Supplementary file 2: annotated R-script for lme-AIC/logLikelihood analysis for non-demographic article-level metrics

Jeff C. Clements, Rémi M. Daigle, Halley E. Froehlich

July 25, 2017

If necessary, install the package `nlme`

```
install.packages("nlme")
```

Load the package:

```
library(nlme)
```

Upload data file for time since inception, number of papers per year

```
nondemographic <- read.csv("Table S2_Pred pubs.csv",stringsAsFactors = FALSE)
names(nondemographic)
```

```
##  [1] "Journal"
##  [2] "Predatory."
##  [3] "Article.."
##  [4] "Article.type"
##  [5] "X..authors"
##  [6] "Corresponding.author.nation.type"
##  [7] "Other.author.nation.types....other.authors."
##  [8] "Article.length....pgs."
##  [9] "Time.sub.accept..days."
## [10] "Time.accept.pub..days."
## [11] "X..of.references"
## [12] "X..figs...tables"
## [13] "Statistics.reported."
## [14] "Statistics.appropriately.reported."
## [15] "Downloadable.PDF."
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the number of authors per article.

```
nondemo1.lme<-lme(X..authors~Predatory.,random=~1|Journal,data=nondemographic)
anova(nondemo1.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1   228 428.8831  <.0001
## Predatory.      1    10  11.0192  0.0078
```

```
nondemo1.gls<-gls(X..authors~Predatory.,data=nondemographic)
anova(nondemo1.lme,nondemo1.gls)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## nondemo1.lme     1  4 994.8539 1008.743 -493.4270
## nondemo1.gls     2  3 999.1325 1009.549 -496.5663 1 vs 2 6.278572  0.0122
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the length of articles

```
nondemo2.lme<-lme(Article.length....pgs.~Predatory.,random=~1|Journal,data=nondemographic)
anova(nondemo2.lme)
```

```
##             numDF denDF   F-value p-value
## (Intercept)     1   228 208.01412  <.0001
## Predatory.      1    10  17.61223  0.0018
```

```
nondemo2.gls<-gls(Article.length....pgs.~Predatory.,data=nondemographic)
anova(nondemo2.lme,nondemo2.gls)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## nondemo2.lme     1  4 1478.859 1492.748 -735.4293
## nondemo2.gls     2  3 1493.649 1504.065 -743.8243 1 vs 2 16.78991  <.0001
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the number of references per article.

```
nondemo3.lme<-lme(X..of.references~Predatory.,random=~1|Journal,data=nondemographic)
anova(nondemo3.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1   228 923.2310  <.0001
## Predatory.      1    10  22.4203   8e-04
```

```
nondemo3.gls<-gls(X..of.references~Predatory.,data=nondemographic)
anova(nondemo3.lme,nondemo3.gls)
```

```
##              Model df      AIC      BIC    logLik   Test      L.Ratio
## nondemo3.lme     1  4 2162.908 2176.797 -1077.454
## nondemo3.gls     2  3 2160.908 2171.325 -1077.454 1 vs 2 2.155007e-07
##              p-value
## nondemo3.lme
## nondemo3.gls  0.9996
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the number of figures and tables per article.

```
nondemo4.lme<-lme(X..figs...tables~Predatory.,random=~1|Journal,data=nondemographic,na.action='na.omit')
anova(nondemo4.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1   207 487.0490  <.0001
## Predatory.      1    10   1.3058  0.2798
```

```
nondemo4.gls<-gls(X..figs...tables~Predatory.,data=nondemographic,na.action='na.omit')
anova(nondemo4.lme,nondemo4.gls)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## nondemo4.lme     1  4 1202.448 1215.967 -597.2238
## nondemo4.gls     2  3 1201.499 1211.639 -597.7495 1 vs 2 1.051348  0.3052
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the time from acceptance to publication

