Distribución espacial y temporal

Jeffrey A. Sibaja-Cordero

27-08-2022

*Texto inicial Federico Bolaños 1999*

*Editado por Federico Bolaños y Eduardo Chacón, septiembre 2009*

*Editado por Jeffrey Sibaja, agosto 2014 y agosto 2022*

*Editado por Andrea Vincent agosto 2022*

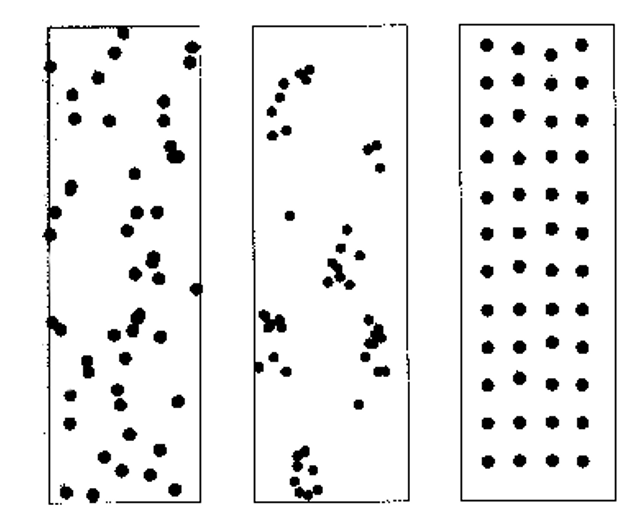
# Introducción

La distribución espacial y temporal estudia los patrones de cómo se presentan los organismos en un área o dentro de un período de tiempo, esto tiene un significado importante para comprender la biología de una especie, así como para el planeamiento de experimentos. Podemos ver si los individuos de una población usan el espacio donde viven en grupos, o son individuos distribuidos regularmente o al azar. En el tiempo se puede determinar si los individuos llegan de forma agrupada, uniforme o al azar a los recursos que utilizan. Cuando se ve la distribución en el espacio, este puede ser en dos o tres dimensiones, para ver el fenómeno en función de área o volumen.

Hay muchos métodos que se han propuesto para estudiar esto, aquí solo se desarrollan algunos, pero se pueden referir a los libros recomendados para ver otros. Podemos contar la frecuencia con que se da los eventos biológicos en dentro de unidades espaciales, intervalos de tiempo, u otro tipo de unidades experimentales, en todos los casos debe existir una uniformidad en el esfuerzo de muestreo, por ejemplo, cuadrantes del mismo tamaño, alícuotas de agua del mismo volumen, unidades de tiempo de la misma duración. Posteriormente, se emplean diferentes distribuciones estadísticas para describir cómo se comportan los datos.

Generalmente, para ver la dispersión se hacen parcelas en el espacio donde se cuentan los individuos de la población de una o varias especies, o en intervalos de tiempo para contar el número de individuos pasando, llegando o emigrando de un sitio. A lo largo de este documento, a estos intervalos de tiempo o a las parcelas, se le llamará unidades experimentales.

Como se mencionó antes los seres vivos siguen tres tipos de distribución (Fig. 1), aleatoria, también conocida como al azar, donde la probabilidad de encontrar un individuo es la misma en cualquier punto, ya sea en el espacio o en el tiempo, y la presencia de un individuo no afecta el que se encuentre otro, pueden o no estar cerca unos a otros. La distribución agrupada es donde la probabilidad de encontrar un individuo aumenta si hay otro, o sea se distribuyen en grupos o parches, este es el patrón más frecuente en la naturaleza. Finalmente, la distribución uniforme o regular, donde la probabilidad de encontrar un individuo donde hay otro es baja o nula y se repelen alejándose unos a otros.



