**Práctica 2, Ecología I. Muestreo de poblaciones naturales: estimación del tamaño poblacional y de la distribución espacial.**

#============================================================================# Escrita por Bernardino González. Área de Ecología. Universidad de Vigo. 2015 #============================================================================

La estimación de parámetros poblacionales suele requerir la obtención de una muestra de la población estudiada. Dependiendo de diversas variables, como el objetivo del estudio, la disponibilidad de medios y tiempo, el tipo de organismo y tamaño del área de distribución de la población, se aplicarán diferentes diseños de muestreo (nº de unidades de muestreo, forma y extensión de las mismas, secuencia temporal, localización espacial, etc.).

En este caso, se realizará un muestreo de árboles en un área reducida, que tomaremos como la población a estudiar, con objeto de aprender a utilizar dos procedimientos básicos en el muestreo de poblaciones naturales: 1) Determinación del tamaño de muestra (Nº de unidades de muestreo) requerido para estimar la abundancia poblacional dentro de un margen de error preestablecido. 2) Estudio de la distribución espacial de una población. Además, se analizará el efecto de las condiciones de muestreo (número y tamaño de la las unidades de muestreo) sobre los resultados obtenidos a partir de esos procedimientos. El desarrollo de esta práctica sirve de introducción a los métodos de Análisis Espacial en Ecología.

**1) FUNDAMENTOS** (deberían leerse antes de la realización de la práctica)

**1.1-** Abundancia poblacional: estimación del tamaño de muestra adecuado

Requiere la realización de un muestreo preliminar al definitivo.

**1.1.1- Método del muestreo en dos etapas**

A partir del muestreo preliminar se calcula la varianza de los valores del número de individuos presentes en las distintas unidades de muestreo (habitualmente denominadas parcelas, cuadrículas, “plots”, “quadrats”, etc.). A continuación, basándose en la teoría estadística, se puede estimar el tamaño de muestra necesario para calcular el tamaño poblacional dentro de un margen de error (precisión) y con un nivel de confianza deseados (1-probabilidad de Error Tipo I), pudiéndose, además, incluir un grado de seguridad en el cumplimiento de este último (1-probabilidad de Error Tipo II). Por ejemplo: se busca estimar el número de árboles por metro cuadrado con un error máximo del 30%, con un grado de confianza del 95% (Error Tipo I, α=0.05) y con un grado de seguridad de este último del 90% (Error Tipo II, β=0.1).

Fórmula a aplicar (Zar, 2009) n = (S2 \* t2α(2), υ \* Fβ(1) (υ, n-1)) / d2

Siendo: n, el tamaño de muestra buscado

S2, la varianza de la muestra

t2α(2), υ , cuadrado del valor tabulado de la t de Student para un nivel de confianza 1-α (de dos colas) y un número de grados de libertad υ igual al tamaño de muestra tentativo menos 1.

Fβ(1) (υ, n-1) , el valor tabulado de la tabla de la F para un nivel de significación, de una cola, β, con grados de libertad υ (tamaño de muestra tentativo menos 1) y n-1 (tamaño de la muestra preliminar (n) menos 1). En el caso de no tener en cuenta el Error de Tipo II, este término se excluye de la fórmula anterior.

d, error admisible en la estimación del tamaño poblacional, corresponde a la mitad del intervalo de confianza de la estimación. El valor de “d” depende de la precisión deseada.

Con este método, la solución del tamaño de muestra necesario se obtiene de forma iterativa: a partir del número inicial de parcelas del muestreo preliminar se calcula un primer tamaño de muestra necesario; el valor obtenido se vuelve a incluir en la ecuación de cálculo, obteniéndose una segunda estimación del tamaño de muestra necesario; y se repite el proceso hasta que el tamaño resultante coincida con el que se ha introducido para el cálculo.

Finalmente se completan las muestras iniciales (número inicial del muestreo preliminar) hasta alcanzar el tamaño de muestra necesario (de ahí lo de muestreo en dos etapas).

**1.2-** Estudio del patrón de distribución

La distribución espacial de una población es el resultado de las características biológicas de la especie a la que pertenece y de las interacciones intrapoblacionales, con otras especies y con el medio físico-químico. Por tanto, el análisis de esta distribución puede ayudar a discernir cuáles son los factores ambientales más relevantes en su dinámica poblacional.

El estudio de distribución espacial a partir del muestreo poblacional por parcelas depende del tamaño y la forma de las mismas y está también sujeto a un tamaño de muestra adecuado, sin embargo, al contrario que en el caso de la estimación del tamaño poblacional, no hay métodos sencillos de estimación del tamaño de muestra adecuado (Krebs, 1999), por lo que en la presente práctica no tendremos en cuenta este aspecto del muestreo.

Los procedimientos que se abordan a continuación, representativos de los denominados análisis de patrones de puntos, forman parte de un conjunto de metodologías diversas integradas en el Análisis Espacial, una disciplina en rápido desarrollo y con amplias aplicaciones en Ecología y otras ciencias que utilizan datos espacialmente explícitos (Maestre et *al*., 2008).

