

## PRÁCTICA 1: MUESTREO

### Introducción

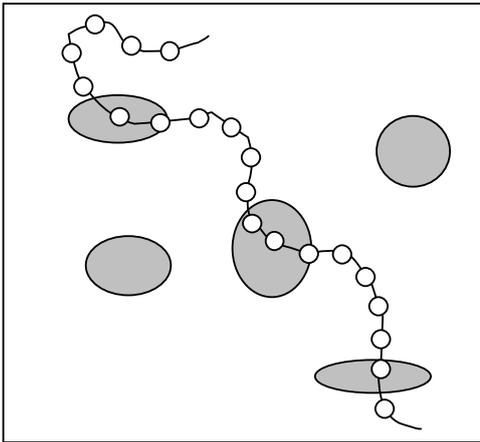
La investigación ecológica se basa en la medición de parámetros de los organismos y del medio en el que viven. Este proceso de toma de datos se denomina muestreo. En la mayoría de los casos, resulta imposible medir todos los organismos o las distintas partes del ambiente implicados en un estudio y el muestreo implica un proceso de selección de organismos o ambientes a medir. A partir de este grupo seleccionado es posible inferir las características de una especie, población o comunidad sin necesidad de examinar la totalidad del colectivo. Suponemos, por tanto, que el muestreo es fiable y representativo, es decir, que existe una alta probabilidad de que las estimas obtenidas se aproximen a los valores reales de las variables examinadas. No obstante, la representatividad del muestreo varía en función del tamaño de muestreo (número de muestras), y en función de la distribución espacial y temporal de las muestras. El tamaño de muestreo responde a un compromiso entre minimizar el esfuerzo durante la fase de recogida de datos y maximizar la fiabilidad de las conclusiones que se deriven de esos datos. La importancia relativa de la distribución de las muestras dependerá de los objetivos concretos de la investigación y de las características de las poblaciones o comunidades estudiadas. El **objetivo** de esta práctica es desarrollar un protocolo básico de muestreo sobre organismos vivos y valorar la influencia de algunos factores que afectan a la representatividad del muestreo.

El trabajo se realizará sobre una **comunidad pratense** y el parámetro básico a muestrear será la **abundancia** de distintas especies herbáceas dentro de esta comunidad. La estima de la abundancia de un organismo puede realizarse utilizando diferentes tipos de estimadores. Algunos, como la densidad ( $n^{\circ}$  de individuos por unidad de superficie), están basados en la noción de individuo. En ocasiones puede resultar más interesante utilizar unidades que no consideran directamente el número de individuos sino otros aspectos que pueden tener importancia funcional, como la cobertura o la biomasa. La cobertura es el porcentaje de la superficie muestreada que está recubierta por la proyección vertical de los organismos estudiados. La biomasa es el peso total de todos los individuos de una especie contenidos en una superficie o volumen estándar de muestreo. La elección del tipo de estimador puede depender de los objetivos del estudio, de la dificultad asociada a la medición de cada tipo, o de condicionantes relativos a la influencia sobre los organismos muestreados. Por ejemplo, la estima de biomasa implica la recolección de la muestra, mientras que la estimación de la cobertura puede tener un efecto nulo sobre los organismos estudiados.

### *Medidas de cobertura*

La medida de cobertura asigna un porcentaje de aparición a una especie concreta en relación a una superficie muestreada, que estará cubierta total o parcialmente por un conjunto de organismos. En el caso de la comunidad pratense, podremos asignar a cada especie un porcentaje de cobertura relativo a una superficie estándar, establecida como "unidad de muestreo", y un porcentaje de cobertura al suelo no cubierto por plantas, de forma que la suma de todas las coberturas sea 100. También podremos considerar que hay superposición entre las hojas de distintas especies y que la suma de todas las coberturas sea superior a 100. En nuestro sistema de estudio podremos calcular el porcentaje de cobertura de las distintas especies mediante dos métodos:

- A) atribuyendo *a ojo* el porcentaje aproximado de superficie ocupada por cada especie presente en la unidad de muestreo
- B) estimando el porcentaje de puntos ocupados por la especie relativo a un conjunto de puntos distribuidos al azar en la unidad de muestreo. Por ejemplo, podemos establecer los puntos de evaluación de presencia mediante un rosario de nudos que se lanza aleatoriamente dentro de la unidad de muestreo (Figura 1)