*Fig. 1. Patrones de distribución de los seres vivos. Aleatorio, agrupado o de contagio y uniforme. Modificado de Krebs (1999).*

Si se supone que las tres parcelas de la figura 1 tienen el mismo número total de individuos, representados con puntos, y se sobreponen cuadrículas sobre estas áreas, es fácil demostrar que las tres parcelas tienen el mismo promedio de número de individuos. Si en total hay 50 individuos en cada una y se hacen 50 cuadrículas, en los tres patrones se tiene un promedio de 1 individuo por cuadrícula. Lo que varía entre estas tres parcelas es la variancia: en el patrón uniforme se tiende a una variancia de 0, en el agregado se tiende a la máxima variancia, y mayor que el promedio. En la dispersión aleatoria la variancia es intermedia y tiende a ser igual que el promedio. Es a partir de esta lógica que surge el único índice de dispersión en este documento. A continuación se les presenta dicho coeficiente de dispersión (CD):

donde s2 representa la variancia, y es el promedio, ambos de la muestra. Esta fórmula indica sila distribución es aleatoria, uniforme, o agrupada.

Un ejemplo para este cálculo en R se muestra para una población de isópodos muestreados a lo largo de una playa con 29 muestras de arena:

Isopoda<-c(0, 0, 0, 0, 2, 1, 2, 1, 2, 2, 1, 1, 0, 3, 0, 1, 0, 1, 2, 0, 0, 1, 0, 4, 4, 2, 1, 1, 1)  
Promedio<-mean(Isopoda)  
Variancia<-var(Isopoda)

Promedio

## [1] 1.137931

Variancia

## [1] 1.337438

Con el promedio y la variancia se calcula el CD

CD<-var(Isopoda)/mean(Isopoda)  
CD

## [1] 1.175325

Cuando se calcula el coeficiente de dispersión , el problema es cómo saber si el valor calculado es significativamente distinto de 1 (distribución aleatoria). El coeficiente multiplicado por sigue una distribución de chi-cuadrado con grados de libertad, donde representa el número de unidades de muestreo sobre las que se calcula el índice, entonces

Con esta prueba de chi-cuadrado se puede rechazar, o no, la hipótesis estadística nula de que el coeficiente es igual a 1. Hay que tener cuidado ya que las pruebas de que generalmente se aplican son las de una cola, puesto que casi siempre el interés es ver si el valor observado es diferente del esperado, no si es mayor o menor. Aquí el interés es una prueba de dos colas puesto que importa cuando el coeficiente es mayor que uno, pero también la relación inversa. Como es una prueba de dos colas, para aceptar la hipótesis nula con una probabilidad alfa de 0.05, el valor de obtenido debe estar entre los valores teóricos de la distribución y .

El cálculo para el ejemplo de los isópodos, es el tamaño de muestra

n<-length(Isopoda)  
n

## [1] 29

La prueba de para el CD observado

CHI<-CD\*(n-1)  
CHI

## [1] 32.90909

Valores tabulares de y

qchisq(0.975,(n-1), lower.tail = FALSE)

## [1] 15.30786

qchisq(0.025,(n-1), lower.tail = FALSE)

## [1] 44.46079

En este ejemplo, el valor de para el CD observado está entre los valores tabulares por lo que no se rechaza la de , por lo que el patrón de distribución de estos animales en la playa es aleatorio.

De esta sencilla manera, ya podemos conocer que patrón presenta nuestra especie, ya sea en el espacio o en el tiempo, depende del tipo de estudio. Es importante tener en cuenta que las causas para que la población de un organismo siga cierto tipo de distribución pueden ser debidas a algún factor ambiental, a factores intrínsecos de las especies como son su comportamiento, estrategias reproductivas, de dispersión o de reclutamiento, así como por la competencia entre los mismos individuos en la población y con poblaciones de otras especies.

