure for each random effect and an estimate of the effect of sex on both birth weight and tarsus length by interacting sex with trait in the fixed effect structure.

Note that there are will now be twelve (co)variance components reported after running the code

```
summary(modelb)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
trait:ped(ANIMAL)!trait.BWT:BWT	1.9893578	1.9893578	0.4410258	4.510751	Unconstrained
trait:ped(ANIMAL)!trait.TARSUS:BWT	3.3170792	3.3170792	0.9032658	3.6723179	Unconstrained
trait:ped(ANIMAL)!trait.TARSUS:TARSUS	10.2297428	10.2297428	2.8078803	3.643226	Unconstrained
trait:BYEAR!trait.BWT:BWT	0.9746359	0.9746359	0.2825691	3.4491944	Unconstrained
trait:BYEAR!trait.TARSUS:BWT	0.1624015	0.1624015	0.4184965	0.3880594	Unconstrained
trait:BYEAR!trait.TARSUS:TARSUS	3.7383626	3.7383626	1.2065555	3.098376	Unconstrained
trait:MOTHER!trait.BWT:BWT	1.1445198	1.1445198	0.2302168	4.9714873	Unconstrained
trait:MOTHER!trait.TARSUS:BWT	-1.5567348	-1.5567348	0.4051864	-3.842021	Unconstrained
trait:MOTHER!trait.TARSUS:TARSUS	4.8206057	4.8206057	1.3201482	3.6515642	Unconstrained
R!variance	1	1	NA	NA	Fixed
R!trait.BWT:BWT	1.8443077	1.8443077	0.3443166	5.3564305	Unconstrained
R!trait.TARSUS:BWT	4.0142617	4.0142617	0.7412593	5.4154621	Unconstrained
R!trait.TARSUS:TARSUS	12.4844258	12.4844258	2.2893707	5.4532129	Unconstrained

### 5) Significance testing:

Under the model above  $r_M$  is estimated at  $-0.663$  ( $-1.56/((4.82*1.14)^{0.5})$ ) and the Component : SE ratio ('z.ratio') associated with the corresponding covariance ( $COV_A$ ) is  $>2$  (in absolute terms). We might therefore infer that there is evidence of a strong negative correlation between the traits with respect to the mother and that while maternal identity explains variance in both traits those mothers that tend to produce heavier offspring actually tend to produce offspring with shorter tarsus lengths. To formally test whether  $COV_M$  was significantly different from zero, we might compare the log-likelihood for this model

```
modelb$loglik
```

which equals  $-2380.25$

to a model in which we specify that  $COV_M=0$ . Since this constraint reduces the number of parameters to be estimated by one, we can use a likelihood ratio test with one degree of freedom. To run the constrained model, modify the G structure definition for the mother random effect to read

```
diag(trait,init=c(1,1)):MOTHER
```

Here we have specified that the covariance matrix contains only diagonal (diag) elements (i.e. no covariance) and reduced the number of starting values by one to fit only values for the variance components. The full model specification is below

```
modelc<-asreml(cbind(BWT,TARSUS)~trait+trait:SEX
,random=~us(trait,init=c(1,0.1,1)):ped(ANIMAL)
+us(trait,init=c(1,0.1,1)):BYEAR
+diag(trait,init=c(1,1)):MOTHER
```

```
,rcov= ~ units:us(trait,init=c(0.1,0.1,0.1))
,data=gryphon
,ginverse=list(ANIMAL=ainv)
,maxiter=20)
```

## Using

```
modelc$loglik
```

We can see that the model log-likelihood is now -2386.05.

And comparing the models using a likelihood ratio test

```
2*(modelb$loglik-modelc$loglik)
```

gives a statistic of  $\chi^2_{IDF} = 11.60$

Which, when assessed against a Chi squared distribution with 1 degree of freedom