Número total de puntos: 20  
 Coincidencias: 6  
 Cobertura estimada:  $(6/20) \times 100 = 30\%$

Figura 1. Representación de muestreo de puntos mediante rosario de nudos. Las formas sombreadas representan la superficie ocupada por una especie determinada

*Medidas de biomasa*

La medida de biomasa de las distintas especies de la comunidad praterense se realizará mediante la siega de toda la materia vegetal aérea en una superficie estándar de muestreo, la separación del material vegetal por especies en el laboratorio, y el peso en la balanza del material de cada especie. Se expresará como **porcentaje de biomasa** (peso de la especie relativo al peso total de la muestra).

**Desarrollo de la práctica**

**Tarea 1. Comparación entre métodos de estima de abundancia**

1.1. Comparación entre medidas de cobertura

El objetivo de esta tarea será evaluar la variabilidad en la medida de cobertura en función del método de estima (a ojo y mediante muestreos de puntos con distinto número de puntos). La unidad básica de muestreo será un **cuadrado de 50x50** cm dispuesto sobre el prado. Sobre ese cuadrado: 1) se identificarán todas las especies presentes; 2) se estimará la cobertura de cada especie a ojo; y 3) se estimará la cobertura de cada especie con tres grupos de puntos (20, 40 y 80 puntos), que se realizarán con 1, 2 y 4 tiradas sucesivas del rosario sobre la superficie de muestreo. Se realizarán 4 réplicas de muestreo, es decir, el cuadrado colocará en 4 puntos aleatorios donde se realizarán todas las medidas. Los datos podrán recolectarse en una tabla similar a esta, calculándose el valor de la media y la desviación típica\* de cada grupo de réplicas para cada especie y método de estima

Especie	Método estima	Réplica 1	Réplica 2	Réplica 3	Réplica 4	Media	Desv. típica
Sp. 1	Ojo						
	20 puntos						
	40 puntos						
	80 puntos						
Sp. 2	Ojo						
	20 puntos						
	40 puntos						
	80 puntos						
Sp. 3	...						

$$* S = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

*Representación gráfica*

Para las cuatro especies más abundantes, se representará el valor de la media ( $\pm S$ ) en función del método de medida.

1.2. Comparación de cobertura y biomasa

En esta tarea se compararán las estimas de abundancia de las especies herbáceas obtenidas a partir de datos de cobertura y de biomasa. La unidad básica de muestreo será un **cuadrado de 30x30** cm dispuesto sobre el prado. Sobre esa unidad: 1) se identificarán las distintas especies; 2) se estimará la cobertura de las mismas mediante 60 puntos (3 tiradas de rosario) distribuidos aleatoriamente; y 3) se estimará la biomasa de cada especie mediante la siega del cuadrado, la separación por especies en el laboratorio, y el pesado de cada especie en el laboratorio.

**Cada grupo intercambiará sus datos con los grupos restantes**, lo que permitirá establecer comparaciones sobre 5-8 réplicas. Los datos podrán recolectarse en una tabla similar a esta, calculándose el valor de la media de cada grupo de réplicas para cada especie y método de estima.

Especie	Método estima	Grupo 1	Grupo 2	Grupo 3	...	Media
Sp. 1	60 puntos					
	% Biomasa					
Sp. 2	60 puntos					
	% Biomasa					
Sp.3	60 puntos					
	% Biomasa					
	...					

