File: AnalysisData_TDTDGRPlines

```
% -----------------------------------------------------------------------

% loads 'TDTDGRPlines.mat': Matlab file of the original file
% "0.1.1. Survival time of Drosophila melanogaster across DGRP lines.cs"
%
% - ANCOVA model that separates slopes and intercepts (b1, b0) to estimate
%                     intercepts and slopes (needed for Var_CTmax and Var_z)
% - ML or REML estimation of variance-covariance parameters (Var_b0,
%                  Var_b1, Cov_b0_b1, Var_e)
%            Var_b0: Variance of intercepts of DGRP genotypes
%            Var_b1: Variance of genotype * testtemp
%            Var_e: Residual variance
% - Estimates variance parameters for CTmax and z using Taylor series
%                     estimator of a ratio (Wolter 2007, p. 240)
% - Jackknife estimates of variance-covariance parameters
% - Jackknife 95% CI
%
% - Written by M. Santos - March 2024 for
%         - MATLAB Version: 9.13.0.2502115 (R2022b) Update 8 -
%
% Needs 'Statistics and Machine Learning Toolbox'
%       - Statistics and Machine Learning Toolbox Version 12.4 (R2022b)-
%
% Reference: Kirk M. Wolter, Introduction to Variance Estimation, 2nd ed.
%            Springer (2007).


% -----------------------------------------------------------------------

% Put the data file and this code file in the same directory. In the
% MATLAB prompt type: AnalysisData_TDTDGRPlines
%
% The ouput are two tables, one for females and other for males (see below)
%
% NOTE: as the code is the linear mixed-effects model fits using ML.
%       To fit with REML replace in the commands:
%       lme = fitlme(tbl,'formula','FitMethod','ML')
%       'ML' by 'REML'

clear
load('TDTDGRPlines.mat'); % loads data file
F_ = TDTDGRPlines{:,4} == 'female';
M_ = TDTDGRPlines{:,4} == 'male';
Data_females = TDTDGRPlines(F_,:);
Data_males = TDTDGRPlines(M_,:);

output = cell(2,6);
output{1,1} = 'Females'; output{2,1} = 'Males';

% Analysis females
n_0Females = 10.2479840885613;
```

```matlab
yF = Data_females{:,14}; % survtime in log10 min.
testtempF = Data_females{:,5}; % covariate
genotypeF = Data_females{:,2}; % grouping factor

[hF,atabF,ctabF,statsF] = aoctool(testtempF,yF,genotypeF,0.05,...
    '','','','off','separate lines');  % ANCOVA model

% fit model
tbl = table(yF,testtempF, genotypeF, 'VariableNames', ...
    {'Survtime','Testtemp','Genotype'});
lme = fitlme(tbl,'Survtime ~ Testtemp + (1|Genotype) + (Testtemp|
Genotype)','FitMethod','ML');
[psiF,mseF] = covarianceParameters(lme);

% Covarianza components & H2 (Direct estimates)
Mean_b0F = mean(statsF.intercepts); Mean_b1F = mean(statsF.slopes);
Var_b0F = psiF{2,1}(1,1); Var_b1F = psiF{2,1}(2,2);
Cov_b0F_b1F = psiF{2,1}(1,2); Var_eF = mseF;
Var_CTmaxF = ((Mean_b0F)^2 / (Mean_b1F)^2) * (Var_b0F/(Mean_b0F)^2 - ...
    (2*Cov_b0F_b1F/(Mean_b0F * Mean_b1F)) + Var_b1F/(Mean_b1F)^2);
Var_zF = Var_b1F / (Mean_b1F)^4;
H2_CTmaxF = Var_CTmaxF / (Var_CTmaxF + Var_zF + Var_eF/n_0Females);
H2_zF = Var_zF / (Var_CTmaxF + Var_zF + Var_eF/n_0Females);

output{1,2}  = {'Var_b0'; 'Var_b1'; 'Cov_b0_b1'; 'Var_e'; ...
    'Var_CTmax'; 'Var_z'; 'H2_CTmax'; 'H2_z'};
output{1,3} = [Var_b0F; Var_b1F; Cov_b0F_b1F; Var_eF; ...
    Var_CTmaxF; Var_zF; H2_CTmaxF; H2_zF];  % estimates using all genotypes

% Delete-one-DRGP genotype Jackknife
lines = unique(Data_females{:,2}); % sorted list with the 20 DGRP genotypes
for ii = 1:numel(lines)
    I = find(Data_females{:,2} ~= lines(ii));
    J = find(statsF.gnames ~= lines(ii));
    yF = Data_females{I,14}; % survtime in log10 min.
    testtempF = Data_females{I,5}; % covariate
    genotypeF = Data_females{I,2}; % grouping factor
    tbl = table(yF,testtempF, genotypeF, 'VariableNames', ...