```
1-pchisq(2*(modelb$loglik-modelc$loglik),1)
```

corresponds to  $P=0.0007$  and we would therefore conclude that the maternal covariance is significantly greater than zero.

We could apply the same procedure to show that the residual (environmental) covariance and the genetic covariance estimates are significantly greater than zero (i.e., heavier individuals tend to have longer tarsus lengths). In contrast, we should find that the BYEAR covariance between the two traits is non-significant (since we did not actually simulate any between-trait covariance of BYEAR effects).

## Tutorial 3 (ASReml-R) – A repeated measures animal model

This tutorial will demonstrate how to run a univariate animal model for a trait with repeated observations using the software ASReml and example data files provided.

### 1) Scenario:

Since gryphons are iteroparous, multiple observations of reproductive traits are available for some individuals. Here we have repeated measures of lay date (measured in days after Jan 1) for individual females of varying age from 2 (age of maturation) up until age 6. Not all females lay every year so the number of observations per female is variable. We want to know how repeatable the trait is, and (assuming it is repeatable) how heritable it is.

### 2) Data files:

The pedigree file `gryphonped.txt` is that used in the preceding tutorials but we now use a new data file `gryphonRM.txt`. Open the latter in your text editor. Columns correspond to individual identity (ANIMAL), birth year (BYEAR), age in years (AGE), year of measurement (YEAR) and lay date (LAYDATE). Each row of the data file corresponds to a single phenotypic observation. Here data are sorted by identity and then age so that the repeated observations on individuals are readily apparent. However this is not a requirement for analysis - data could equally be sorted by some other variable (e.g., measurement year) or be in a random order.

**3) Estimating repeatability:** With repeated measures on individuals it is often of interest, prior to fitting a genetic model, to see how repeatable a trait is. We can estimate the repeatability of a trait as the proportion of phenotypic variance explained by individual identity by entering the code below into the R console. The parts describing data structure and source are similar to the command file used in tutorial 1 but we have now removed extraneous annotation. Note also that the file pathway (i.e. `C://...`) will have to be changed to direct R towards where you have stored the tutorial data files.

```
library(asreml)
gryphonRM<-read.table("C://Documents and Settings//My Documents//gryphonRM.txt",
header=T, na.strings="NA");

gryphonped<-read.table("C://Documents and Settings//My Documents//gryphonped.txt",
header=T, na.strings="NA");

ainv<-asreml.Ainverse(gryphonped)$ginv

gryphonRM$ANIMAL<-as.factor(gryphonRM$ANIMAL)
gryphonRM$BYEAR<-as.factor(gryphonRM$BYEAR)
gryphonRM$AGE<-as.factor(gryphonRM$AGE)
gryphonRM$YEAR<-as.factor(gryphonRM$YEAR)
gryphonRM$LAYDATE<-as.numeric(gryphonRM$LAYDATE)

modelv<-asreml(fixed=LAYDATE~ 1
, random= ~ide(ANIMAL,var=T,init=1)
, data=gryphonRM
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

Note that since we have associated the term ANIMAL with the pedigree file here, estimating the amount of variance explained by individual identity (rather than by additive effects) requires use of the code `ide( )` which effectively just means ignore the pedigree and treat ANIMAL as a normal random effect. The resulting partition of the phenotypic variance is seen using the code

```
summary(modelv)$varcomp
```

which outputs

	gamma	component	std.error	z.ratio	constraint
ide(ANIMAL, var = T)	0.5205728	11.08634	1.1794316	9.39973	Positive
R!variance	1	21.29643	0.8896197	23.9388	Positive

Between-individual variance is given by the ide(ANIMAL) component, while the residual component (R!Variance) therefore represents within-individual variance. Here then the repeatability of the trait can be determined by hand as 0.340 (i.e., as  $11.0863/(11.0863+21.2964)$ ). Given that we set up the simulation such that mean lay date changes with age (initially increasing to age 5 before a late life decline) we might ask what the repeatability of lay date is after conditioning on age effect. This would be done by adding age into the model as a fixed effect.

```
modelw<-asreml(fixed=LAYDATE~ 1 + AGE
, random= ~ide(ANIMAL,var=T,init=1)
, data=gryphonRM
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

Fitting the above gives

```
summary(modelw)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ide(ANIMAL, var = T)	0.750299	12.28986	1.156084	10.6306	Positive
R!variance	1	16.37995	0.6866362	23.85535	Positive

So that the repeatability of lay date, after accounting for age effects, is now estimated as 0.429 (i.e.,  $12.2898/(12.2898+16.3799)$ ). So, just as we saw when estimating  $h^2$  in tutorial 1, the inclusion of fixed effects will alter the estimated effect size if we determine total phenotypic variance as the sum of the variance components. Thus, proper interpretation is vital.