*Representación gráfica y análisis*

Se representará la biomasa en función de la cobertura, en diagramas X,Y. Se realizará diagramas considerando la variación "dentro de especie" aparecida al considerar las réplicas de los distintos grupos, para las dos especies más representativas (cada punto del diagrama X,Y corresponderá a una réplica muestral), así como un diagrama contrastando los valores promedio de cobertura y biomasa de todas las especies

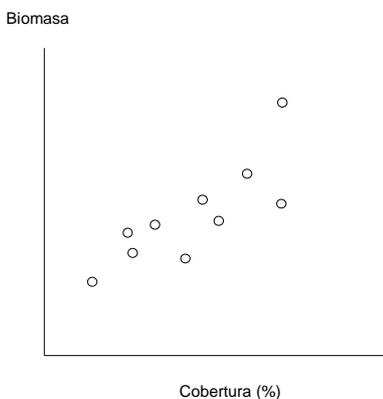


Figura 2. Representación de la biomasa en función de la cobertura, cada punto representa los valores promedio de una especie

Para cada contraste biomasa-cobertura representado anteriormente, se calculará el grado de **correlación entre ambos parámetros**, y el nivel de significación estadística de dicha correlación mediante el coeficiente de correlación de rangos de Spearman:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$$

Donde  $d_i$  es la diferencia entre los rangos de  $x_i$  e  $y_i$ , y  $n$  es el número de pares de valores (véanse anexos de explicación del Coeficiente de Spearman y tabla de niveles de significación).

**Tarea 2. Análisis de la influencia de la distribución espacial de las muestras**

La disposición espacial de las muestras también puede afectar a los resultados del muestreo. La elección de un diseño determinado puede hacerse en función del enfoque de la investigación o de las características de la comunidad y el medio físico en donde se desarrolla el trabajo. En general se habla de *muestreo aleatorio* cuando las unidades de muestreo se disponen al azar sobre el área de estudio, y de *muestreo dirigido* cuando las unidades se colocan siguiendo una estructura espacial concreta. El *transecto* es una modalidad de muestreo dirigido y consiste en colocar las unidades de muestreo siguiendo una disposición rectilínea. Suele utilizarse cuando se observa un gradiente de algún factor del medio. La disposición espacial de las muestras a lo largo del gradiente permite examinar si existe alguna relación entre la variable medida en el muestreo y los factores que conforman el gradiente o, simplemente, con la posición a lo largo del transecto.

Se realizará un muestreo en transecto disponiendo 10 unidades de 50x50 cm a lo largo de 13.5 m, separando las unidades entre sí 1.5 m, y disponiendo el transecto desde la parte alta a la parte baja del prado, paralelo a la carretera. En cada unidad de muestreo se estimará la cobertura de las distintas especies presentes mediante 80 puntos (4 tiradas de rosario). Cada grupo realizará un transecto, y los distintos transectos se dispondrán paralelos y separados entre sí 1.5 m. De esta forma, todos los grupos muestrearán 10 clases de distancia respecto al punto inicial del transecto. **Cada grupo intercambiará sus datos con los grupos restantes.** Los datos podrán recolectarse en una tabla similar a esta, calculándose el valor de la media y la desviación típica de cada grupo de réplicas para cada especie.

Especie	Distancia	Grupo 1	Grupo 2	Grupo 3	...	Media	Desv. típica
Sp. 1	0	% cobertura					
	1.5						
	3						
	4.5						
	6						
	7.5						
	9						
	10.5						
	12						
	13.5						
Sp. 2	0						
	1.5						
	...						

*Representación gráfica y análisis*

Se seleccionarán las dos especies más representativas y para cada una de ellas se representará gráficamente el valor de la media ( $\pm S$ ) en función de la distancia a lo largo del transecto. Se calculará el coeficiente de correlación de Spearman entre la distancia a lo largo del transecto y las coberturas de las dos especies más representativas.

### Cuestiones a desarrollar en relación a los datos obtenidos:

- 1.- ¿Qué ventajas e inconvenientes pueden tener las distintas formas de estimar la abundancia?
- 2.- ¿Existen diferencias entre las estimaciones de cobertura total (“a ojo”) y las obtenidas mediante puntos distribuidos al azar? En caso afirmativo ¿a qué pueden deberse?
- 3.- ¿Cómo influye el número de puntos de muestreo en la estimación de cobertura? ¿Cuál sería el número de puntos mínimo que se debería utilizar en este caso? ¿Es útil el valor de S para estimar ese número mínimo? ¿por qué?
- 4.- ¿Cómo se explicaría la correspondencia entre la biomasa y la cobertura como estimas de abundancia?
- 5.- ¿Existen diferencias en la fuerza de la relación entre biomasa y cobertura en los diseños “dentro de especie” y “entre especies”? En caso afirmativo ¿a qué pueden deberse?
- 6.- ¿Existen tendencias claras de estructura espacial en alguna de las especies muestreadas en el transecto? Además de los cambios en el valor de la media a lo largo del transecto ¿cambia S? ¿a qué crees que pueden deberse los cambios?
- 7.- Las tendencias espaciales ¿son semejantes, independientes o contrarias entre especies? ¿a qué crees que pueden deberse?