**1.2.1- Método de la Relación Varianza/Media**

Este método consiste en hallar un índice de dispersión (a veces llamado Índice de Fisher), derivado del cociente entre la varianza del número de individuos por parcela respecto a la media de ese número.

En el caso de una distribución espacial aleatoria, como se explicó anteriormente, se espera que la distribución de frecuencias de parcelas con 0, 1, 2, etc. individuos siga una distribución de Poisson. Esta distribución presenta la característica de que su varianza coincide con su media; de forma que el cociente de ambas se espera que sea 1. Si se multiplica este cociente por el número de parcelas contadas (n) menos 1, se obtiene un índice que, mediante un test de Chi-cuadrado, permite evaluar la hipótesis nula “el índice vale 1x(n-1)”, con dos alternativas, menor o mayor de 1x(n-1). Si no se puede rechazar la hipótesis nula se concluirá que la distribución es aleatoria (conforme a una distribución de Poisson). La hipótesis nula se contrastará comparando el valor del índice obtenido con los valores tabulados en una distribución Chi-cuadrado, con un número de grados de libertad (gl) igual al número de parcelas menos uno (n-1), y con niveles de significación 1-(α/2) y α/2, siendo α el error tipo I admitido.

Si el valor del Índice está entre ambos valores de Chi-cuadrado, o es igual a uno de ellos, no se rechazará la hipótesis nula de aleatoriedad en la distribución.

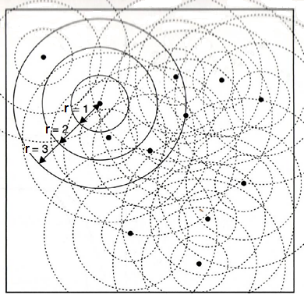
Si el Índice < Chi-cuadrado con un nivel de significación 1-(α/2) y con gl=n-1, se rechazará la hipótesis nula. Esto ocurrirá cuando el cociente varianza/media sea menor que 1, es decir, si la distribución es uniforme.

Si el Índice > Chi-cuadrado con un nivel de significación α/2 y con gl=n-1, se rechazará la hipótesis nula. Esto ocurrirá cuando el cociente varianza/media sea mayor que 1, es decir, si la distribución es agregada.

**1.2.2- Método de la K de Ripley**

La distribución espacial de los individuos de una población puede cambiar según la escala a la que se estudie. Esta es la razón de que, en ocasiones, dependiendo del tamaño de unidad de muestreo utilizada (cuadrícula en este caso), se obtengan diferentes patrones de distribución para una misma población. De ahí que la distribución espacial de los individuos de una población debe estudiarse a diferentes escalas.

El método de la K de Ripley efectúa ese análisis evaluando diversas escalas a la vez. Se puede aplicar si se dispone de las coordenadas de cada uno de los individuos de la población estudiada. En este caso la vamos a emplear para estudiar el patrón de distribución espacial de los individuos de una población, pero sirve también para evaluar hipótesis acerca de este patrón, estimar parámetros y ajustar modelos del efecto de variables ambientales sobre la población (Rozas y Camarero, 2005).

Para aplicar este método, se construye un círculo de radio “r” en torno a cada individuo, se cuenta el número de otros individuos que quedan dentro del círculo, se suman los números de individuos incluidos en cada uno de los círculos, se calcula la K para ese radio “r” (K(r)), se aumenta éste en una pequeña cantidad y se vuelve a hacer el cálculo de K, hasta un valor de r igual a 1/4 de la dimensión espacial menor del área estudiada (cuando ésta es rectangular).

Finalmente, se representan los sucesivos K(r) frente a los respectivos r. Comparando la línea que resulta de unir esos puntos con otra correspondiente a la que se esperaría obtener si los individuos se distribuyeran al azar, se puede estimar si el patrón de distribución de los individuos de la población cambia en función de la escala (representada aquí por el radio r).

La fórmula de este índice K es, para un radio r:

K(r) = [A/(n\*n)] ∑∑ [I( dij<=r)/eij)] , siendo

A: el área total estudiada

n: el número de individuos en el área global estudiada

∑∑: el segundo sumatorio calcula el total de individuos dentro de un círculo y el primero (desde i=1 hasta n) suma esos totales, de todos los círculos de radio r. Un mismo individuo puede estar incluido en más de un círculo por lo que, para el segundo sumatorio, contará tantas veces como círculos en los que se incluya. Estos sumatorios divididos por n representan el número medio de individuos que se encuentra en áreas circulares de radio r trazadas alrededor de cada uno de los individuos de la población.

I( dij<=r): variable que toma el valor 1 si dij <= r y 0 en caso contrario

eij: es un factor de corrección del efecto de borde del área de estudio, corresponde a la proporción del perímetro del círculo en torno a un punto i, y que contiene a un punto j, que está dentro del área de estudio, de forma que si la circunferencia está completa el valor de eij es 1 (esto nunca ocurrirá en los círculos cuyo radio sea mayor que la distancia de su centro, individuo i, al borde del área).