Here age is modelled as a 5 level factor (specified using the function `as.factor` at the beginning of the analysis). We could equally have fitted it as a continuous variable instead in which case, given the late life decline, we would probably also include a quadratic term.

#### 4) Partitioning additive and permanent environment effects:

Generally we expect that the repeatability will set the upper limit for heritability since, while additive genetic effects will cause among-individual variation, so will other types of effect. Non-additive contributions to fixed among-individual differences are normally referred to as “permanent environment effects”. If a trait has repeated measures then it is necessary to model permanent environment effects in an animal model to prevent upward bias in  $V_A$ . To illustrate this fit the animal model

```
modelx<-asreml(fixed=LAYDATE~ 1 + AGE
, random= ~ped(ANIMAL,var=T,init=1)
, data=gryphonRM
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

Note that, by using `ped()` instead of `ide()`, we are now fitting the additive effect

Variance components are almost unchanged:

```
summary(modelx)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	0.826448	13.91659	1.4437353	9.639293	Positive
R!variance	1	16.83904	0.7074338	23.802987	Positive

This suggests that all of the among-individual variance is –rightly or wrongly – being partitioned as  $V_A$  here. In fact here the partition is wrong since the simulation included both additive genetic effects and additional fixed heterogeneity that was not associated with the pedigree structure (i.e. permanent environment effects). An unbiased estimate of  $V_A$  is given by the model:

```
modely<-asreml(fixed=LAYDATE~ 1 + AGE
, random= ~ped(ANIMAL,var=T,init=1)+ide(ANIMAL,var=T,init=1)
, data=gryphonRM
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

Which yields variance components of:

```
summary(modely)$varcomp
```

	gamma	component	std.error	z.ratio	constraint
ped(ANIMAL, var = T)	0.2976828	4.876102	1.8087252	2.695878	Positive
ide(ANIMAL, var = T)	0.4518251	7.400982	1.7279724	4.283044	Positive
R!variance	1	16.380192	0.6866352	23.855741	Positive

The estimate of  $V_A$  is now much lower (reduced from 13.92 to 4.88) since the additive and permanent environment effects are being properly separated. We can estimate  $h^2$  and the repeatability from this model

$$h^2 = 4.876102 / (4.876102 + 7.400982 + 16.380192) = 0.17$$

$$\text{rep} = 7.400982 / (4.876102 + 7.400982 + 16.380192) = 0.43$$

### 5) Adding additional effects and testing significance:

Models of repeated measures can be extended to include other fixed or random effects. For example try including year of measurement (YEAR) and birth year (BYEAR) as random effects.

```
modelz<-asreml(fixed=LAYDATE~ 1 + AGE
, random=~ped(ANIMAL,var=T,init=1)+ide(ANIMAL,var=T,init=1) + YEAR + BYEAR
, data=gryphonRM
,ginverse=list(ANIMAL=ainv)
, na.method.X="omit", na.method.Y="omit")
```

This model will return additional variance components corresponding to variation in lay dates between years of measurement and between birth cohorts of females. The latter were simulated as should be apparent from the model fit (very low  $V_{\text{BYEAR}}$  and no change in log-likelihood if you compare to a reduced model with BYEAR excluded). Note that the z.ratio of the BYEAR component in modelz is high, although the magnitude of the component is very small and non-significant, emphasising the need to check significance using likelihood ratio tests as detailed above. However, YEAR effects were simulated as should be apparent from the COMP/SE (z.ratio) reported in `summary(modelz)$varcomp` and this could be formally confirmed by likelihood ratio test (see Tutorial 1).

YEAR effects could alternatively be included as fixed effects (try this!). Since we simulated large year of measurement effects this treatment will reduce  $V_R$  and increase the estimates of heritability and repeatability which must now be interpreted as proportions of phenotypic variance after conditioning on both age and year of measurement effects.