### Bibliografía

- Hairston, N. G. (1989). Ecological Experiments. Purpose, Design, and Execution. Cambridge University Press.  
Sutherland, W. (2006). Ecological census techniques: a handbook. Cambridge University Press.  
Zar, J. H. (1984). Biostatistical Analysis. Prentice-Hall, Inc.

for each  $A$ , in the same sequence as described in Section 12.4. The appropriate standard error is

$$SE = \sqrt{\frac{2}{n_A - 3}} \quad (19.29)$$

if the samples  $A$  and  $B$  are of the same size, or

$$SE = \sqrt{\frac{1}{n_A - 3} + \frac{1}{n_B - 3}} \quad (19.30)$$

if  $n_A \neq n_B$ . The critical value is  $q'_{\alpha(1), \infty, p}$  (from Table B.6) or  $q'_{\alpha(2), \infty, p}$  (from Table B.7) for the one-tailed or two-tailed test, respectively.

**Multiple Contrasts Among Correlation Coefficients.** Section 12.5 introduces the concepts and procedures of multiple contrasts among means; these are multiple comparisons involving groups of means. In a similar fashion, multiple contrasts may be examined among correlation coefficients (Marascuilo, 1971: 454-455). We again employ the  $z$ -transformation and calculate, for each contrast, the test statistic

$$S = \frac{\sum c_i z_i}{SE} \quad (19.31)$$

where

$$SE = \sqrt{\sum c_i^2 \sigma_i^2} \quad (19.32)$$

and  $c_i$  is a contrast coefficient, as described in Section 12.5. (For example, if we wished to test the hypothesis  $H_0: (\rho_1 + \rho_2)/2 - \rho_3 = 0$ , then  $c_1 = \frac{1}{2}$ ,  $c_2 = \frac{1}{2}$ , and  $c_3 = -1$ .) The critical value for this test is\*

$$S_\alpha = \sqrt{\chi_{\alpha, (k-1)}^2} \quad (19.34)$$

### 19.9 RANK CORRELATION

If we have data obtained from a bivariate population that is far from normal, then the correlation procedures discussed thus far are generally inapplicable. Instead, we may operate with the ranks of the measurements for each variable. Two different rank correlation methods are commonly encountered, that proposed by Spearman† (1904), and that of Kendall‡ (1962).

Example 19.11 demonstrates Spearman's rank correlation procedure. After each measurement of a variable is ranked, as done in previously described nonparametric

\*Since  $\chi_{\alpha, \nu}^2 = \nu F_{\alpha(1), \nu, \infty}$ , it is equivalent to write

$$S_\alpha = \sqrt{(k-1)F_{\alpha(1), (k-1), \infty}} \quad (19.33)$$

but Equation 19.34 is preferable, because it engenders less rounding error in the calculations.

†Charles Edward Spearman (1863-1949), English psychologist and statistician.

‡Maurice George Kendall (1907-1983), English statistician.

#### Example 19.11

The Spearman rank correlation coefficient, computed for the data of Example 19.1.

$X$	Rank of $X$	$Y$	Rank of $Y$	$d_i$	$d_i^2$
10.4	4	7.4	5	-1	1
10.8	8.5	7.6	7	1.5	2.25
11.1	10	7.9	11	-1	1
10.2	1.5	7.2	2.5	-1	1
10.3	3	7.4	5	-2	4
10.2	1.5	7.1	1	0.5	0.25
10.7	7	7.4	5	2	4
10.5	5	7.2	2.5	2.5	6.25
10.8	8.5	7.8	9.5	-1	1
11.2	11	7.7	8	3	9
10.6	6	7.8	9.5	-3.5	12.25
11.4	12	8.3	12	0	0
$n = 12$					$\sum d_i^2 = 42.00$

$$r_s = 1 - \frac{6 \sum d_i^2}{n^3 - n}$$

$$= 1 - \frac{6(42.00)}{1716}$$

$$= 1 - 0.147$$

$$= 0.853$$

To test  $H_0: \rho_r = 0, H_1: \rho_r \neq 0$ ,

$$(r_s)_{0.05(2), 12} = 0.587 \text{ (from Table B.19)}$$

Therefore, reject  $H_0$ .