        {'Survtime','Testtemp','Genotype'});
    lme = fitlme(tbl,'Survtime ~ Testtemp + (1|Genotype) + (Testtemp|
Genotype)','FitMethod','ML');
    [psiF,mseF] = covarianceParameters(lme);

    Mean_b0F = mean(statsF.intercepts(J)); Mean_b1F = mean(statsF.slopes(J));
    Var_b0F = psiF{2,1}(1,1); Var_b1F = psiF{2,1}(2,2);
    Cov_b0F_b1F = psiF{2,1}(1,2); Var_eF = mseF;
    Var_CTmaxF = ((Mean_b0F)^2 / (Mean_b1F)^2) * (Var_b0F/(Mean_b0F)^2 - ...
        (2*Cov_b0F_b1F/(Mean_b0F * Mean_b1F)) + Var_b1F/(Mean_b1F)^2);
    Var_zF = Var_b1F / (Mean_b1F)^4;
    H2_CTmaxF = Var_CTmaxF / (Var_CTmaxF + Var_zF + Var_eF/n_0Females);
    H2_zF = Var_zF / (Var_CTmaxF + Var_zF + Var_eF/n_0Females);
    JdummyF(:,ii) = [Var_b0F; Var_b1F; Cov_b0F_b1F; Var_eF; Var_CTmaxF; ...
    Var_zF; H2_CTmaxF; H2_zF]; % keeps the delete-one estimates
end
```

```matlab
pseudo_valuesF = numel(lines).*repmat(output{1,3},1,numel(lines)) - ...
    (numel(lines)-1).*JdummyF;
Jack_estimatesF = mean(pseudo_valuesF,2);
output{1,4} = Jack_estimatesF; % Jackknife estimates
SE_F = sqrt(var(pseudo_valuesF,0,2)/numel(lines));
low_F = Jack_estimatesF - 2*SE_F;
up_F = Jack_estimatesF + 2*SE_F;
output{1,5} = low_F; % lower 95% CI
output{1,6} = up_F; % upper 95% CI

% Analysis males
n_0Males = 10.8054730372430;
yM = Data_males{:,14}; % survtime in log10 min.
testtempM = Data_males{:,5}; % covariate
genotypeM = Data_males{:,2}; % grouping factor

[hM,atabM,ctabM,statsM] = aoctool(testtempM,yM,genotypeM,0.05,...
    '','','','off','separate lines');  % ANCOVA model

% fit model
tbl = table(yM,testtempM, genotypeM, 'VariableNames', ...
    {'Survtime','Testtemp','Genotype'});
lme = fitlme(tbl,'Survtime ~ Testtemp + (1|Genotype) + (Testtemp|
Genotype)','FitMethod','ML');
[psiM,mseM] = covarianceParameters(lme);

% Covarianza components & H2 (Direct estimates)
Mean_b0M = mean(statsM.intercepts); Mean_b1M = mean(statsM.slopes);
Var_b0M = psiM{2,1}(1,1); Var_b1M = psiM{2,1}(2,2);
Cov_b0M_b1M = psiM{2,1}(1,2); Var_eM = mseM;
Var_CTmaxM = ((Mean_b0M)^2 / (Mean_b1M)^2) * (Var_b0M/(Mean_b0M)^2 - ...
    (2*Cov_b0M_b1M/(Mean_b0M * Mean_b1M)) + Var_b1M/(Mean_b1M)^2);
Var_zM = Var_b1M / (Mean_b1M)^4;
H2_CTmaxM = Var_CTmaxM / (Var_CTmaxM + Var_zM + Var_eM/n_0Males);
H2_zM = Var_zM / (Var_CTmaxM + Var_zM + Var_eM/n_0Males);

output{2,2}  = {'Var_b0'; 'Var_b1'; 'Cov_b0_b1'; 'Var_e'; ...
    'Var_CTmax'; 'Var_z'; 'H2_CTmax'; 'H2_z'};
output{2,3} = [Var_b0M; Var_b1M; Cov_b0M_b1M; Var_eM; ...
    Var_CTmaxM; Var_zM; H2_CTmaxM; H2_zM]; % estimates using all genotypes

% Delete-one-DRGP genotype Jackknife
lines = unique(Data_males{:,2}); % sorted list with the 20 DGRP genotypes
for ii = 1:numel(lines)
    I = find(Data_males{:,2} ~= lines(ii));
    J = find(statsM.gnames ~= lines(ii));
    yM = Data_males{I,14}; % survtime in log10 min.
    testtempM = Data_males{I,5}; % covariate
    genotypeM = Data_males{I,2}; % grouping factor
    tbl = table(yM,testtempM, genotypeM, 'VariableNames',...
        {'Survtime','Testtemp','Genotype'});
    lme = fitlme(tbl,'Survtime ~ Testtemp + (1|Genotype) + (Testtemp|
Genotype)','FitMethod','ML');
    [psiM,mseM] = covarianceParameters(lme);
```

```matlab
        Mean_b0M = mean(statsM.intercepts(J)); Mean_b1M = mean(statsM.slopes(J));
        Var_b0M = psiM{2,1}(1,1); Var_b1M = psiM{2,1}(2,2);
        Cov_b0M_b1M = psiM{2,1}(1,2); Var_eM = mseM;
        Var_CTmaxM = ((Mean_b0M)^2 / (Mean_b1M)^2) * (Var_b0M/(Mean_b0M)^2 - ...