La literatura con estudios previos sobre su ecología y notas de su historia natural pueden ayudar a explicar los patrones observados. Pero al mismo tiempo ante la carencia de información previa el obtener datos de su distribución espacial o temporal puede ser el punto de partida para conocer como la especie se desenvuelve en su medio. También, adicionalmente se puede evaluar si la distribución observada en la población de estudio sigue alguna teórica, lo que da mucho mejor posibilidad de predicción o bien para describir probabilísticamente como ocurren las cantidades de individuos. Estas distribuciones estadísticas son descritas por medio de la relación entre la variancia y el promedio, o coeficiente de dispersión.

## Ajuste a distribuciones de probabilidad

A continuación, se detallan tres distribuciones de probabilidad para datos discretos (también llamados datos de conteos), que más se utilizan para explicar los tres tipos básicos de dispersión en que se pueden encontrar los seres vivos.

Por ejemplo, en la distribución de Poisson, que explica apropiadamente la dispersión aleatoria, tiene la variancia igual al promedio ; en la distribución binomial negativa la variancia es mayor que el promedio , por lo que muchas veces es correcta para modelar poblaciones con patrón agregado. En la distribución binomial, la variancia es menor que el promedio , así puede aplicarse a poblaciones de patrón regular o uniforme. Existe la posibilidad de que la variancia sea 0, o sea que todos los valores de la población sean iguales, en estos casos hay una distribución constante .

## Distribución Poisson

La distribución Poisson se utiliza para ver si una serie de eventos ocurren de forma aleatoria. En el contexto de distribución espacial, esta se utiliza para resolver cuando los patrones de dispersión son al azar. Esta distribución tiene como parámetro , que es igual al promedio y a la variancia, , y el único parámetro que se estima. Entonces podemos estimar cual es la probabilidad de que en una unidad de muestreo se encuentren de 0 a infinito número de individuos, con

El problema entonces es tener una estimación de , lo que logramos con el promedio del número de individuos . Entonces la probabilidad de tener de 0 individuos en adelante estaría dada por

Con lo que llegamos a la siguiente ecuación, después de conocer la probabilidad de 0 individuos.

En R se pueden calcular todas estas probabilidades para los datos observados de los isópodos con el comando siguiente,

poisp<-dpois(0:max(Isopoda+1),Promedio)

Se debe calcular la probabilidad de encontrar un individuo más que el máximo. Las probabilidades se usan en la prueba de bondad de ajuste que se explica más adelante.

## Distribución Binomial

La distribución binomial se utiliza para ver la probabilidad de que uno de dos eventos pueda ocurrir. Un ejemplo es la probabilidad de que una moneda caiga con escudo o corona , probabilidades que se relacionan con , o con un ejemplo biológico la probabilidad de que un individuo de la población de un animal sea macho o hembra. Si tiramos 10 veces una moneda, y lo repetimos varias veces, no en todas las oportunidades saldrán cinco escudos, aún cuando la probabilidad de que salga cualquiera de ellos sea la misma .[[1]](#footnote-1) Entonces podemos estimar la probabilidad de que al tirar una moneda salgan 0 escudos y 10 coronas, o 1 escudo y 9 coronas, … o 10 escudos y 0 coronas. En este ejemplo entonces pueden salir un máximo de 10 escudos, lo que se denota como que son la cantidad de posibles eventos.

Escudos<-c(0,1,2,3,4,5,6,7,8,9)  
x<-0:max(Escudos+1)

Entonces la probabilidad de que obtengamos cualquier número de escudos, entre 0 y 10 , está dada por:

donde y

En el contexto de distribución espacial de organismos, esta distribución de probabilidad es de utilidad cuando la dispersión de los organismos es uniforme y por lo tanto el promedio es mayor que la variancia. Entonces se estima la probabilidad de que en cualquier punto de la unidad de muestreo se encuentre un individuo. Para esto se estima el número máximo de individuos que pueda aparecer en una unidad de muestreo redondeado sin decimales, lo que sería una estimación de , por

y se estiman de las ecuaciones anteriores. En R se puede calcular como:

k<-round(Promedio^2/(Promedio-Variancia))

Entonces

y

p<-Promedio/k

Esto da como parámetros estimados a y . Con el significado de , que es el número máximo de individuos que se espera en la unidad de muestreo. Hay que tener cuidado de no confundirlo con el número de unidades de muestreo realizadas.