Al ser A/n el inverso de la densidad media de individuos en el área de estudio, K(r) representa el número de éstos que se observan como media en un área circular de radio r, dividido por esa densidad media. Si la distribución fuera aleatoria, el número de individuos que se esperaría encontrar en un círculo de radio r sería (n/A) x r2 (esto es, número de individuos esperados por unidad de superficie multiplicado por la superficie circular considerada), por lo que su K debería ser K(r)= r2. Así, el valor de K disminuirá, respecto al de una distribución aleatoria (K(r)= r2), cuando la distribución es uniforme, y aumentará cuando es por agregados. Si a una escala r el valor de K(r), junto con sus intervalos de confianza, está por encima o debajo del valor esperado (que viene dado por el área del círculo de radio r), se concluirá que, a esa escala, el patrón de distribución no es aleatorio, sino que es agregado, si está por encima, o uniforme, si está por debajo.

La K de Ripley es uno de los varios métodos existentes que tienen en cuenta la escala en el análisis del patrón de distribución espacial de una población; la elección de uno u otro depende del objetivo del estudio y de las características espaciales de la población estudiada, tales como la homogeneidad de la densidad de individuos a lo largo del área considerada. Ejemplos de estos estadísticos son el ”pcf” (pair correlation function) o el “O-ring” (Wiegand and Moloney, 2004). Con objeto de simplificar la práctica solo se aplicará la K de Ripley, sin entrar a considerar si es o no el método más idóneo en este caso.

**2) DESARROLLO DE LA PRÁCTICA**

Se estructura en dos etapas: 1) Recogida de datos en el campo. 2) Análisis de los datos en el laboratorio, con ayuda de ordenador, para la estimación del tamaño de muestra necesario, la abundancia y el tipo de distribución espacial de la población muestreada.

En la etapa 1 se delimita la población sobre la que se van a simular en la etapa 2, con ayuda del ordenador, diferentes muestreos a partir de los cuales se estimará el tamaño de muestra necesario, la abundancia de la población original (aunque realmente ya la conocemos, es la que hemos establecido en el muestreo de campo. Un muestreo real se haría en el campo a partir de una serie de unidades de muestreo seleccionadas en el área de distribución de la población) y su tipo de distribución espacial. El protocolo de muestreo que se simulará se conoce como *muestreo simple al azar sin reemplazamiento*, en el que las unidades de muestreo (cuadrículas) se eligen al azar dentro del área de estudio, asumida homogénea, pero eliminando la posibilidad de que una misma unidad sea elegida más de una vez en cada proceso de muestreo independiente. El uso de cuadrículas como unidades de muestreo, en las que se cuenta el número de individuos que aparecen en ellas, es adecuado para organismos inmóviles o relativamente poco móviles durante el tiempo en que se efectúa su censo (Krebs, 1999) como es el caso presente, con árboles.

**2.1- Recogida de datos (muestreo)**

**2.1.1-** Delimitar un área cuadrada de 20 m de lado con una cuerda marcada cada metro (o con cintas métricas). Establecer uno de sus vértices como origen de coordenadas espaciales (x,y). Esta superficie se tomará como el área total de distribución de la población a estudiar (árboles dentro del área, sin diferenciar entre especies (sólo para simplificar la práctica)).

**2.1.2-** Dividir el área de estudio en 4 subáreas iguales de 10x10 m con cuerdas marcadas cada metro (o con cintas métricas). Los alumnos se dividirán en un número equivalente de grupos que se harán cargo de las medidas a realizar en su respectiva subárea.

**2.1.3-** Con ayuda de una cinta métrica se medirán las coordenadas de cada árbol dentro de una subárea. Las coordenadas estarán referidas a las distancias en perpendicular (expresadas en m con dos decimales) desde el centro de la sección basal del árbol a cada uno de los bordes del área principal. De esta forma, la distancia perpendicular al eje X corresponderá a la coordenada “y” que se asignará al árbol. Estas medidas se anotarán a lápiz en un papel sobre una tabla con tres columnas (nº de árbol dentro de la subárea, coordenada x, coordenada y). Una vez anotadas las medidas de un árbol, éste se marcará para no repetir la medición de sus coordenadas. Al finalizar las medidas de los árboles de una subárea, se recogerán todas las marcas colocadas.

En el caso de árboles cuyo “centro” se encuentre en el límite de una subárea (efecto de borde), se asumirá dentro de ella si se localiza en el límite superior del eje X (borde derecho), o en el límite inferior del eje Y (borde inferior), de la subárea. De esta forma, si el “centro” de un árbol se encuentra en el límite inferior del eje X o en el superior del eje Y de una subárea (en los lados que forman su ángulo superior izquierdo), se considerará que está fuera de la misma. Este criterio se aplica tanto entre subáreas vecinas como a las subáreas con bordes coincidentes con él área total de estudio.

