

Prueba de evaluación continua 1 Estadística Multivariante

Francesc Carmona y Miriam Mota

19 de noviembre de 2025

Caracterización inorgánica de las setas *Amanita ponderosa*

Amanita ponderosa son hongos silvestres comestibles que crecen en algunos microclimas de la Península Ibérica. Gastronómicamente esta especie es muy relevante, debido no sólo al consumo tradicional de las poblaciones rurales, sino también a su valor comercial en los mercados gourmet. La caracterización mineral de las setas comestibles es extremadamente importante para los procesos de certificación y comercialización. En el estudio de Cátia Salvador et al. (2018), se evaluó la composición inorgánica de los cuerpos frutales de *Amanita ponderosa* (Ca, K, Mg, Na, P, Ag, Al, Ba, Cr, Cu, Fe, Mn, Pb y Zn) y sus respectivos sustratos de suelo de 24 diferentes lugares de muestreo del suroeste de la Península Ibérica (por ejemplo, Alentejo, Andalucía y Extremadura). Con los siguientes ejercicios y la aplicación de algunas técnicas de Análisis Multivariante vamos a obtener algunos resultados del estudio.



Figura 1: Representación geográfica de las áreas de muestreo de las setas *A. ponderosa* en el área de Alentejo, Andalucía y Extremadura. En detalle, observamos el lugar concreto de la región de Évora (Alentejo, Portugal), con la vegetación circundante de *Quercus suber*, *Cistus ladanifer* y con una muestra del cuerpo fructífero recogido en este lugar.

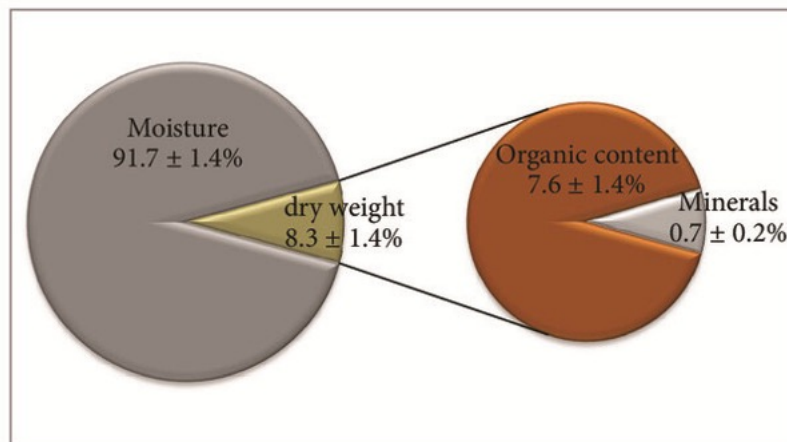


Figura 2: Moisture, dry weight, and organic and minerals content of *A. ponderosa* fruiting bodies from 24 different sampling sites.

Ejercicio 1 (30 pt.)

- (a) Obtener los datos a partir del archivo `amanita_ponderosa.RData`.

Reproducir los resultados (medias y desviaciones estándar) de la **tabla 2** del trabajo en estudio. La tabla se puede expresar en 6 columnas, aunque se valorará si se consigue en 3. No hace falta reproducir la tabla entera. Con mostrar las primeras filas será suficiente.

Obtener también los estadísticos de la figura 2. Nuestros valores serán ligeramente distintos.

Nota: Los resultados serán parecidos a los del estudio, aunque no idénticos. El motivo es que no tenemos los datos originales y los del archivo `amanita_ponderosa.RData` se han simulado.

- (b) En la misma tabla 2 se hace un ANOVA sin demasiadas explicaciones. Como ejercicio adicional se puede realizar un análisis de comparación múltiple con ayuda de la función `LSD.test()` del paquete `agricolae` para obtener grupos de “sites” (con las letras a, b, c,...). Como se indica en el texto del estudio, ¿las muestras de Almendres (1) y Serpa (20) son similares?
- (c) Indicar qué estadísticos robustos se pueden utilizar para describir los datos de una variable y poner un ejemplo con los datos de Moisture.
- (d) Dibujar el gráfico de la **figura 3** pero con un boxplot múltiple.

Aunque en general siempre se utilizan logaritmos neperianos, obsérvese que la escala del eje *Y* es en logaritmos decimales.