Ahora se puede estimar la probabilidad de que se encuentren 0, 1, … individuos en una unidad de muestreo. Cabe la posibilidad de que en una unidad de muestreo se encuentren más individuos de lo que estima por lo que entonces sería ese número máximo, el que se usa como estimador de .

Con lo que llegamos a la siguiente ecuación, después de conocer la probabilidad de 0 individuos.

La fórmula para obtener las probabilidades en R es

binomialp<-dbinom(x, k, p)

## Distribución Binomial Negativa

La distribución Binomial Negativa se utiliza muchas veces para explicar patrones de dispersión de contagio o agrupados en el contexto de distribución espacial. Esta tiene como parámetros a y y donde el promedio y la variancia se definen como

y

Para un ejemplo se presenta los datos de distribución de un gusano poliqueto de 29 muestras de arena de una playa

Poliqueto<-c(0, 0, 0, 0, 0, 0, 0, 0, 0, 10, 4, 7, 2, 9, 6, 3, 9, 0, 1, 1, 2, 1, 0, 0, 0, 0, 0, 1, 1)  
Promedio<-mean(Poliqueto)  
Variancia<-var(Poliqueto)

Promedio

## [1] 1.965517

Variancia

## [1] 9.748768

Con el promedio y la variancia se calcula el CD

CD<-var(Poliqueto)/mean(Poliqueto)  
CD

## [1] 4.9599

El CD en este caso es mayor que 1, lo que indica patrón agregado, pero debe apoyarse en la prueba de según el tamaño de muestra

n<-length(Poliqueto)  
n

## [1] 29

La prueba de para el CD observado

CHI<-CD\*(n-1)  
CHI

## [1] 138.8772

Valores tabulares de y

qchisq(0.975,(n-1), lower.tail = FALSE)

## [1] 15.30786

qchisq(0.025,(n-1), lower.tail = FALSE)

## [1] 44.46079

Para esta especie de poliqueto, ciertamente el valor de excede los valores tabulares. Por lo que tenemos evidencia que la especie tiene un patrón agregado. Se puede entonces ver si los datos se ajustan a la distribución binomial negativa.

Para utilizar la binomial negativa, se tiene que estimar la constante . Se logra una aproximación inicial de con el uso de la siguiente ecuación, usando a y como estimadores de y ,

Para el ejemplo de los poliquetos el cálculo sería:

k=mean(Poliqueto)^2/(var(Poliqueto)-mean(Poliqueto))  
k

## [1] 0.4963553

Para obtener una mejor estimación de debe darse la siguiente igualdad,

donde, es el número de unidades de muestreo,

así como

con se denota la frecuencia de la clase .

Para cumplir la igualdad se debe hacer por prueba y error, modificando poco el valor de hasta llegar a un grado satisfactorio (similitud entre los valores), en la igualdad. Este procedimiento se debe usar solo cuando el número de unidades de muestreo es mayor de 20 y se pueden acomodar en una distribución de frecuencias, ver Krebs (1999) cuando no se cumple esta condición.

En R se puede utilizar para encontrar un valor mejor estimado de los comandos de Crawley (2007) por medio de máxima verosimilitud. Pero primero debe hacerse una tabla de frecuencia con la primera fila la cantidad de individuos y la segunda fila la cantidad de muestras con ese número de individuos.