**2.2) Análisis de los datos**

**2.2.1-** Datos de partida

Crea, en el escritorio del ordenador, una carpeta denominada Ecología I (si hubiera una anterior creada con este nombre, bórrala). Incluye en ella el archivo Excel “Eucaliptal” disponible en los archivos FAITIC de la asignatura.

Abre el archivo “Eucaliptal” y copia en su HOJA 1 las distancias (coordenadas) x e y de cada uno de los árboles del área estudiada (no copies el número de árbol):

En las celdas A1 y B1 escribe x e y, respectivamente, ambas en minúsculas.

Debajo, escribe las coordenadas x e y de cada árbol (usa el punto para los decimales). El programa Excel deberá estar configurado para aceptar la coma decimal como punto.

Guarda los cambios. Recuerda borrar, al término de la práctica, estos datos y todo lo que añadas en él.

Estos datos se van a analizar empleando el programa R (R Development Core Team, 2015)

**2.2.2-** Apertura del programa R

“Clica” en el icono del mismo, una R mayúscula, a partir del escritorio o en el despegable de Inicio. En pasos sucesivos de la práctica, se irá copiando y ejecutando por partes el código (script) que se indica en las distintas secciones de este documento. Una vez abierta la hoja de trabajo de R (RConsole), se copiará de una vez todo el texto que aparece en rojo y azul de cada sección de las que siguen. El texto que figura en negro tras el símbolo # no es necesario copiarlo, pero, por comodidad, para no tener que estar seleccionando solo el texto en rojo y azul, se puede copiar también a la vez, no afecta a los resultados, se añade a título informativo sobre lo que hace el programa con el código que se copia. Las partes del texto en negro, no precedidas por #, no pueden copiarse, darían un error.

Si tras copiar en la consola de R una parte de este texto, el cursor permanece inmediatamente a continuación, presiona la tecla Entrar para que se ejecute el código copiado, de forma que el cursor saltará al siguiente renglón, y aparecerá precedido por el símbolo “>” (denominado “prompt” ≡ “da entrada a”). En ese momento, se pueden seguir añadiendo (copiando) nuevas instrucciones. En los casos que una instrucción del código no quepa en un único renglón, en vez del “>” aparecerá un “+” al inicio del siguiente, es ese momento se podrá seguir copiando el texto de la misma. Ten en cuenta que R diferencia entre mayúsculas y minúsculas de una misma letra, por lo que hay que mantener siempre la forma en que está escrito cualquier carácter.

**2.2.3-** Carga de los programas de R necesarios y los datos con los que se va a trabajar

**Salvo que se cierre R durante la práctica, no hace falta volver a repetir esta sección de la práctica (solo habrá que copiarla una vez)**.

Para poder trabajar con los datos del archivo “Eucaliptal” en R habrá que importarlos desde Excel. Para eso, sigue los siguientes pasos:

1- Copia y pega en la consola de R el siguiente texto, pero no lo ejecutes aún, es decir, no presiones Entrar a continuación (el cursor debería permanecer justo a continuación del texto).

Eucaliptal<-read.table ("clipboard", header=TRUE)

2- Abre el archivo “Eucaliptal”, selecciona todas las casillas con datos (incluidas las celdas A1 y B1 de su HOJA 1) y cópialas (presionando el botón derecho del ratón y eligiendo “Copiar”).

3- Vuelve a la Consola de R y, ahora, presiona Entrar. Lo único que ocurrirá es que el cursor pasará al renglón siguiente al del texto que copiaste anteriormente.

4- Copia y pega el siguiente texto para comprobar que has importado con éxito los datos del archivo Excel: se mostrarán solo los datos iniciales del archivo.

head(Eucaliptal)

5- Copia y pega

library(wisp) # aparecerá el mensaje Welcome to WiSP version 1.2.6-2 January 2013

**2.2.4-** Introducción de los parámetros necesarios

Debajo aparecen en azul los valores numéricos de los parámetros que se pueden modificar. **Si a lo largo de la práctica quieres cambiar alguno de estos parámetros,** escribe en la ventana de R (RConsole) el nombre del parámetro seguido de un igual o de los símbolos “menor que” y “menos” juntos (<-), también lo puedes copiar directamente desde aquí y pegarlo en R sin incluir el valor numérico; añade a continuación el valor deseado y presiona la tecla “Entrar” para ejecutar el cambio. Después, copia de nuevo la sección 2.2.4.2 y, si no aparece un mensaje de error, y no se ha cambiado de población, continúa copiando el resto del código (script) a partir del apartado 2.2.6.1 (inclusive). Recuerda que para copiar o escribir los parámetros en R hay que hacerlo en donde aparezca parpadeando el cursor precedido por el símbolo “>”.