            (2*Cov_b0M_b1M/(Mean_b0M * Mean_b1M)) + Var_b1M/(Mean_b1M)^2);
        Var_zM = Var_b1M / (Mean_b1M)^4;
        H2_CTmaxM = Var_CTmaxM / (Var_CTmaxM + Var_zM + Var_eM/n_0Males);
        H2_zM = Var_zM / (Var_CTmaxM + Var_zM + Var_eM/n_0Males);
        JdummyM(:,ii) = [Var_b0M; Var_b1M; Cov_b0M_b1M; Var_eM; Var_CTmaxM; ...
            Var_zM; H2_CTmaxM; H2_zM]; % keeps the delete-one estimates
end
pseudo_valuesM = numel(lines).*repmat(output{2,3},1,numel(lines)) - ...
    (numel(lines)-1).*JdummyM;
Jack_estimatesM = mean(pseudo_valuesM,2);
output{2,4} = Jack_estimatesM; % Jackknife estimates
SE_M = sqrt(var(pseudo_valuesM,0,2)/numel(lines));
low_M = Jack_estimatesM - 2*SE_M;
up_M = Jack_estimatesM + 2*SE_M;
output{2,5} = low_M; % lower 95% CI
output{2,6} = up_M; % upper 95% CI

clear a* c* C* d* F* g* H* h* I* i* J* l* m* M* p* S* s* t* u* V* y*
% keeps only the files 'Data_females', 'Data_males','TDTDGRPlines',
%        'output' in the Workspace

Component = output{1,2};
Females = table(output{1,3},output{1,4},output{1,5},...
    output{1,6},'VariableNames',["Estimate","Jackknife",...
    "Lower 95% CI","Upper 95% CI"],'RowNames',Component);

Males = table(output{2,3},output{2,4},output{2,5},...
    output{2,6},'VariableNames',["Estimate","Jackknife",...
    "Lower 95% CI","Upper 95% CI"],'RowNames',Component);
```

Output should be:

```matlab
%{
IF Linear mixed-effects model fit by ML
    lme = fitlme(tbl,'formula','FitMethod','ML')

Females =

  8×4 table
```

	Estimate	Jackknife	Lower 95% CI	Upper 95% CI
Var_b0	3.6272	3.7986	0.85169	6.7456
Var_b1	0.0026675	0.0028046	0.00061141	0.0049977
Cov_b0_b1	-0.098364	-0.10337	-0.18372	-0.023008
Var_e	0.025342	0.025379	0.017435	0.033323
Var_CTmax	0.48689	0.44496	-0.0090153	0.89894

```
        Var_z        0.14118      0.14958     0.038018       0.26114
        H2_CTmax     0.77218      0.75877       0.6636       0.85394
        H2_z          0.2239      0.23811      0.14571        0.3305


Males =

  8×4 table

                  Estimate      Jackknife    Lower 95% CI    Upper 95% CI
                  _____      _____    _____    _____


        Var_b0        7.5089       7.9066       1.9438         13.87
        Var_b1      0.0054776    0.0057813    0.0014722        0.01009
        Cov_b0_b1   -0.20281     -0.21405     -0.37438        -0.053725
        Var_e       0.022591     0.022629     0.016262         0.028995
        Var_CTmax     1.9845       1.9008      0.57054          3.231
        Var_z        0.65621      0.68576      0.17015          1.2014
        H2_CTmax      0.7509      0.73051      0.67047          0.79055
        H2_z          0.2483      0.26879      0.20839          0.3292

%   END ML fit
%

IF Linear mixed-effects model fit by REML
     lme = fitlme(tbl,'formula','FitMethod','REML')

Females =

  8×4 table

                  Estimate      Jackknife    Lower 95% CI    Upper 95% CI
                  _____      _____    _____    _____


        Var_b0        3.8551       3.7909      0.68863          6.8932
        Var_b1      0.0028334    0.0028034    0.00048933       0.0051175
        Cov_b0_b1   -0.10451     -0.10332     -0.18811        -0.018528
        Var_e       0.025341     0.025378     0.017435         0.033321
        Var_CTmax    0.51497      0.40129     -0.0099748        0.81256
        Var_z        0.14996      0.14956      0.031902         0.26722
        H2_CTmax      0.7716      0.73667      0.66995          0.80339
        H2_z         0.22469      0.25967      0.19372          0.32563


Males =

  8×4 table

                  Estimate      Jackknife    Lower 95% CI    Upper 95% CI
                  _____      _____    _____    _____


        Var_b0        7.9343       7.9066       1.6065         14.207
        Var_b1      0.0057872    0.0057814    0.0012311        0.010332
        Cov_b0_b1   -0.21428     -0.21405     -0.38336        -0.044751
```

```
    Var_e          0.022591       0.022629       0.016262       0.028995
    Var_CTmax       2.0967         1.8939         0.42215        3.3657
    Var_z           0.6933         0.68534        0.14073        1.2299
    H2_CTmax        0.75095        0.73192        0.67942        0.78442
    H2_z            0.24831        0.26739        0.21457        0.32021
%
%  END REML fit
%}
```

Published with MATLAB® R2022b