$$P < 0.001$$

To employ the correction for ties (see Equation 19.37):

among the  $X$ 's there are 2 measurements of 10.2 cm and 2 of 10.8 cm, so

$$\sum T_x = \frac{(2^3 - 2) + (2^3 - 2)}{12} = 1;$$

among the  $Y$ 's there are 2 measurements tied at 7.2 cm and 2 at 7.8 cm, so

$$\sum T_y = \frac{(2^3 - 2) + (2^3 - 2)}{12} = 1;$$

therefore,

$$r_s = \frac{(12^3 - 12)/6 - 42.00 - 1 - 1}{\sqrt{[(12^3 - 12)/6 - 2(1)][(12^3 - 12)/6 - 2(1)]}} = \frac{242}{284} = 0.852;$$

and the hypothesis test proceeds exactly as above.

testing procedures, Equation 19.1 can be applied to the ranks to obtain the Spearman rank correlation coefficient,  $r_s$ . However, a simpler computation is

$$r_s = 1 - \frac{6 \sum d_i^2}{n^3 - n} \quad (19.35)$$

where  $d_i$  is a difference between  $X$  and  $Y$  ranks.

The value of  $r_s$ , as an estimate of the population rank correlation coefficient,  $\rho_r$ , may range from  $-1$  to  $+1$ , and it has no units; however, its value is not to be expected

**TABLE B.19** CRITICAL VALUES OF THE SPEARMAN RANK CORRELATION COEFFICIENT,  $r_s$

$\alpha(2)$ :	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
$\alpha(1)$ :	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
$n$									
4	0.600	1.000	1.000						
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.217	0.406	0.503	0.587	0.678	0.727	0.769	0.818	0.846
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.791	0.824
14	0.200	0.367	0.464	0.538	0.626	0.679	0.723	0.771	0.802
15	0.189	0.354	0.446	0.521	0.604	0.654	0.700	0.750	0.779
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.729	0.762
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580
31	0.126	0.236	0.301	0.356	0.418	0.459	0.496	0.541	0.571
32	0.124	0.232	0.296	0.350	0.412	0.452	0.489	0.533	0.563
33	0.121	0.229	0.291	0.345	0.405	0.446	0.482	0.525	0.554
34	0.120	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35	0.118	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36	0.116	0.219	0.279	0.330	0.388	0.427	0.462	0.504	0.533
37	0.114	0.216	0.275	0.325	0.383	0.421	0.456	0.497	0.526
38	0.113	0.212	0.271	0.321	0.378	0.415	0.450	0.491	0.519
39	0.111	0.210	0.267	0.317	0.373	0.410	0.444	0.485	0.513
40	0.110	0.207	0.264	0.313	0.368	0.405	0.439	0.479	0.507
41	0.108	0.204	0.261	0.309	0.364	0.400	0.433	0.473	0.501
42	0.107	0.202	0.257	0.305	0.359	0.395	0.428	0.468	0.495
43	0.105	0.199	0.254	0.301	0.355	0.391	0.423	0.463	0.490
44	0.104	0.197	0.251	0.298	0.351	0.386	0.419	0.458	0.484
45	0.103	0.194	0.248	0.294	0.347	0.382	0.414	0.453	0.479
46	0.102	0.192	0.246	0.291	0.343	0.378	0.410	0.448	0.474
47	0.101	0.190	0.243	0.288	0.340	0.374	0.405	0.443	0.469
48	0.100	0.188	0.240	0.285	0.336	0.370	0.401	0.439	0.465
49	0.098	0.186	0.238	0.282	0.333	0.366	0.397	0.434	0.460
50	0.097	0.184	0.235	0.279	0.329	0.363	0.393	0.430	0.456