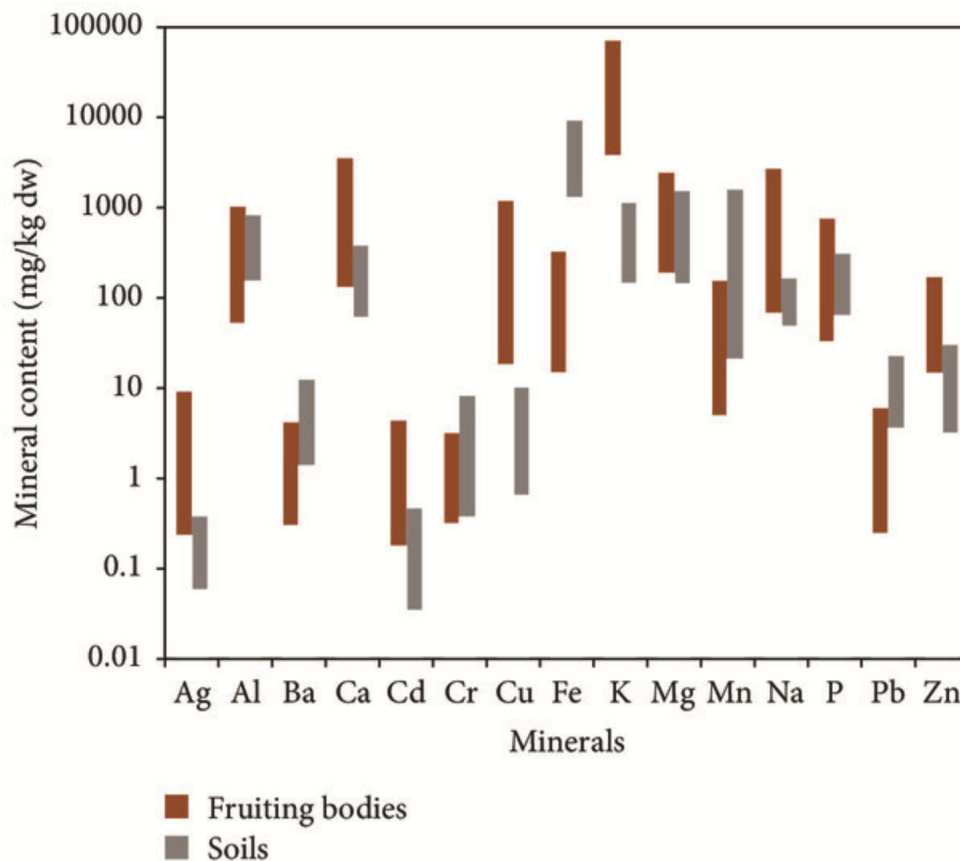


Figura 3: Concentration of mineral content of *A. ponderosa* fruiting bodies and soil substrates. The values are represented in logarithmic scale.

Ejercicio 2 (30 pt.)

En este ejercicio vamos a utilizar el análisis de componentes principales o PCA. Debemos obtener las componentes principales de los datos de composición de minerales de las tablas 4, 5 y 6 del estudio. Estos datos se hallan en los `data.frame` `FB.df` y `SS.df`.

- Realizar el PCA para los datos medios de la composición mineral de los *fruiting bodies* (FB) de cada *site*.
Valorar el resultado y comentar el gráfico de dispersión con las dos primeras componentes.
¿Qué ocurre en Azaruja?
- Realizar el PCA para los datos medios de la composición mineral del *soil substrate* (SS) de cada *site*.
Valorar el resultado y comentar el gráfico de dispersión con las dos primeras componentes.
¿Qué ocurre en Azaruja? ¿Y en Serpa?
- En el estudio se utiliza el factor de bioconcentración (BCF) o ratio del contenido en el cuerpo fructífero respecto al contenido del substrato respectivo para evaluar la capacidad de acumulación de los hongos.

Utilizaremos la ratio “FB/SS” de las medias para todos los minerales evaluados en cada lugar, es decir, el contenido relativo de la seta respecto al suelo.

- Si $\text{ratio} > 1$, la seta acumula más que el suelo.
- Si $\text{ratio} < 1$, el elemento está menos concentrado en la seta que en el suelo.

Al tomar logaritmos hacemos que los dos items sean:

- Si $\log(\text{ratio}) > 0$, la seta acumula más que el suelo.
- Si $\log(\text{ratio}) < 0$, el elemento está menos concentrado en la seta que en el suelo.

Realizar el PCA para los logaritmos de las ratios de cada **site**.

Valorar el resultado y comentar el gráfico de dispersión con las dos primeras componentes.

¿Qué ocurre en Azaruja? ¿Y en Serpa?

Ejercicio 3 (20 pt.)

- (a) Realizar un análisis de proximidades o MDS con los datos de la composición mineral de los FB del estudio. Para ello habrá que utilizar las medias de cada lugar de muestreo (**site**).

Como se trata de una composición utilizaremos la distancia de Bray-Curtis que no es euclídea.

- (b) Dibujar un gráfico del análisis en dos dimensiones donde veamos la región a la que pertenece cada lugar de muestreo.

Valorar el resultado e identificar los lugares más extremos.

Ejercicio 4 (20 pt.)

Realizar un análisis de correspondencias (CA) sobre los datos de la composición en peso fresco de los FB (tabla 2). También aquí se utilizan las medias de composición de cada lugar de muestreo para las tres componentes: Moisture, Organic y Minerals.

Como los valores calculados son medias, se puede comprobar que las filas no suman el 100 % y, por ello, deberemos ajustar los valores para que sean proporciones de suma 100.

Referencias

- [1] Cátia Salvador et al. (2018), A Data Mining Approach to Improve Inorganic Characterization of *Amanita ponderosa* Mushrooms. *Int J Anal Chem.*, Jan 31; doi: 10.1155/2018/5265291