Poliqueto1 <-factor(Poliqueto, levels = c(0:max(Poliqueto+1)))  
table(Poliqueto1)

## Poliqueto1  
## 0 1 2 3 4 5 6 7 8 9 10 11   
## 15 5 2 1 1 0 1 1 0 2 1 0

TablaConteos<-data.frame(table(Poliqueto1))  
freq<-c(TablaConteos$Freq)  
freq

## [1] 15 5 2 1 1 0 1 1 0 2 1 0

Los comandos de Crawley (2007) buscan cumplir la igualdad explicada anteriormente

x<-0:max(Poliqueto+1)  
kfit <-**function**(x)  
 {lhs<-numeric() *#lado izquierdo igualdad*  
 rhs<-numeric() *#lado derecho igualdad*   
 y <-0:(length(x) -1)  
 j<-0:(length(x)-2)  
 m <-sum(x \* y)/(sum(x))  
 s2 <-(sum(x \* y^2) -sum(x \* y)^2/sum(x))/(sum(x)-1)  
 k1 <-m^2/(s2 -m)  
 a<-numeric(length(x)-1)  
 **for**(i **in** 1:(length(x) -1)) a[i] <-sum(x [-c(1:i)])  
 i<-0  
 **for** (k **in** seq(k1/1.2,2\*k1,0.001)) {  
 i<-i+1  
 lhs[i] <-sum(x) \* log(1 + m/k) *#lado izquierdo igualdad*  
 rhs[i] <-sum(a/(k + j))} *#lado derecho igualdad*  
 k<-seq(k1/1.2,2\*k1,0.001)  
 plot(k, abs(lhs-rhs),xlab="k",ylab="Difference",type="l")  
 d<-min(abs(lhs-rhs)) *#mínima diferencia entre lados de la igualdad*  
 sdd<-which(abs(lhs-rhs)==d)  
 k[sdd]}  
K<-kfit(freq) *#uso de la tabla de frecuencias en la función para buscar k*

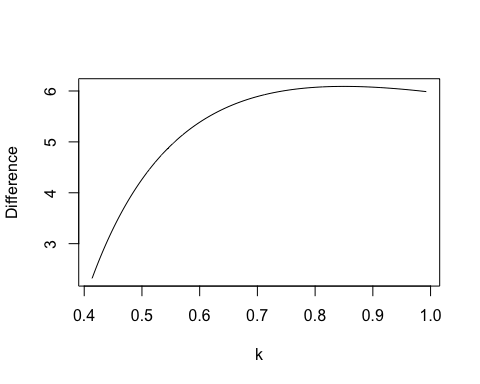


Fig. 2. Valor de K por estimador de máxima verosimilitud

K *#k mejor estimado*

## [1] 0.4136294

Con este valor mejor estimado de se pueden calcular las probabilidades de encontrar 0 individuos o de encontrar un individuo, y en este caso la probabilidad de encontrar un grupo de dos individuos, o un grupo de tres individuos:

Con lo que llegamos a la siguiente ecuación, después de conocer la probabilidad de 0 individuos.

En R se puede calcular las probabilidades con la función general para la distribución binomial negativa

la que involucra el cálculo de la función gama , con la mejor estimada y el promedio

Mean<-Promedio  
binneg<-**function**(x,Mean,K) (1+ Mean /K)^(-K)\*( Mean /( Mean +K))^ x \*gamma(K+ x)/(factorial(x)\*gamma(K))  
binnegprob<-sapply(x, **function**(i) binneg(i,Mean,K))

## Pruebas de Bondad y Ajuste

Una vez que se tiene la probabilidad de que se encuentren de 0 a infinito número de individuos en una unidad de muestreo, es importante que la última probabilidad debe ser la de ese número o más, excepto en la binomial donde representa el número mayor de individuos posible en una unidad. Estos valores de probabilidad se multiplican por el número de unidades de muestreo realizadas .

número esperado de muestras con 0 individuos

número esperado de muestras con 1 individuo

número esperado de muestras con dos individuos

número esperado de muestras con x de individuos

Para el ejemplo de los isópodos se usan las probabilidades de Poisson por el número de muestras

Promedio<-mean(Isopoda)  
poisfreq<-dpois(0:max(Isopoda+1),Promedio)\*n

Para el ejemplo de los poliquetos con patrón agregado se usan las probabilidades de la binomial negativa por el total de muestras

binnegfreq<-n\*(1+ Mean /K)^(-K)\*( Mean /( Mean +K))^ x \*gamma(K+ x)/(factorial(x)\*gamma(K))

y con la tabla de conteos observados en las muestras de la playa y estos valores esperados por Poisson o por la binomial negativa se hace una tabla de contingencia.