2.2.4.1

xmax= 20 # longitud X, en m, del área estudiada

ymax= 20 # longitud Y, en m, del área estudiada

nsamples= 20 # número de muestras (“cuadrículas”) a utilizar

xsize= 2 # ancho de la cuadrícula de muestreo (en el eje de la x). Solo se pueden emplear valores que hagan que xmax/xsize=número entero>1

ysize= 2 # alto de la cuadrícula de muestreo (en el eje de la y). Solo se pueden emplear valores que hagan que ymax/ysize=número entero>1

a= 0.05 # alpha (α): nivel de significación para estimar el tipo de distribución y para establecer el grado de error tipo I que se admite para estimar el tamaño poblacional

beta= 0.1 # grado de error tipo II que se admite para estimar el tamaño poblacional. Si no quieres considerar este error, iguálalo a 0.5

d= 0.3 # proporción de error que se admite en el tamaño poblacional estimado

2.2.4.2

if (((xmax/xsize)-round(xmax/xsize) )| ((ymax/ysize)-round(ymax/ysize)) !=0)

stop("Con este tamaño de cuadrículas

\nno se puede dividir el área en un número entero de las mismas.

\nCambia xsize y/o ysize") # Si tras ejecutar el código anterior aparece un mensaje de texto, realiza los cambios indicados en él. Si no se muestra el mensaje, continúa copiando el código de la siguiente sección

**2.2.5-** Representación de la distribución de los ejemplares en el área estudiada y recuento de su número.

Salvo que se cierre R durante la práctica, independientemente de los cambios que se hagan en los parámetros, **no hará falta volver a copiar** **esta parte del texto** (es decir, **sólo** hace falta **copiarla la primera vez** que se carga el código)

mireg<-generate.region(x.length=xmax,y.width=ymax)

midens<-generate.density(mireg)

mipars<-setpars.population(midens, number.groups = length(Eucaliptal$x), size.method = "user", size.values = 1, size.prob = 1, size.min = 1, size.max = 1)

mipob <- generate.population(mipars)

mipob$posx<-Eucaliptal$x

mipob$posy<-Eucaliptal$y

plot(Eucaliptal$x, Eucaliptal$y, pch=16) # la gráfica que se muestra representa la posición de los distintos individuos en el área de estudio

*Pregunta*

# 1) *A la vista del gráfico de distribución espacial de la población, ¿qué tipo de distribución crees que presenta?*

**2.2.6-** Muestreo de la población con cuadrículas elegidas al azar en el área de distribución de la misma: Estimación del área muestreada

2.2.6.1

if (xsize\*ysize\*nsamples>xmax\*ymax){

print(c("El área de muestreo es mayor que la de distribución de la población",

"Tiene que cumplirse que: área de cuadrícula x nº de cuadrículas <= área total de la población",

"Disminuir nsamples y/o xsize y/o ysize"), quote=FALSE)

} else { xsize\*ysize\*nsamples/(xmax\*ymax)}

2.2.6.2

Si tras ejecutar el código anterior aparece un mensaje de texto, realiza los cambios indicados en él; si aparece un número, ejecuta la línea de código que sigue, y escribe el número anterior en el renglón en que se mostrará tras su ejecución, a continuación de “1:” que se muestra en la consola de R. Tras hacerlo, **presiona 2 veces la tecla Entrar (tras “2:” no se pone nada). Si se efectúa correctamente, aparecerá el mensaje “Read 1 item”**

proporción.área.muestreada<-scan()

**2.2.7-** Muestreo de la población con cuadrículas elegidas al azar en el área de distribución de la misma: Realización del muestreo

**Si no se cambia ningún parámetro y solo se quiere volver a muestrear la población, basta con copiar el código a partir de aquí**

midispars<-setpars.design.pl(mireg, n.interval.x= xmax/ xsize, n.interval.y= ymax/ ysize,

method="random", area.covered=proporción.área.muestreada)

midis<-generate.design.pl(midispars)

misamp<-generate.sample.pl(mipob,midis)

plot.sample.pl(misamp, whole.population=T) # la gráfica que se muestra representa la posición de las cuadrículas de muestreo sobre la población estudiada

**2.2.8-** Cuenta del número de individuos por cuadrícula y estimación del tamaño de la población (En toda el área de distribución; el real es el número de ejemplares contados en el campo)

summary(misamp) # muestra los resultados e información sobre el muestreo

ns<-as.vector(table(misamp$unit))

ns<-if (length(ns)<nsamples) c(ns,rep(0,nsamples-length(ns))) else ns

mipobest<-point.est.pl(misamp)

print (c("Tamaño de población estimado = ",mipobest$Nhat.ind, "Tamaño real de la población = ", length(mipob$posx)), quote= FALSE)

**2.2.9-** Evaluación del tipo de distribución espacial de esta población

index<-(nsamples-1)\*(var(ns)/mean(ns)) # calcula el índice de dispersión para evaluar el tipo de distribución

print(c("Media de individuos por cuadrícula =",round(mean(ns),2)), quote=FALSE)

print(c("Varianza de individuos por cuadrícula =",round(var(ns),3)), quote=FALSE)

print(c("Índice de dispersión =",round(index,2)), quote=FALSE)

res<-ifelse(index<qchisq(a/2, nsamples-1), "Distribución uniforme", ifelse(index>qchisq(1-(a/2) , nsamples-1), "Distribución agregada", "Distribución aleatoria")) # compara el index, para el nivel de significación elegido, con los valores tabulados en una tabla de Chi-square al nivel de significación alfa elegido (parámetro “a” inicial)

print(c("Tipo de distribución = ",res), quote=FALSE)

