Supplementary File 2: *Tutorial for ASReml*

Running ASReml using Context

Having downloaded and installed the software our experience is that people are often confused about how to actually run it. In fact there are several options available here and to a large extent the way a user chooses to interact with the program is a matter of personal preference. The first option is to use a command-line interface to send jobs to ASReml. To do this you will first need to prepare input files using a text editor and, after running the job from the command line, you will then use the text editor to review the various output files produced. The second option is to use ASReml-W which is an inbuilt graphical user interface (GUI) that allows the user to edit command files, run them and save results as desired.

Although convenient and intuitive for users more accustomed to windows based interfaces we have found that the GUI can sometimes behave erratically (at least on Windows platforms). For this reason our normal preference is to use a third party text editor that communicates directly with ASReml. This allows the user to prepare and edit input files, run the job, and view the output files in a single software environment. The makers of ASReml suggest using a freeware text-editor called ConTEXT and we have found this to be a very easy and intuitive way to run analyses. Instructions for configuring ConTEXT appropriately come with the current distribution of ASReml. Below we have appended the text of the `Readme.txt` file which together with the file `ASReml.chl` can be found at `C:\Program files\ASReml2\Context` after a standard installation.

```
Context Notes on using ASReml v2 from Context
======= ====================================

ConText is a popular choice among ASReml users on the Windows platform. This is
a third party product and ASReml support does not include user support with
respect to the use of Context.

ConText is available from http://www.context.cx/. To install Context, download
and run this file. Then copy the file ASReml.chl to
C:\Program Files\Context\Highlighters.

A little bit of setting up is required to run ASReml conveniently. We assume
ASReml is installed in the standard location:

      C:\Program files\ASReml2\bin\asreml.exe.

To enable ASReml to run from Context under <F9>, type

<Options> <Environment Options> <Execute keys>
<Add> file extensions: as, asr
<F9>Execute "C:\Program files\ASReml2\bin\asreml.exe"
Start In    %p
Parameters  %f
Hint        ASReml

To access the User Guide by typing <F12>, find <ADOBE ACROBAT path>
by right clicking the ADOBE READER icon and selecting <Properties>. Then
continue with:

<F12>Execute <ADOBE ACROBAT path>
Start In    %p
Parameters  "C:\Program files\ASReml2\doc\UserGuide.pdf"
Hint        User Guide
```

```
In a similar way, you can setup Context to process a standard .pin file by
typing:

<Options> <Environment Options> <Execute keys>
<Add> file extensions: pin
<F9>Execute "C:\Program files\ASReml\bin\asreml.exe"
Start In    %p
Parameters -p     %f
Hint        ASReml PIN
```

Users who prefer to use the command line of ASReml-W GUI options should consult the user manual for more details. In case of problems technical support is also available from the publishers (VSN) and includes tutorials on installation and getting started (http://www.vsni.co.uk/software/asreml/user-area/tutorials/.

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

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

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 `gryphon.ped` and `gryphon.dat` in your text editor. The structure and contents of these files is fairly self-explanatory. The pedigree file `gryphon.ped` 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.dat`. 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 write the command file below using your text editor. We suggest saving as `gryphon1.as` (extension `.as` denoting ASReml command files). Note that we also recommend keeping all relevant files in the same folder on your computer when running ASReml for example `C:\Program files\ASReml2\tutorial1` or similar under Microsoft windows). With the input files stored here then when an analysis is performed ASReml will write the output files to the same directory.

A useful feature of the command file syntax is that any content on a line after # is ignored by ASReml. We have taken advantage of this to annotate the example below and highlight some key features. An ASReml command file has the following sections: project title, data description, name (and location) of the pedigree file, name (and location) of the data file, animal model specification. In more complex models there will also be a final section specifying the variance model (we will meet this in Tutorial 2). Run this command file (e.g., by pressing F9 if you are using Context and have configured it as described above).

```
# provide a title - failure to do this will
# prevent the programme from running properly

ASReml analysis of Gryphon birth weight

# next section describes contents of data file
# note that column headings are indented (use at least one space)
# !P associates a term with the pedigree file
# !A means this term is a factor and not continuous
```

```
   ANIMAL       !P
   MOTHER       !A
   BYEAR        !A
   SEX          !A
   BWT
   TARSUS

# then specify the pedigree and data files
# for the data file we tell ASReml to skip the first line since
# these are headers not data
# !FCON  !DDF 1  is asking for conditional tests of fixed effects
# Note that the pedigree and data file names below should not be indented
# see ASReml manual for more details

gryphon.ped    !skip 1
gryphon.dat    !skip 1   !FCON  !DDF 1

# Then we specify the model
# mu is the mean & any term after !r is treated as random

BWT ~ mu !r ANIMAL
```

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 command file and open the primary results file `gryphon1.asr` in your text editor to view the parameter estimates reported.