Tabla para los isópodos

Isopoda1 <-factor(Isopoda, levels = c(0:max(Isopoda+1)))  
TableCounts<-data.frame(table(Isopoda1))  
TablaIsopoda<-data.frame(x=TableCounts,poisfreq, poisp)  
TablaIsopoda

## x.Isopoda1 x.Freq poisfreq poisp  
## 1 0 10 9.2939606 0.32048140  
## 2 1 10 10.5758862 0.36468573  
## 3 2 6 6.0173146 0.20749361  
## 4 3 1 2.2824297 0.07870447  
## 5 4 2 0.6493119 0.02239007  
## 6 5 0 0.1477744 0.00509567

Tabla para los poliquetos

SummaryTable<-data.frame(x=TablaConteos,binnegfreq,binnegprob)  
SummaryTable

## x.Poliqueto1 x.Freq binnegfreq binnegprob  
## 1 0 15 14.0643062 0.484976077  
## 2 1 5 4.8060178 0.165724753  
## 3 2 2 2.8063808 0.096771752  
## 4 3 1 1.8653124 0.064321117  
## 5 4 1 1.3151146 0.045348780  
## 6 5 0 0.9590585 0.033070983  
## 7 6 1 0.7148880 0.024651311  
## 8 7 1 0.5411274 0.018659564  
## 9 8 0 0.4142820 0.014285585  
## 10 9 2 0.3199577 0.011033023  
## 11 10 1 0.2488315 0.008580395  
## 12 11 0 0.1946123 0.006710770

Esto genera una frecuencia esperada de unidades de muestreo según la distribución teórica utilizada para modelar los datos. Que se puede representar con un gráfico de barras con los siguientes comandos:

Para el caso de los isópodos

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="Cantidad de muestras",xlab="Número de individuos")  
legend(x = "topright",c("observados","esperados"),fill=c("black","white"))

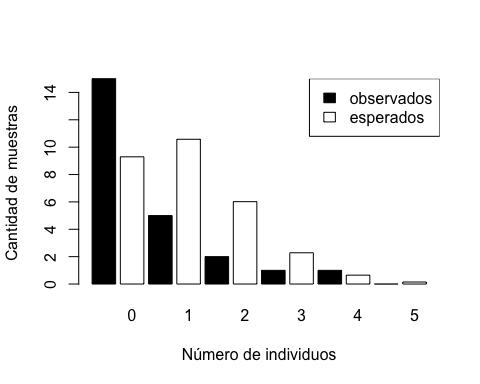


Fig. 3. Distribución de frecuencias de la cantidad de individuos de isópodos en muestras de arena.

Y para los poliquetos

E<-(max(Poliqueto+1)+1)\*2  
both<-numeric(E)  
both[1:E %% 2 != 0]<-freq  
both[1:E %% 2 == 0]<-binnegfreq  
labels<-character(E)  
labels[1:E %% 2 == 0]<-as.character(0:max(Poliqueto+1))  
par(mfrow=c(1,1), pch=22)  
Fig<-barplot(both,col=rep(c("black","white"),max(Poliqueto+1)),names=labels, ylab="Cantidad de muestras",xlab="Número de individuos")  
legend(x = "topright",c("observados","esperados"),fill=c("black","white"))

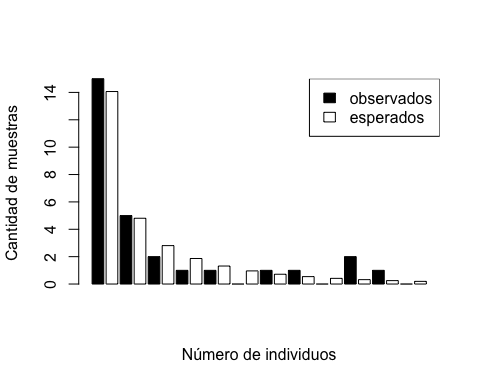


Fig. 4. Distribución de frecuencias de la cantidad de individuos de una especie de poliqueto en muestras de arena.