**2.2.10-** Cálculo del tamaño de muestra necesaria

Número de cuadrículas de muestreo necesario para estimar, con los niveles de significación deseados (errores tipo I y II, y rango de error máximo aceptado), el tamaño de la población

dens<-ns/(xsize\*ysize)

densmed<-mean(dens)

densvar<-var(dens)

di<-d\*mipobest$Nhat.ind/(xmax\*ymax)

ntent=nsamples

for (i in 1:10){

t<-qt(a/2,ntent-1,lower.tail=FALSE)

F<-qf(beta,ntent-1,nsamples-1, lower.tail=FALSE)

nest<- densvar\*(t^2)\*F/(di^2)

ntent=nest

}

print(c("Tamaño de muestra necesario (Población grande) = ", ceiling(nest)), quote= FALSE)

print (c("Tamaño de muestra necesario (Población pequeña) = ",ceiling(nest/(1+( nest-1)/(xmax\*ymax/(xsize\*ysize))))), quote=FALSE)

**2.2.11-** Cálculo de intervalos de confianza de la estimación de abundancia por Bootstrap

Aplica el código que se muestra a continuación (el cálculo se realiza por un procedimiento no paramétrico, perteneciente a los métodos de muestreo repetido (“resampling methods”), en concreto mediante un método de “Bootstrap”.

miicest<-int.est.pl(misamp, vlevels = c(a/2, 1-a/2)) # estima el intervalo de confianza para el tamaño de la población

print(c("Población media estimada = ",round(miicest$boot.mean$Nhat.grp[1])),quote=FALSE)

print(c("Límite inferior del intervalo de confianza = ",floor(miicest$ci$Nhat.grp[1])),quote= FALSE)

print(c("Límite superior del intervalo de confianza = ",ceiling(miicest$ci$Nhat.grp[2])),quote= FALSE)

Pregunta

*#* 2) *¿Qué efecto produciría sobre el tamaño de muestra necesario (estimado a partir del muestreo con cuadrículas) el utilizar un a (alfa) menor (por ej. 0.01 en vez de 0.05? ¿Por qué? ¿Cómo podrías comprobar que tu respuesta es correcta?*

**2.2.12-** Aplicación del método de la K de Ripley

Se utilizará el programa “spatstat” de R (Véase un ejemplo de empleo de este programa con este método en Davies et *al*., 2014). El análisis de los resultados se basará, en este caso, exclusivamente en observar sobre una gráfica la disposición de los valores de K obtenidos en relación a los intervalos de confianza de los K esperados si la distribución fuese de tipo aleatoria. **Si no se cambia de población de estudio esta parte de la práctica solo es necesaria aplicarla una vez**, el resultado no cambia al modificar los parámetros de partida que determinan el tamaño y número de las cuadrículas de muestreo ya que no intervienen en su cálculo, sí estará influido por el nivel de significación α elegido.

2.2.12.1

library(spatstat)

Eucaliptalppp<-ppp(Eucaliptal$x,Eucaliptal$y,xrange=c(0,xmax),yrange=c(0,ymax))

plot(Kest(Eucaliptalppp,correction="best"), col=c(2,1)) # la línea roja representa los valores de K observados, la negra los valores esperados si la distribución fuese aleatoria; cuando la línea roja está por encima de la negra es muy posible que a ese “r” (escala) el patrón sea por agregados; si está por debajo, uniforme. El valor de r está escalado a un máximo de 1 (que equivaldría a 20 en este caso.

2.2.12.2

plot(envelope(Eucaliptalppp,fun=c("Kest")),col=c(2,1,8,2)) # igual al gráfico anterior, pero se añade una franja gris que representan el rango de significación para una distribución aleatoria. En las zonas en las que la banda gris incluya la línea roja, el patrón es aleatorio, si la línea roja queda fuera y por encima, por agregados, fuera y por debajo, uniforme.

Código para la K transformada (Se denomina función L, no cambia el resultado de K, solamente se incluye aquí por si se quiere obtener una representación gráfica un poco más fácil de interpretar, el patrón aleatorio aparece como una recta)

plot(Lest(Eucaliptalppp,correction="best"),col=c(2,1))

plot(envelope(Eucaliptalppp,fun=c("Lest")),col=c(2,1,8,2))

**2.3- Efecto sobre los resultados del muestreo del número y tamaño de las cuadrículas empleadas**

Utilizando el código de este apartado, se efectuarán sucesivas simulaciones (actualmente su número está fijado en 10) de muestreos en los que se combinan 3 dimensiones de lado de cuadrícula (1.0, 2.0 y 4.0) con 3 tamaños de muestra (10, 15 y 20 cuadrículas). Los resultados de esas simulaciones, recogidos en un archivo Excel, se utilizarán para responder a las preguntas que se plantean al final de esta sección. Estas combinaciones están adaptadas al área de distribución establecida para la población estudiada (20x20 m), en caso de variarla convendría revisar las dimensiones de cuadrícula y tamaños de muestra a emplear.