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ANIMAL                 1309   1309  0.886849       3.39541       5.35   0 P
 Variance                854    853   1.00000       3.82863       7.38   0 P

 Analysis of Variance              NumDF    DenDF_con F_inc    F_con M P_con
   7 mu                                1     224.8  2978.10  2978.10 . <.001
```

We fitted a single random effect and thus have partitioned the phenotypic variance into two variance components - an additive genetic (V_A) and a residual (V_R) component. The ANIMAL is V_A and estimated as 3.40. Given that the ratio of V_A to its standard error (listed under 'Comp/SE') is considerable larger than 2 (i.e. the parameter estimate is more than 2 SEs from zero) this looks likely to be highly significant. The information in the row described as "Variance" component refers to the residual variance V_R, which is estimate as 3.83. The lower section of the output ('Analysis of Variance') describes a conditional Wald F-test performed on the mean also indicates this is significantly greater than zero.

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, $h^2 = 3.39541/(3.39541+3.82863) = 0.47$

However, in practice it is easier to get ASReml to do this calculation for you and return the estimated heritability along with its standard error. These additional jobs are requested in a post-processing command file with the extension `.pin`. Write the following command file and save as `gryphon1.pin`

```
# Post-processing file to estimate heritability
# Primary results file contains two variances the first is VA & the second VR

F VP   1+2     #  Sum components to give VP (becomes component 3)
F VA   1       #  Writes out component 1 (VA) with SE
F VR   2       #  Writes out component 1 (VA) with SE
```

```
H h2   1 3     #  Ratio of Va to Vp (i.e. heritability) with SE
```

Here `F` requests a simple function of one or more variance components (addition, subtraction etc) while `H` requests a ratio. Running this file outputs the requested parameters (with SEs) to a further results file `gryphon1.pvc`:

```
ASReml 2.00a [01 Jul 2006]  ASReml analysis of Gryphon birth weight
          gryphon.pvc created 12 Jun 2009 17:24:39.453

   3 VP  1          7.224      0.3728
   4 VA  1          3.395      0.6350
   5 VR  2          3.829      0.5186
     h2           = ANIMAL     1/VP  1      3=          0.4700    0.0765
 Notice: The parameter estimates are followed by
          their approximate standard errors.
```

So in this model h^2 of birth weight equals 0.47 ± 0.08 SE.

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

```
BWT ~ mu SEX !r ANIMAL
```

Doing this obviously alters the output in `gryphon1.asr`, and the F-test shows that indeed SEX is highly significant

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ANIMAL                 1309   1309   1.04153       3.06044       5.84   0 P
 Variance                854    852   1.00000       2.93842       7.06   0 P

Analysis of Variance              NumDF    DenDF_con F_inc    F_con M P_con
   7 mu                               1     251.0  3491.44  3491.44 . <.001
   4 SEX                              1     831.0   185.60   185.60 A <.001
 Notice: The DenDF values are calculated ignoring fixed/boundary/singular
             variance parameters using numerical derivatives.

                     Estimate       Standard Error    T-value     T-prev
   4 SEX
     2                   2.20700       0.161997         13.62
   7 mu
                     1    6.05867       0.171825         35.26
```

However, it is also worth noting that the variance components have changed slightly. 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. Rerunning `gryphon1.pin` outputs the following to `gryphon1.pvc.`

```
ASReml 2.00a [01 Jul 2006]  ASReml analysis of Gryphon birth weight
          gryphon.pvc created 12 Jun 2009 17:47:26.906
```

```
   3 VP  1          5.999      0.3114
   4 VA  1          3.060      0.5244
   5 VR  2          2.938      0.4161
     h2           = ANIMAL     1/VP  1      3=          0.5102    0.0743
 Notice: The parameter estimates are followed by
          their approximate standard errors.
```

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

```
BWT ~ mu SEX !r ANIMAL BYEAR
```

Results in an additional variance component reported in `gryphon.asr`

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 BYEAR                   34     34  0.383782      0.886183       3.29   0 P
 ANIMAL                1309   1309   1.17215       2.70659       6.12   0 P
 Variance               854    852   1.00000       2.30908       6.69   0 P
```

Here the variance in BWT explained by birth year is 0.886 and, based on COMP/SE 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; this will test for any differences between offspring of different mothers, over and above estimated as being due to additive genetic effects:

```
BWT ~ mu SEX !r ANIMAL BYEAR MOTHER
```

Gives estimated variance components of

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 BYEAR                   34     34  0.541454      0.882030       3.35   0 P
 MOTHER                 429    429  0.686595       1.11847       4.69   0 P
 ANIMAL                1309   1309   1.41100       2.29853       4.63   0 P
 Variance               854    852   1.00000       1.62900       4.39   0 P
```

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.

