



UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO
FACULTAD DE BIOLOGÍA

PROGRAMA EN EXTENSO DE LA MATERIA DE DISEÑO Y ANÁLISIS EXPERIMENTAL

Semestre en el que se imparte: Quinto

Carga horaria. 4 horas semanales

Créditos SATCA: 4

Área Académica: Física y Matemáticas

INTRODUCCIÓN

En la Biología, la mayor parte de los fenómenos están afectados por factores casuales que son incontrolables en cuanto a su variación. Por ello, se necesita de la Estadística para medir estos fenómenos variables con un error predecible y para detectar importantes diferencias entre los mismos.

Muchos problemas biológicos exigen la planeación de un experimento, el cual incluye, entre otros aspectos, el diseño muestral aleatorio, la definición de las variables aleatorias relacionadas con el problema científico a resolver, el análisis y ejecución del diseño experimental más adecuado, y el análisis descriptivo e inferencial de los datos.

Después de estudiar en Matemáticas I y Matemáticas II algunos de los elementos mencionados, resta prestar atención al diseño experimental y sus técnicas estadísticas inferenciales correspondientes, con énfasis en el análisis de varianza, así como introducir algunos experimentos factoriales y técnicas de análisis multivariado que en la actualidad son muy usadas por los biólogos.

Sin lugar a dudas, aquí el uso de un software estadístico es más que importante, es indispensable. La interrelación efectiva entre la Estadística, la Computación y la Biología, contribuirá al cumplimiento de los objetivos y al logro de las habilidades previstas en el perfil del profesional.

OBJETIVOS

GENERAL

EDUCATIVOS

Contribuir a que los estudiantes:

- Consoliden la concepción científica del mundo, mostrando como los métodos estadísticos están vinculados a la práctica y la vida social del hombre, en particular al ejercicio de su profesión.
- Adquieran la convicción de que las técnicas de computación y de la Estadística sirven de base a otras asignaturas de la carrera y constituyen medios de gran utilidad para el desarrollo de su actividad profesional y preprofesional.
- Desarrollen habilidades y hábitos de proceder reflexivamente al transmitir sus ideas con precisión y exactitud y evaluar críticamente los resultados de su trabajo y así garantizar la científicidad en sus análisis.
- Desarrollar cualidades de la personalidad como la constancia, la voluntad y la capacidad de trabajo independiente.

INSTRUCTIVOS

- Describir los conceptos fundamentales del análisis de varianza y el diseño experimental.
- Aplicar los métodos y procedimientos estadísticos de planeación del diseño experimental y de análisis inferencial de pruebas de hipótesis, incluyendo algunos multivariados, a la solución de fenómenos biológicos y sociales, utilizando para ello la computación.
- Interpretar la esencia de los fenómenos biológicos mediante la aplicación de los métodos estadísticos mencionados, con el apoyo de la computación.
- Comprender la literatura publicada sobre investigaciones biológicas que involucran diseño experimental o análisis estadísticos de datos.
- Utilizar software estadísticos profesionales en el procesamiento de la información.

PARTICULARES

- Describir los conceptos fundamentales del diseño experimental, así como las formas de control del error y los modelos de los diseños experimentales más usados.
- Describir las premisas del análisis de varianza y justificar el uso de las transformaciones de escala en el caso del incumplimiento de algunas de estas premisas.

- Identificar el diseño experimental y la técnica inferencial que mejor resuelve el problema objeto de interés.
- Aplicar las técnicas inferenciales de comparación de promedios de más de dos grupos para una variable o factores o para varios factores (ANOVA y ANOVA factorial) a la solución de fenómenos biológicos o sociales, apoyándose en la computación.
- Aplicar la Chi-cuadrada para demostrar la independencia, homogeneidad entre dos variables o grupos y para probar normalidad en la población, usando la computación.
- Aplicar la regresión y correlación lineal simple a si como la múltiple.
- Determinar el modelo matemático lineal que relacione dos variables.
- Exponer los métodos de regresión y correlación lineal simple a un procedimiento para predecir una variable y de respuesta con base a los valores de dos o más variables independientes.
- Aplicar la ji-cuadrada para las pruebas de bondad de ajuste así, como para demostrar independencia y homogeneidad entre dos variables o grupos.
- Diferenciar los métodos paramétricos de los no paramétricos para su aplicación en problemas biológicos.
- Interpretar los resultados en los términos del problema biológico resuelto una vez de haber aplicado los métodos estadísticos estudiados.
- Utilizar un paquete estadístico profesional para el procesamiento de la información y ejecución de los métodos estadísticos objeto de estudio.
- Diseñar un experimento y procesar la información mediante los métodos estadísticos descriptivos o inferenciales tratados, resolviendo un problema biológico relacionado en su vida profesional, apoyándose en un software estadístico profesional.