Primero, haz los tres tamaños de muestra, en orden creciente, de las cuadrículas de 1.0x1.0; después, aumenta el tamaño de cuadrícula y repite el proceso, hasta completar las 9 combinaciones posibles.

Copia y ejecuta el código que se muestra a continuación, de forma progresiva; en cada ocasión hasta la línea de trazos que separa cada parte. Se asume que los datos de coordenadas del archivo “Eucaliptal”, el programa wisp y los parámetros “xmax”, “ymax”, “a”, “beta” y “d” ya están cargados en R tras la realización del apartado anterior de la práctica, en caso de que no sea así, o de que se quieran modificar, hay que volver a ejecutarlos.

nsamples= 10

xsize= 1

ysize= 1

if (((xmax/xsize)-round(xmax/xsize) )| ((ymax/ysize)-round(ymax/ysize)) !=0)

stop("Con este tamaño de cuadrículas

\nno se puede dividir el área en un número entero de las mismas.

\nCambia xsize y/o ysize")

nsimul<- 10

mireg<-generate.region(x.length=xmax,y.width=ymax)

midens<-generate.density(mireg)

mipars<-setpars.population(midens, number.groups = length(Eucaliptal$x), size.method = "user", size.values = 1, size.prob = 1, size.min = 1, size.max = 1)

mipob <- generate.population(mipars)

mipob$posx<-Eucaliptal$x

mipob$posy<-Eucaliptal$y

plot(Eucaliptal$x, Eucaliptal$y, pch=16)

if (xsize\*ysize\*nsamples>xmax\*ymax){

print(c("El área de muestreo es mayor que la de distribución de la población",

"Tiene que cumplirse que: área de cuadrícula x nº de cuadrículas <= área total de la población",

"Disminuir nsamples y/o xsize y/o ysize"), quote=FALSE)

} else { xsize\*ysize\*nsamples/(xmax\*ymax)}

#============================================================================

proporción.área.muestreada<-scan() # Véase sección 2.2.6.2

#============================================================================

midispars<-setpars.design.pl(mireg, n.interval.x= xmax/ xsize, n.interval.y= ymax/ ysize,

method="random", area.covered=proporción.área.muestreada)

Mdisp<- NULL

Vdisp<- NULL

Idisp<- NULL

Tdist<-NULL

Pobests<-NULL

ns<-NULL

ngran<- NULL

npeq<- NULL

nmicest<- NULL

ninf<- NULL

nsup<- NULL

for (v in 1:nsimul){

midis<-generate.design.pl(midispars)

misamp<-generate.sample.pl(mipob,midis)

plot.sample.pl(misamp, whole.population=T)

ns<-as.vector(table(misamp$unit))

ns<-if (length(ns)<nsamples) c(ns,rep(0,nsamples-length(ns))) else ns

mipobest<-point.est.pl(misamp)

Pobests<-c(Pobests,mipobest$Nhat.ind)

index<-(nsamples-1)\*(var(ns)/mean(ns))

dist<-ifelse(index<qchisq(a/2, nsamples-1), " Uniforme", ifelse(index>qchisq(1-(a/2) , nsamples-1), "Agregada", "Aleatoria"))

dens<-ns/(xsize\*ysize)

densmed<-mean(dens)

densvar<-var(dens)

di<-d\*mipobest$Nhat.ind/(xmax\*ymax)

ntent=nsamples

for (i in 1:10){

t<-qt(a/2,ntent-1,lower.tail=FALSE)

F<-qf(beta,ntent-1,nsamples-1, lower.tail=FALSE)

nest<- densvar\*(t^2)\*F/(di^2)

ntent<-nest

}

miicest<-int.est.pl(misamp,plot=FALSE)

nmest<-round(miicest$boot.mean$Nhat.grp[1])

nicinf = floor(miicest$ci$Nhat.grp[1])

nicsup = ceiling(miicest$ci$Nhat.grp[2])

ngran<-c(ngran, floor (nest))

npeq<- c(npeq, floor(nest/(1+( nest-1)/(xmax\*ymax/(xsize\*ysize)))))

Mdisp<-c(round(mean(ns),2),Mdisp)

Vdisp<-c(round(var(ns),3),Vdisp)

Idisp<-c(Idisp, round(index,2))

Tdist<-c(Tdist,dist)

nmicest<-c (nmicest,round(nmest,0))

ninf<- c(ninf,round(nicinf,0))

nsup<-c(nsup,round(nicsup,0))

}

# En caso de que aparezca el mensaje de debajo, repite el proceso (hasta que no se muestre). Esto ocurrirá principalmente con el tamaño de cuadrícula pequeño (el de 1x1). Bastará con que vuelvas a pegar el último tramo copiado (desde la anterior #=====…).