Note that since we now have more variance components in our model output, using the post-processing options to estimate heritability requires some modification to `gryphon1.pin`. This should be rewritten as

```
# Post-processing file to estimate heritability
```

```
# Primary results file contains 4 variances the order given is
# 1-BYEAR, 2-MOTHER, 3-ANIMAL, 4-Variance

F VP   1+2+3+4     #  Sum all components to give VP (becomes component 5)
F VBYEAR    1
F VMOTHER   2
F VA        3
F VR        4
H h2   3 5     #  Ratio of Va to Vp is heritability
H m2   2 5     # corresponding ratio for maternal effect size
```

Which returns:

```
ASReml 2.00a [01 Jul 2006]  ASReml analysis of Gryphon birth weight
          gryphon.pvc created 12 Jun 2009 18:15:07.046

   5 VP  1          5.928      0.3802
   6 VBYEAR  1     0.8820      0.2632
   7 VMOTHER  2     1.118      0.2386
   8 VA  3          2.299      0.4963
   9 VR  4          1.629      0.3714
     h2           = ANIMAL     3/VP  1      5=          0.3877    0.0778
     m2           = MOTHER     2/VP  1      5=          0.1887    0.0380
 Notice: The parameter estimates are followed by
          their approximate standard errors.
```

7) Testing significance of random effects

A final point to note in this tutorial is that while the COMP/SE reported in the primary results file can be used as an 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 we could fit models with and without the inclusion of maternal identity as a random effect and compare the final log-likelihoods as reported in `gryphon1.asr`.

`BWT ~ mu SEX !r ANIMAL BYEAR MOTHER` gives a final LogL = -1094.81

`BWT ~ mu SEX !r ANIMAL BYEAR` gives a final LogL = -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:

χ^2_{1DF} = 2 x (-1094.81 – -1113.77) = 37.92 which corresponds to P<0.0001

Using this test then for α =0.05 we would 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) – A bivariate animal model

This tutorial will demonstrate how to run a multivariate animal model using the software ASReml 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, `gryphon.ped` and `gryphon.dat` respectively).

3) Running the model:
For running multivariate analyses in ASReml the code is slightly more complex than for the univariate case. This is because ASReml 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 in the main command file. To run a bivariate analysis write the following command file using your text editor. We suggest saving this file as `gryphon2.as` (extension `.as` denoting ASReml command files). The parts describing data structure and source are almost identical the command file used in tutorial 1 but we have now removed extraneous annotation. We have also increased the default maximum number of iterations (`!MAXIT`) which can help to achieve convergence for more complicated models and turned off the fixed effect testing.

```
ASReml analysis of Gryphon birth weight and tarsus length

   ANIMAL       !P
   MOTHER       !A
   BYEAR        !A
   SEX          !A
   BWT
   TARSUS

gryphon.ped     !skip 1
gryphon.dat    !skip 1   !MAXIT 20

# Then we specify the model
# Trait is the multivariate equivalent of mu
# So here we fit a mean for each trait (as fixed; "Trait" in the right-hand side of
# the equation), and an additive genetic effect for each trait as random (Trait.ID)

BWT TARSUS ~ Trait   !r  Trait.ANIMAL


#Then we have to specify our model of the variance structure
1 2 1         # Variance header: 1 residual structure with two dimensions,
              # 1 additional random effects
```

```
#Residual (R) structure defined as an unstructured covariance matrix

0
Trait 0 US  !S2== 1      # Starting values supplied as lower triangle
0.1                      # VR1,
0.1 0.1                  # COVR, VR2


#Random effect (G) structure defined as an unstructured covariance matrix

Trait.ANIMAL 2
Trait 0  US
1                        # Starting values supplied as lower triangle
0.1 1                    # VA1
ANIMAL                   # COVA, VA2
```

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 look at the primary results file `gryphon2.asr`. Notice that there are now six (co)variance components reported in the table, and these are also written out in matrix form below.

