APPENDIX 1

Sequence of commands for spatial distribution of count data (Poisson Distribution). Random distribution of isopod individuals in a sandy beach on the Caribbean, Costa Rica.

```
#The data file is called beach.fauna in format .csv, and it is available at :
http://www.kerwa.ucr.ac.cr/handle/10669/29428

#Commands based in Krebs (1999) and Crawley (2007)
attach(beach_fauna)

#STEP 1) SAMPLE SIZE.
Isopoda<-Cirolana+Ancinus
n<-length(Isopoda)
n

#STEP 2) POPULATION PARAMETERS.
Mean<-mean(Isopoda)
Mean

Variance<-var(Isopoda)
Variance

#STEP 3) OBSERVED VALUES
#Table of observed values = first row: number of individuals, and second row: number of sampling units.
table(Isopoda)
#Observed values including the full range of number of individuals.
Isopoda1 <- factor(Isopoda, levels = c(0:max(Isopoda+1)))
table(Isopoda1)
#Frequency of sampling units by counts of individuals.
TableCounts<-data.frame(table(Isopoda1))
freq<- c(TableCounts$Freq)
freq

#STEP 4) INDEX OF DISPERSION:
#random pattern (ID = 1), uniform pattern (ID < 1), and clumped pattern (ID > 1).
ID<-var(Isopoda)/mean(Isopoda)
ID

#Index of dispersion test, based in a chi-squared.
CHI<-ID*(n-1)
CHI
#if chi-squared value is between tabular values the null hypothesis of random pattern (ID = 1) is not rejected.
qchisq(0.975,(n-1), lower.tail = FALSE) #tabular value of chi-squared 0.975
qchisq(0.025,(n-1), lower.tail = FALSE) #tabular value of chi-squared 0.025
#In this case, If the value of the index of dispersion is between the tabulated values the pattern is random (ID = 1)

#STEP 5) POISSON FIT FOR RANDOM PATTERN (ID = 1). AGGREGATION PROBABILITIES.
#The probability of obtaining X individuals in a sample follows a random pattern modeled by the Poisson distribution.
#density function to obtain the probability of found an x number of individuals in a sampling unit, with a parameter λ
#(λ = mean = variance):
```

$$p(x) \ = (\mathrm{e})^{-\lambda}\left(\frac{\lambda^{x}}{x!}\right)$$

```r
poisp<-dpois(0:max(Isopoda+1), Mean)

#STEP 6) EXPECTED FREQUENCIES FOR A THEORETICAL POISSON DISTRIBUTION
#The Poisson probabilities can be expressed as frequencies by multiplying each by the sample size (n)
poisfreq<-dpois(0:max(Isopoda+1), Mean)*n

#STEP 7) SUMMARY TABLE WITH OBSERVED AND EXPECTED VALUES
SummaryTable<-data.frame(x=TableCounts,poisfreq, poisp)
SummaryTable

#STEP 8) COMPARISON OF THE OBSERVED AND EXPECTED FREQUENCIES WITH A BARPLOT
par(mfrow=c(1,1), pch=22)
E<-(max(Isopoda+1)+1)*2
both<-numeric(E)
both[1:E %% 2 != 0]<-freq
both[1:E %% 2 == 0]<-poisfreq
labels<-character(E)
labels[1:E %% 2 == 0]<-as.character(0:max(Isopoda+1))
barplot(both,col=rep(c("black","white"),max(Isopoda)),names=labels,ylab="Frequency of samples",
        xlab="Counts")
legend(x = "topright",c("observed","expected"),fill=c("black","white"))

#STEP 9) GOODNESS OF FIT (G TEST):
#The null hypothesis is that observed values follow a Poisson distribution
#To follow the assumption of all expected values should be 1 or more, the last observations need be pooled (Krebs,
1999):

cs<-factor(0:max(Isopoda+1))
a<-max(which(SummaryTable$poisfreq>1))
a
levels(cs)[a:max(Isopoda+a)]<-"+"
levels(cs)

#expected values for G test:
ef<-as.vector(tapply(SummaryTable$poisfreq,cs,sum))

#observed values for G test:
of<-as.vector(tapply(SummaryTable$x.Freq,cs,sum))

#Table of contingence for G test:

p<-as.vector(tapply(SummaryTable$poisp,cs,sum))
ContigenceTable<-data.frame(of,ef,p)
ContigenceTable

#G test:
G<-2*sum((of)*log(of/ef))
```

```
G
#Degrees of freedom = Frequency classes (counts) - 2 statistics (total and mean) used to fit the Poisson distribution.

df<-sum(ef>0)-2
df
qchisq(0.95,df)
pchisq(G,df, lower.tail = FALSE)

#In this case p>0.05, and null hypothesis can not be rejected.
#Conclusion: The data follows a Poisson distribution.

#G test adjusted by Williams' correction for continuity for small samples (n < 200)(Krebs, 1999).
q<-1+((a+1)/(6*sum(of)*df))
q
Gadj<-G/q
Gadj
qchisq(0.95,df)
pchisq(Gadj,df, lower.tail = FALSE)
#In this case p > 0.05, and null hypothesis can not be rejected.
#Conclusion: The data follows a Poisson distribution.
```