Es entonces cuando se puede realizar una prueba de chi-cuadrado para comprobar la hipótesis estadística nula de sí la distribución observada sigue la distribución teórica (esperados).

representa el número de categorías o clases de la distribución de frecuencias, por ejemplo, si son 0, 1, 2 y hasta 3 individuos posibles de encontrar en las unidades de muestreo sería un número de 4 categorías.

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

## [1] "0" "1" "2" "+"

Se calcula con base a esas 4 categorías de conteo de isópodos los valores observados y esperados en las unidades de muestreo

ef<-as.vector(tapply(TablaIsopoda$poisfreq,cs,sum)) *#esperados*  
of<-as.vector(tapply(TablaIsopoda$x.Freq,cs,sum)) *#observados*  
p<-as.vector(tapply(TablaIsopoda$poisp,cs,sum))   
ContigenceTable<-data.frame(of,ef,p)  
ContigenceTable

## of ef p  
## 1 10 9.293961 0.3204814  
## 2 10 10.575886 0.3646857  
## 3 6 6.017315 0.2074936  
## 4 3 3.079516 0.1061902

La fórmula de chi cuadradro es la siguiente

Para la población de isópodos se obtiene el siguiente valor de

pruebachi<-sum(((of-ef)^2)/ef)  
pruebachi

## [1] 0.08709767

La prueba sigue la distribución con grados de libertad cuando se usa para comparar los datos observados con la distribución Poisson.

df<-sum(ef>0)-2  
df

## [1] 2

Por lo que se obtiene la probabilidad con

pchisq(pruebachi, df, lower.tail = FALSE)

## [1] 0.9573858

Al obtenerse un no se rechaza la hipotesis nula que los datos de isópodos siguen una distribución aleatoria (CD=1) que puede ser modelada con la distribución de Poisson.

Se podría decir que la distribución se aplica en este tipo de situaciones con mucho éxito, sin embargo, tiene la limitación de que las frecuencias esperadas no deben ser pequeñas. Algunos autores sugieren que pequeño es cuando este valor es menor de 5, otros cuando es menor de 1, otros buscan un número intermedio y defienden que no debe ser menor de 3. Otra prueba de mucha utilidad y que no se ve tan afectada cuando las frecuencias esperadas son menores de 5 pero que no pueden ser menores de 1 es la , que sigue, aproximadamente, una distribución con los mismos grados de libertad expuestos anteriormente, la ecuación es

Cuando se tienen valores esperados menores de 5 en una prueba de , o menores de 1 en la de , deben agruparse las categorías cercanas para que no se violen estos requerimientos.

Al observar los datos en la tabla para los poliquetos hay varios esperados con menos de 1 unidad de muestreo, por lo que deben agruparse las últimas categorías de conteo de individuos (Krebs, 1999):

cs<-factor(0:max(Poliqueto+1))  
a<-min(which(SummaryTable$binnegfreq<1))  
levels(cs)[a:max(Poliqueto+a)]<-"+";levels(cs)

## [1] "0" "1" "2" "3" "4" "+"

En este ejemplo resulta en las categorías de 0 individuos, 1, 2, 3, 4 individuos y una última categoría de >4 individuos presentes en las muestras. Con estas categorías se puede sumar los valores de observados y esperados y presentar la tabla de contingencia

ef<-as.vector(tapply(SummaryTable$binnegfreq,cs,sum)) *#esperados*  
of<-as.vector(tapply(SummaryTable$x.Freq,cs,sum)) *#observados*  
p<-as.vector(tapply(SummaryTable$binnegprob,cs,sum))  
ContigenceTable<-data.frame(of,ef,p)  
ContigenceTable