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 Residual         UnStructured  1  1   3.84987       3.84987       7.40   0 U
 Residual         UnStructured  2  1   3.31313       3.31313       3.63   0 U
 Residual         UnStructured  2  2   17.6460       17.6460       6.62   0 U
 Trait.ANIMAL       UnStructured 1 1   3.36845       3.36845       5.31   0 U
 Trait.ANIMAL       UnStructured 2 1   2.46001       2.46001       2.29   0 U
 Trait.ANIMAL       UnStructured 2 2   12.3463       12.3463       4.02   0 U
 Covariance/Variance/Correlation Matrix UnStructured Residual
   3.850      0.4020
   3.313       17.65
 Covariance/Variance/Correlation Matrix UnStructured Trait.ANIMAL
   3.368      0.3815
   2.460       12.35
```

The first three terms relate to the residual matrix and, since BWT is our first trait and TARSUS our second these are $V_{R.BWT}$, COV_R, $V_{R.TARSUS}$. These are followed by the corresponding genetic terms $V_{A.BWT}$, COV_A, $V_{A.TARSUS}$. Based on our quick and dirty check (is Comp/SE > 2) all components look to be statistically significant. Below the table the same components are written in matrix form with the covariance reported below the diagonal and the correlation reported above the diagonal. Thus this model gives an estimate of $r_G = +0.3815$.

Just as with the univariate model, we can extract (co)variance components and derived parameters (e.g., h^2, r_G) along with their approximate standard errors using the post-processing options. For instance to obtain estimates of h^2 for each trait along with the genetic variances we might save the following as `gryphon2.pin`

```
F VP_BWT   1+4     # becomes component 7
F VP_TAR   3+6     # becomes component 8
H h2_BWT   4 7     # Ratio of VA_BWT to VP_BWT is heritability
H h2_TAR   6 8     # Ratio of VA_TAR to VP_TAR is heritability
R rG       4 5 6   # R returns a correlation given components V1 COV12 V2
```

Running this will produce a results file called `gryphon2.pvc` containing

```
ASReml 2.00a [01 Jul 2006]  ASReml analysis of Gryphon birth weight
          gryphon.pvc created 14 Jun 2009 13:49:39.765
```

```
   7 VP_BWT  1      7.218      0.3723
   8 VP_TAR  3      29.99       1.700
     h2_BWT       = Trait.AN   4/VP_BWT     7=          0.4667    0.0767
     h2_TAR       = Trait.AN   6/VP_TAR     8=          0.4116    0.0931
     rG           = Trait.AN/SQR[Trait.AN*Trait.AN]=    0.3815    0.1300
 Notice: The parameter estimates are followed by
         their approximate standard errors.
```

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:

```
BWT TARSUS ~ Trait Trait.SEX !r  Trait.ANIMAL  Trait.BYEAR Trait.MOTHER


1 2 3        # note 3 here not 1 as previously since now 3 random effects

0
Trait 0 US  !S2== 1
0.1
0.1 0.1

Trait.ANIMAL 2    # Specifies G structure for additive effects
Trait 0  US
1
0.1 1
ANIMAL

Trait.BYEAR 2     # specifies G structure for BYEAR effects
Trait 0  US
1
0.1 1
BYEAR

Trait.MOTHER 2    # specifies G structure for MATERNAL effects
Trait 0  US
1
0.1 1
MOTHER
```

Note that there will now be twelve (co)variance components reported in `gryphon2.asr`.

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 Residual        UnStructured  1  1   1.84431       1.84431       5.36   0 U
 Residual        UnStructured  2  1   4.01427       4.01427       5.42   0 U
 Residual        UnStructured  2  2   12.4845       12.4845       5.45   0 U
 Trait.ANIMAL      UnStructured 1 1   1.98935       1.98935       4.51   0 U
 Trait.ANIMAL      UnStructured 2 1   3.31706       3.31706       3.67   0 U
 Trait.ANIMAL      UnStructured 2 2   10.2297       10.2297       3.64   0 U
 Trait.BYEAR       UnStructured 1 1  0.974638      0.974638       3.45   0 U
 Trait.BYEAR       UnStructured 2 1  0.162405      0.162405       0.39   0 U
 Trait.BYEAR       UnStructured 2 2   3.73836       3.73836       3.10   0 U
 Trait.MOTHER      UnStructured 1 1   1.14452       1.14452       4.97   0 U
 Trait.MOTHER      UnStructured 2 1  -1.55673      -1.55673      -3.84   0 U
 Trait.MOTHER      UnStructured 2 2   4.82061       4.82061       3.65   0 U
 Covariance/Variance/Correlation Matrix UnStructured Residual
   1.844      0.8366
   4.014       12.48
 Covariance/Variance/Correlation Matrix UnStructured Trait.ANIMAL
   1.989      0.7353
   3.317       10.23
 Covariance/Variance/Correlation Matrix UnStructured Trait.BYEAR
  0.9746      0.8508E-01
```

```
  0.1624       3.738
 Covariance/Variance/Correlation Matrix UnStructured Trait.MOTHER
   1.145     -0.6628
  -1.557       4.821
```

