

ON DRILLING FREQUENCY AND MANLY'S ALPHA: TOWARDS A NULL MODEL FOR PREDATOR PREFERENCE IN PALEOECOLOGY

JANSEN A. SMITH,^{1*} JOHN C. HANDLEY,² and GREGORY P. DIETL³

1. Cornell University, Department of Earth and Atmospheric Sciences, Ithaca, NY 14853,
jas933@cornell.edu

2. Paleontological Research Institution, Ithaca, NY, 14850, handley@priweb.org;

3. Cornell University, Department of Earth and Atmospheric Sciences, Ithaca, NY 14853,
Paleontological Research Institution, Ithaca, NY, 14850, gpd3@cornell.edu

ESTIMATION OF MANLY'S ALPHA IN R

This appendix provides the R code used to estimate Manly's alpha, as well as the code to create the types of figures presented in the main text.

```
# Data should be formatted as a simple data frame where each row is a taxon
# Be sure to remove any records with zero TaxonCounts
# TaxonCounts are counts of specimens per taxon
# PredationCounts are dill hole counts per taxon
# Below is a header to demonstrate proper data formatting
head(Counts)
```

Taxon	TaxonCounts	PredationCounts
<i>Abra palmeri</i>	1	0
<i>Acorylus rickettsi</i>	5	0
<i>Anomia peruviana</i>	1	0
<i>Argopecten</i> sp. 1	1	0
<i>Calliostoma palmeri</i>	1	0
<i>Calyptrea mamillaris</i>	14	0

```
# Compute Manly's alpha with posterior mean and credible intervals in JAGS (Just Another
# Gibbs Sampler – a program for analysis of Bayesian hierarchical models using Markov Chain
# Monte Carlo simulation that serves as a cross-platform engine for the BUGS language)
# Compute MLE (Maximum Likelihood Estimation) for comparison
# Specify model in BUGS (Bayesian inference Using Gibbs Sampling) language: store in a file
# Follows style of: Introduction to WinBUGS for Ecologists. 2010. Academic Press, Burlington
# Marc Kéry
# To perform these operations, load the following libraries:
```

```
library(lattice)
library(coda)
library(R2jags)
```

```
sink("ManlyAlpha.jags")
cat(""
```

```

model {

# The following Priors are necessary for the formulation:

alpha[1:m] ~ ddirch(rep(1/m,m))
# m is the number of taxa
# Output values should be between 0 and 1

# Likelihood: need the multinomial expression
PredationCounts[1:m] ~ dmulti(p,N)
p <- alpha[]*TaxonCounts/sum(alpha[]*TaxonCounts)
}
",fill = TRUE)
sink()

ManlyBayes = function(PredationData, file = "filename")
{
TaxonCounts = PredationData$TaxonCounts
#xtabs(~Species, PredationData)

PredationCounts = PredationData$PredationCounts
#xtabs(Drilled~Species, PredationData)

jags.data <- list(PredationCounts = PredationCounts, TaxonCounts = TaxonCounts, m =
length(PredationCounts), N = sum(PredationCounts))

# Initial values
inits <- function() list(alpha = rep(1/jags.data$m,jags.data$m))
params <- c("alpha") # Parameters monitored
ni <- 100000;nt <- 5;nb <- 50000;nc <- 11
alpha.out <- jags(jags.data, inits, params, "ManlyAlpha.jags", n.chains = nc, n.thin = nt, n.iter =
ni, n.burnin = nb, working.directory = getwd())
ProbGreater = numeric(jags.data$m)
for( ii in 1:jags.data$m){
ProbGreater[ii] = sum(alpha.out$BUGSoutput$sims.list$alpha[,ii]>1/jags.data$m
)/length(alpha.out$BUGSoutput$sims.list$alpha[,ii])
}
Result= data.frame( Taxon = PredationData$Taxon,
MLE =

# The following steps will generate data as reported in Table 1 of the main text
# Figure 2 in the main text is generated from these results
round((PredationCounts/TaxonCounts)/sum(PredationCounts/TaxonCounts),3),
Lo2.5 = round(alpha.out$BUGSoutput$summary[1:jags.data$m,"2.5%"],3),
Hi97.5 = round(alpha.out$BUGSoutput$summary[1:jags.data$m,"97.5%"],3),

```

```

    Bayes = round(alpha.out$BUGSoutput$mean$alpha,3),
    Neutral = round(rep(1/jags.data$m,jags.data$m),3),
    ProbGreaterNeutral= round(ProbGreater,3),
    ProbLessNeutral = round(1-ProbGreater,3) ,
    PredationCounts = PredationCounts,
    TaxonCounts = TaxonCounts)
write.csv( Result, file = sprintf("%s.csv",file))

}

# "The proportion of pseudo-BCD values which exceed the observed value is the p-value (exact
# significance level) for the test"
# From: PLEDGER, S., GEANGE, S., HOARE, J., and PEREZ-MATUS, A., 2007, Resource
# selection: Tests and estimation using null models: Victoria University of Wellington,
# Wellington, New Zealand.
# http://www.mcs.vuw.ac.nz/research/publications/reports/mscs/mscs07-04.pdf
# Pledger et al. suggest that Randomization (below) is a superior approach that MLE because it
# better incorporates cases where PredationCounts are 0 (that is, drilling frequency = 0)

RandomizationTest = function(Counts, file = "filename")
{
  TaxonCounts = Counts$TaxonCounts #xtabs(~Species, PredationData)
  PredationCounts = Counts$PredationCounts #xtabs(Drilled~Species, PredationData)
  # Make random draws under the NULL
  uProb= TaxonCounts/sum(TaxonCounts)
  Fsample = rmultinom(1000, size = sum(PredationCounts), uProb)

  #Bray Curtis Dissimilarity under the null
  uProb= TaxonCounts/sum(TaxonCounts) # the selection probabilities when no preference
  ExpectedPredationCountsNull = uProb*sum(PredationCounts)
  BCD = numeric(1000)
  for(ii in 1:1000){

    BCD[ii] = sum(abs(Fsample[,ii] - ExpectedPredationCountsNull))/(2*sum(PredationCounts))
  }

  # the BCD for the sample we have compared to the NULL
  ExpectedPredationCountsNull = uProb*sum(PredationCounts)
  BCDactual = sum(abs(PredationCounts -
    ExpectedPredationCountsNull))/(2*sum(PredationCounts))

  #calculate p-value
  pvalue = 1-min(range(which(sort(c(BCD, BCDactual)) == BCDactual))/length(c(BCD,
    BCDactual)))
  alpha = (PredationCounts/TaxonCounts)/sum(PredationCounts/TaxonCounts)

```

```

# Plot to PDF file
# This will generate a plot similar to Figure 1 in the main text
pdf(sprintf("%s.pdf",file) ,width=7,height=8,paper='special')
par(mfrow = c(2,1))
hist(c(BCD, BCDactual), main = sprintf("p-value: %2.3f",pvalue), xlab = "Bray-Curtis
Dissimilarity")
abline(v =BCDactual, col = "red")
par(mar = c(12,4,2,2))
plot(alpha, las = 2, xlab = "", type = "l")
abline(h = 1/length(alpha), lty = 2)

#back to default
par(mar=c(5.1, 4.1, 4.1, 2.1), mgp=c(3, 1, 0), las=0)
par(mfrow = c(1,1))
dev.off()
return(list(pvalue = pvalue,alpha = alpha, BCD = BCD, BCDactual = BCDactual) )
}

```