# Error in generate.sample.pl(mipob, midis) :

# There were zero detections.

Tabla<- data.frame(xsize,ysize, nsamples, Párea= proporción.área.muestreada, Pobests, Mdisp,Vdisp,Idisp,Tdist,CV=round(Vdisp\*100/Mdisp,2),ngran,npeq,nmicest,ninf,nsup,rangoic=nsup-ninf,snecpeq=round(npeq\*xsize\*ysize,2), Error=abs(length(Eucaliptal$x)-nmicest))

#============================================================================

# Finalmente si tras la ejecución del código anterior no se generó ningún mensaje de error, copia y pega:

write.table(Tabla,"clipboard",sep="\t",col.names=FALSE)

# **Después de escribir esta sentencia en R, abre el archivo de Excel “Eucaliptal”, en donde se copiaron los datos, y ejecuta “Pegar” en la casilla A3 de la HOJA 2. Si se efectúan más simulaciones con otros tamaños, pégalas en la primera casilla en blanco de la izquierda, justo por debajo de los datos de la simulación anterior. Una vez acabada la práctica deja el archivo tal como estaba (Sin nada en la HOJA 1 y solo con el texto original de las filas 1 y 2 de la HOJA 2)**

#============================================================================

Preguntas

3) *¿Influye el área total muestreada en el error de estimación del tamaño de la población (calculado como la diferencia entre el estimado y el real)?*

4) *¿Cuál sería el tamaño de cuadrícula más adecuado para estimar la abundancia de esta población? ¿Coincide ese tamaño con el más adecuado si se tuviera en cuenta el coste de muestreo (estimado como proporción del área total de la población que habría que muestrear dado un tamaño de muestra adecuado y el tamaño de la cuadrícula correspondiente)?*

5) *¿Qué conclusión se obtiene sobre la distribución espacial de los árboles? ¿Coinciden los resultados según el tamaño de cuadrícula elegido?*

Preguntas generales

6) *Supón dos poblaciones con el mismo número de individuos, una con distribución aleatoria y otra con distribución uniforme, ¿se requerirá el mismo número de muestras (cuadrículas) para estimar las abundancias de ambas con iguales errores?*

7) *¿Conoces algún método de estimación del patrón de distribución de una población que no se base en el empleo de cuadrículas de muestreo?*

8) *La distribución espacial de los individuos de la población estudiada, ¿cambia con la escala?*

**4) BIBLIOGRAFÍA**

Baddeley, A. and Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns Journal of Statistical Software, **12 (**6): 1-42.

Brower, J. E.; Zar, J. H. and von Ende, C. N. (1990). Field and Laboratory Methods for General Ecology. WCB/McGraw-Hill. 273.

Davies, A. B.; Levick, S. R.; Asner, G. P.; Robertson, M. P.; Rensburg, B. J. and Parr, C. L. (2014). Spatial variability and abiotic determinants of termite mounds throughout a savanna catchment. Ecography 37: 852–862.

Elzinga, C. L.; Salzer, D. W; Willoughby, J. W. and Gibbs, J. P. (2001). Monitoring Plant and Animal populations. Blackwell Science, Inc. 360 pp.

Haase, P. (1995). Spatial Pattern Analysis in Ecology Based on Ripley's K-Function: Introduction and Methods of Edge Correction. Journal of Vegetation Science, 6(4): 575-582.

Krebs, C. J. (1999). Ecological methodology. Addison Wesley Lohgman. 620 pp.

Maestre, F. T.; Escudero, A. y Bonet, A. (eds.) (2008). Introducción al Análisis Espacial de Datos en Ecología y Ciencias Ambientales. Métodos y Aplicaciones. Universidad Rey Juan Carlos. Dykinson, S. L.: 849 pp.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>

Ripley, B.D. (1976). The second–order analysis of stationary point processes. Journal of Applied Probability 13: 255-266.

Rozas, V. y Camarero J. J: (2005). Técnicas de análisis espacial de patrones de puntos aplicadas en ecología forestal. Invest Agrar. Sist. Recur. For., 14(1): 79-97.

Seber, G. A. F. (1982). The estimation of animal abundance and related parameters. Edward Arnold. 654 pp.

Southwood, T. (1978). Ecological methods, with particular reference to the study of insect populations. ELBS. 524 pp.

Sutterland, W. J. (Ed.) (2006). Ecological Census Techniques, a handbook. Cambridge University Press. 432 pp.

Wiegand T., and K. A. Moloney (2004). Rings, circles and null-models for point pattern analysis in ecology. Oikos 104:209-229.

Zar, J. H. (2009). Biostatistical Analysis. Pearson. 960 pp.

Zucchini, W., Borchers, D.L., Erdelmeier, M., Rexstad, E. and Bishop, J. (2007). WiSP 1.2.4. Institut fur Statistik und Okonometrie, Geror-August-Universitat Gottingen, Platz der Gottinger Seiben 5, Gottingen, Germany.