5) Significance testing

Under the model above the maternal covariance (r_M) is estimated at -0.6628 and the Comp/SE associated with the corresponding covariance (COV_A) is >2 in absolute terms (second last in list of terms estimated). We might therefore infer that there is evidence of a strong negative maternal correlation between the traits 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 (LogL=-2380.25) to one 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 additive effects to read

```
Trait.MOTHER 2      # Specifies G structure for MATERNAL effects
Trait 0  US  !GPZP  # constrains elements to be P(ositive), Z(ero), P(ositive)
1
0.1 1
MOTHER
```

The model log-likelihood will now be given as LogL=-2386.05. Comparing the models gives χ^2_{1DF} = 2 x (-2380.25 – -2386.05) = 11.6 which corresponds to P<0.001, and we would therefore conclude that the maternal covariance is significantly greater than zero.

We could apply the same procedure to show that both the residual (environmental) covariance and the additive genetic covariance are significantly positive. 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) – 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.

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 `gryphon.ped` is that used in the preceding tutorials but we now use a new data file `gryphonRM.dat`. 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 using the command file below (we suggest saving this as `gryphon3.as`)

```
ASReml analysis of Gryphon female laydate

 ANIMAL          !P
 BYEAR           !A
 AGE             !A
 YEAR            !A
 LAYDATE

gryphon.ped        !skip 1
gryphonRM.dat      !skip 1      !DDF 1 !FCON


LAYDATE ~ mu   !r   ide(ANIMAL)
```

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 in `gryphon3.asr` as

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ide(ANIMAL)            1309   1309  0.520573       11.0863       9.40   0 P
 Variance               1607   1606   1.00000       21.2964      23.94   0 P
```

The between-individual variance is given by the `ide(ANIMAL)` component, while the residual component (Variance) therefore represents within-individual variance. Here then the repeatability of the trait can be determined by hand (or using a post-processing file) as 0.340 (i.e., as 11.0863/(11.30863+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.

```
LAYDATE ~ mu  AGE !r  ide(ANIMAL)
```

Fitting the above gives

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ide(ANIMAL)            1309   1309  0.750299       12.2898      10.63   0 P
 Variance               1607   1602   1.00000       16.3799      23.86   0 P
```

So that the repeatability of laydate, 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 due to specifying `!A` after AGE in the command file. 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:

```
LAYDATE ~ mu AGE  !r   ANIMAL # note we are now fitting the additive effect
```

Variance components reported in `gryphon3.asr` are almost unchanged:

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ANIMAL                 1309   1309  0.826417       13.9171       9.64   0 P
 Variance               1607   1602   1.00000       16.8403      23.80   0 P
```

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:

```
LAYDATE ~ mu AGE  !r  ANIMAL  ide(ANIMAL) # include an addition non-additive effect
```

Which yields variance components of:

```
Source                Model  terms     Gamma     Component    Comp/SE   % C
 ANIMAL                 1309   1309  0.297683       4.87610       2.70   0 P
 ide(ANIMAL)            1309   1309  0.451825       7.40098       4.28   0 P
 Variance               1607   1602   1.00000       16.3802      23.86   0 P
```

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 could use the post-processing facility to estimate h^2 and the repeatability from this model using the following command file (save as `gryphon3.pin).`

```
F VP     1+2+3         # becomes component 4
F Vind   1+2           # VA+VPE = between individual variance (now component 5)
H rep    5 4           # Ratio of Vind to VP is repeatability
H h2     1 4           # Ratio of VA to VP is heritability
```

Parameters requested will be written to `gryphon3.pvc`

```
ASReml 2.00a [01 Jul 2006]  ASReml analysis of Gryphone female laydate
          gryphon2.pvc created 14 Jun 2009 16:11:13.406

   4 VP  1          28.66       1.261
   5 Vind  1        12.28       1.186
     rep          = Vind  1    5/VP  1      4=          0.4284    0.0274
     h2           = ANIMAL     1/VP  1      4=          0.1702    0.0607
 Notice: The parameter estimates are followed by
          their approximate standard errors.
```

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.

```
LAYDATE ~ mu  AGE  !r   ANIMAL  ide(ANIMAL)  YEAR  BYEAR
```

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 not simulated as should be apparent from the model fit (very low V_{BYEAR} and no change in log-likelihood if you compare to a reduced model with BYEAR excluded). However, YEAR effects were simulated as should be apparent from the COMP/SE reported in `gryphon3.asr` 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.