## of ef p  
## 1 15 14.064306 0.48497608  
## 2 5 4.806018 0.16572475  
## 3 2 2.806381 0.09677175  
## 4 1 1.865312 0.06432112  
## 5 1 1.315115 0.04534878  
## 6 5 3.392757 0.11699163

Para un informe debe presentarse como cuadro, esto con formato limpio y los valores correctamente redondeados.

Se calcula los grados de libertad = categorías-total-promedio-k

df<-sum(ef>0)-3  
df

## [1] 3

En este caso se usa grados de libertad, a diferencia con Poisson que se usa grados de libertad. La razón de esta diferencia es que en todas debemos conocer el número total de unidades de muestreo para calcular las frecuencias esperadas, pero se estima también y en la Binomial, y y en la Binomial Negativa, mientras que en la de Poisson solo se estimó .

La prueba G para el ejemplo de los poliquetos:

G<-2\*sum((of)\*log(of/ef))  
G

## [1] 3.056244

La importancia de hacer la prueba de observados contra esperados es que si obtenemos un valor de no rechazamos la hipótesis nula que nuestros datos siguen la distribución teórica.

pchisq(G, df, lower.tail = FALSE)

## [1] 0.3830345

Este caso la especie de poliqueto sigue una distribución agregada que se puede explicar por la binomial negativa . Podemos usar esas probabilidades teóricas para explicar cómo se da la presencia de individuos en esa población. Por ejemplo, conoceríamos así que la probabilidad de encontrar grupos de cuatro individuos es en una unidad de muestreo dentro del sitio de estudio. Otro ejemplo en el caso de distribución temporal donde se puede conocer cuál es la probabilidad que 5 individuos de un insecto visiten una flor en un intervalo de tiempo definido.

Un último aspecto es que Krebs (1999) y Crawley (2007) recomiendan hacer una corrección en la prueba G cuando el número de muestras es menos de 200.

q<-1+((a+1)/(6\*sum(of)\*df))  
Gadj<-G/q  
Gadj

## [1] 3.015802

pchisq(Gadj, df, lower.tail = FALSE)

## [1] 0.3891952

Tras la corrección con el valor de y la probabilidad = 0.389, se mantiene la misma conclusión.

En otros casos puede que la prueba de chi-cuadrado o la prueba G den un valor de , por lo que se rechaza la hipótesis nula y se entiende que nuestros datos no se ajustan a esa distribución teórica, pero si tienen el patrón de distribución que probamos con el para el CD, por lo que si quisiéramos predecir o conocer probabilidades tendríamos que buscar otra forma de modelar la población.

## Referencias

Crawley, M. J. (2007). *The R book*. John Wiley & Sons.

Brower, J., Zar, J., y von Ende, C. (1997). *Field and Laboratory Methods for General Ecology* (4a ed.). WCB/McGraw Hill.

Fowler, J. y Cohen, L. (1990). *Practical Statistics for Field Biology*. John Wiley & Sons.

Krebs, C. J. (1999). *Ecological Methodology* (2a ed.). Addison-Welsey Educational Publishers, Inc.

Sibaja-Cordero, J. A. (2018). Spatial distribution of macrofauna within a sandy beach on the Caribbean coast of Costa Rica. *Revista de Biología Tropical*, *66*(S1), S176–S186.

Zar, J. H. (1996). *Biostatistical Analysis* (3a ed.). Prentice Hall, Inc.

1. Solo tirando una gran cantidad de veces la moneda (tendiendo al infinito), la probabilidad que sale escudo o corona se iguala a 0.5 por un mayor muestreo (ley de los números grandes). [↑](#footnote-ref-1)