```
nondemo5.lme<-lme(Time.sub.accept..days.~Predatory.,random=~1|Journal,data=nondemographic,na.action='na
anova(nondemo5.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1   171 72.20690  <.0001
## Predatory.      1     7  8.29146  0.0237
```

```
nondemo5.gls<-gls(Time.sub.accept..days.~Predatory.,data=nondemographic,na.action='na.omit')
anova(nondemo5.lme,nondemo5.gls)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## nondemo5.lme     1  4 2053.091 2065.818 -1022.545
## nondemo5.gls     2  3 2069.731 2079.276 -1031.865 1 vs 2 18.64016  <.0001
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on the time from submission to acceptance

```
nondemo6.lme<-lme(Time.accept.pub..days.~Predatory.,random=~1|Journal,data=nondemographic,na.action='na
anova(nondemo6.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1   171 8.430415  0.0042
## Predatory.      1     7 0.025689  0.8772
```

```
nondemo6.gls<-gls(Time.accept.pub..days.~Predatory.,data=nondemographic,na.action='na.omit')
anova(nondemo6.lme,nondemo6.gls)
```

```
##              Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## nondemo6.lme     1  4 1726.690 1739.418 -859.3452
## nondemo6.gls     2  3 1809.035 1818.581 -901.5176 1 vs 2 84.34492  <.0001
```

Model df AIC BIC logLik Test L.Ratio p-value nondemo6.lme 1 4 1726.690 1739.418 -859.3452 nondemo6.gls 2 3 1809.035 1818.581 -901.5176 1 vs 2 84.34492 <.0001

Upload data file for demographic proportions

```
demographic <- read.csv("TableS3_proportions.csv",stringsAsFactors = FALSE)
names(demographic)
```

```
## [1] "Predatory."
## [2] "Journal"
## [3] "X.a..Prop.developing.prim.authors"
## [4] "X.b..Prop.developing.corr.authors"
## [5] "X.c..Prop.developing.other.authors"
## [6] "X.d..Prop.statistics.reported"
## [7] "X.e..Prop.stats.adequate"
```

Arcsine transform the proportion data

```
demographic$arcProp.statistics.reported <- asin(demographic$X.d..Prop.statistics.reported)
demographic$arcProp.stats.adequate <- asin(demographic$X.e..Prop.stats.adequate)
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on proportion of article reporting general statistical analyses

```
nondemo7.lme<-lme(arcProp.statistics.reported~Predatory.,random=~1|Journal,data=demographic)
anova(nondemo7.lme)
```

```
##             numDF denDF   F-value p-value
## (Intercept)     1    10 184.78480  <.0001
## Predatory.      1    10  33.87698   2e-04
```

```
nondemo7.gls<-gls(arcProp.statistics.reported~Predatory.,data=demographic)
anova(nondemo7.lme,nondemo7.gls)
```

```
##              Model df       AIC      BIC    logLik   Test      L.Ratio
## nondemo7.lme     1  4 11.344608 12.55495 -1.672304
```

```
## nondemo7.gls     2  3  9.344608 10.25236 -1.672304 1 vs 2 3.552714e-15
##              p-value
## nondemo7.lme
## nondemo7.gls       1
```

lme and AIC/logLikelihood for effect of journal [random] and journal type [fixed] on proportion of article reporting details of statistical analyses

```
nondemo8.lme<-lme(arcProp.stats.adequate~Predatory.,random=~1|Journal,data=demographic)
anova(nondemo8.lme)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1    10 52.65127  <.0001
## Predatory.      1    10 40.53909   1e-04
```

```
nondemo8.gls<-gls(arcProp.stats.adequate~Predatory.,data=demographic)
anova(nondemo8.lme, nondemo8.gls)
```

```
##              Model df      AIC      BIC     logLik   Test      L.Ratio
## nondemo8.lme     1  4 8.653626 9.863966 -0.3268129
## nondemo8.gls     2  3 6.653626 7.561381 -0.3268129 1 vs 2 1.776357e-15
##              p-value
## nondemo8.lme
## nondemo8.gls       1
```