



## RESEÑA DE EVENTOS CIENTÍFICOS

## Curso teórico-práctico: “Estadística para la investigación en Biología” Statistics for biological research workshop

Orlando Advíncula Zeballos

Universidad Nacional Mayor de San Marcos  
Universidad Nacional Agraria La Molina

Correo: [orlando\\_ad@outlook.com](mailto:orlando_ad@outlook.com)

Recibido 15 diciembre 2015, Aceptado 6 enero 2016.

DOI: <https://dx.doi.org/10.6084/m9.figshare.2380939.v3>

© 2016 Todos los derechos reservados

**Resumen**

Durante el mes de noviembre del 2015, el Instituto de Genética Barbara McClintock organizó el curso teórico-práctico de “Estadística para la investigación en Biología”, con el objetivo de proporcionar a los participantes las herramientas básicas de la investigación científica. La temática desarrollada abarcó temas como la metodología de la investigación científica, la inferencia estadística, diseños experimentales y el análisis multivariado. En cada parte del curso se revisó el marco teórico de manera profunda, para luego resolver problemas a las aplicaciones en Ciencias Biológicas. Se usaron los software Infostat y Primer-E.

**Palabras clave:** estadística, biología, investigación.**Abstract**

On November 2015, the Institute of Genetics Barbara McClintock organized “Statistics for biological research” workshop, with the aim of providing participants the basic tools of scientific research. The themes developed included the methodology of scientific research, inference statistical inference, experimental designs and multivariate analysis. In each part of the course the theoretical framework deeply, and then solve problems related to applications in Biological Sciences was reviewed. Softwares Infostat and Primer-E were used.

**Key words:** statistics, biology, research.**Método Científico**

En la primera parte del curso, se desarrollaron los conceptos básicos de la investigación científica; el método científico, su importancia y trascendencia. Es necesario tener en cuenta el objetivo general de la investigación, así como los específicos, pues en base a ellos se escogerá la metodología. Por lo que debemos tener bien claro las condiciones respecto a la eficiencia y estandarización.

Esto implica que el trabajo será bien planificado antes de su ejecución. En esta etapa el método seleccionado deberá responder al objetivo, teniendo en cuenta las limitaciones, como éstas serán manejadas, y si el tamaño de muestra será el requerido para la evaluación (1).

Otro aspecto importante a la hora de realizar la investigación es evitar la pseudoréplica; que es un error bastante común (1,2), esto se da cuando las réplicas no son independientes entre sí. Por lo que es muy importante tener bien claras las nociones del diseño experimental. Estos casos fueron profundamente analizados con diversos ejemplos.

## **Inferencia Estadística**

Se analizó los principios de la inferencia estadística, este es el procedimiento inductivo mediante el cual a partir de las observaciones realizadas, se describen las diversas características de la distribución bajo estudio, lo cual es la base para el modelamiento (3). Se explicó que para los casos en los cuales las variables poseen distribución normal, se usarán las pruebas paramétricas, de no haber normalidad se usará pruebas No Paramétricas.

Para el curso se analizaron ejemplos para dos muestras independientes, analizando sus diversas variantes. Se hizo gran énfasis en las pruebas de hipótesis.

## **Diseños Experimentales**

Se analizaron los principios básicos del diseño experimental. Conceptos como la naturaleza de las unidades experimentales, número y clase de tratamientos (incluyendo controles), asignación de tratamientos a unidades experimentales, número de unidades experimentales (réplicas) que recibirán los tratamientos, arreglo espacial de las unidades experimentales y la secuencia temporal de la aplicación de tratamientos (1).

El primer diseño experimental analizado fue el DCA (Diseño Completo al Azar), el cual tiene una fuente de variabilidad (factor), que a su vez tiene sus niveles. Este modelo también se le denomina ANOVA de una vía.

Para la aplicación de este modelo, se tuvo que analizar los supuestos básicos como la distribución normal de los residuales y la homogeneidad de varianzas. En los casos que se encuentren diferencias entre los niveles del tratamiento, se analizaron las diversas pruebas de comparación. Pruebas como LSD, Dunnet y contrastes fueron usadas para este caso.

El segundo diseño experimental analizado fue el DBCA (Diseño de Bloques Completo al Azar), el cual tiene dos fuentes de variabilidad, uno de ellos es el tratamiento (factor) y la otra es aquella que se desea controlar, mediante el bloque (4). A este diseño también se le denomina ANOVA de dos vías. El cual reduce la probabilidad de un efecto influyente pero no detectado de gradientes presentes y de eventos azarosos (1). Debe cumplir con los mismos supuestos del DCA.

El tercer diseño experimental analizado fue el Experimento Factorial, el cual tiene dos fuentes de variabilidad, busca analizar las interacciones entre los diversos niveles de ambos factores (las mejores combinaciones). Para el curso se estudiaron modelos con dos factores.

## **Análisis Multivariado**

Esta herramienta es muy usada para las investigaciones en Ecología y medio ambiente. El análisis multivariado permite analizar gran cantidad de información. Entre los métodos analíticos de ordenación y clasificación, permiten identificar similitud entre entidades: sitios de muestreo, comunidades, dietas, etc (1). La gran ventaja en el presente es que los softwares pueden ejecutar de manera muy rápida estos procedimientos, lo cual no era posible antiguamente.

Entre los métodos estudiados se empezó por el Análisis de Componentes Principales (PCA), el cual permite reducción de variables en dos componentes principales.

El segundo tema desarrollado fue el de Semejanza Ecológica, ideal para poder comparar diversos puntos de muestreo. Para este procedimiento se puede usar varios tipos de distancias: Bray-Curtis es el más usado en Ecología y Biología Marina, otro método usado es el Jaccard cuando se quiere analizar presencia/ausencia.

El Análisis Cluster permite realizar clasificación, por ejemplo entre zonas de muestreo. Es muy útil para poder encontrar patrones de agrupación en los ecosistemas, incluso puede incorporar factores abióticos en el Cluster para poder ver una clasificación más clara.

Finalmente, se usaron las pruebas no paramétricas ANOSIM (*Analysis of Similarity*) (5) para comparar dos o más lugares en torno a un factor, que podría ser por ejemplo la salinidad, profundidad, etc. De encontrar diferencias significativas podría usarse la prueba SIMPER (Similarity Percentages) (5), para poder analizar cuales son las especies que caracterizan a una determinada zona.

## Agradecimientos

El autor quiere agradecer al Instituto de Genética Bárbara McClintock, por la confianza en la elaboración del curso Estadística para la investigación en Biología.

## Referencias Bibliográficas

1. Galindo-Leal C. Diseño y Análisis de Proyectos para el Manejo y Monitoreo de la Diversidad Biológica. 2003; 2003:1–176.
2. Programa de Investigación Tropical del Centro para la Biología de la Conservación. El diseño de Proyectos de Conservación. Ecotono. 1996; 1996:12.
3. Di Rienzo J, Casanoves F, González L, Tablada E, Díaz M, Robledo C, et al. Estadística para las ciencias agropecuarias. 2005; 2005:372.
4. Balzarini MG, Gonzalez LA, Tablada EM, Casanoves F, Di Rienzo JA, Robledo CW. InfoStat, versión 2008. Córdoba, Argentina.; 2008.
5. Clarke KR. Non-parametric multivariate analyses of changes in community structure. Aust J Ecol. 1993; 18:117-142.