CONTENIDOS

SISTEMA DE CONOCIMIENTOS

CAPITULO I

DISEÑOS EXPERIMENTALES

1.1 Introducción.

1.2 Conceptos básicos del diseño experimental.

1.3 Análisis de varianza.

1.3.1 Diseño completamente aleatorizado.

- 1.3.2 Diseño de bloques al azar.
- 1.3.2 Diseño de cuadrado latino
- 1.3.3 Experimentos factoriales.
- 1.3.4 Diseño de parcelas divididas
- 1.3.5 Diseño jerárquico

CAPITULO II REGRESIÓN Y CORRELACIÓN LINEAL SIMPLE

- 2.1 Introducción
- 2.2 Modelo de regresión lineal simple
- 2.3 Propiedades de los estimadores de mínimos cuadrados.
- 2.4 Inferencias acerca de los coeficientes de regresión.
- 2.5 Evaluación de la ecuación de regresión lineal.
- 2.6 Uso de la ecuación de regresión lineal.
- 2.7 Coeficiente de correlación.
- 2.8 Coeficiente de determinación.

CAPITULO III REGRESIÓN Y CORRELACIÓN MÚLTIPLE

- 3.1 Introducción
- 3.2 El modelo de regresión múltiple
- 3.3 Propiedades de los estimadores de mínimos cuadrados.
- 3.4 Inferencias acerca de los coeficientes de regresión múltiple.
- 3.5 Evaluación de la ecuación de regresión múltiple.
- 3.6 Uso de la ecuación de regresión múltiple.
- 3.7 Coeficiente de correlación.
- 3.8 Coeficiente de determinación.

CAPÍTULO IV DISTRIBUCIÓN JI-CUADRADA

- 6.1 Introducción.
- 6.2 Propiedades matemáticas sobre la distribución ji-cuadrada.
- 6.3 Prueba de bondad de ajuste.
- 6.4 Prueba de independencia.
- 6.5 Prueba de homogeneidad.

CAPÍTULO V ESTADÍSTICA NO PARAMÉTRICA

- 5.1 Introducción.
- 5.2 Prueba de U Mann-Whitney.
- 5.3 Prueba de signos para comparaciones pareadas.
- 5.4 Prueba Wilcoxon.
- 5.5 Prueba de bondad de ajuste de Kolmogorov-Smirnov.
- 5.6 Análisis unilateral de la varianza por jerarquías de Kruskal-Wallis.
- 5.7 Análisis bilateral de la varianza por jerarquías de Friedman.
- 5.8 Coeficiente de correlación por rangos de Spearman.
- 5.9 Análisis de regresión no paramétrico.

SISTEMA DE HABILIDADES

- Aplicar las transformaciones de escala correspondiente a diferentes casos y probar su efectividad.
- Identificar el tipo de diseño experimental que mejor se ajuste a las diferentes situaciones que aparecen en los ejercicios.
- Verificar las premisas del análisis de varianza y del resto de los análisis multivariados.
- Desarrollar el análisis de varianza a partir de los datos que se deriven de los diferentes diseños experimentales estadísticos.
- Interpretar los resultados obtenidos a partir de las técnicas inferenciales aplicadas en términos del fenómeno biológico y de acuerdo con los objetivos que se persigan.
- Utilizar las ventajas y desventajas de cada diseño experimental para lograr identificar aquel más idóneo.
- Desarrollar los métodos de comparación múltiple de las medias cuando es significativo el ANOVA.
- Diferenciar los análisis de varianza de clasificación doble con réplicas (experimentos factoriales) de los que no tienen réplicas.
- Analizar los efectos de las interacciones en los diseños factoriales.
- Determinar cuándo el análisis de regresión es la herramienta estadística apropiada para analizar un problema biológico.
- Aplicar la ji-cuadrada de independencia en demostrar la independencia y homogeneidad entre dos variables o grupos, así como probar normalidad de la población.
- Diferenciar los métodos paramétricos de los no paramétricos de comparación de medias de varios grupos con una variable independiente y una dependiente.
- Calcular e interpretar el grado de asociación, así como la mejor ecuación que ajusta la relación entre varias variables independientes a nivel reproductivo y productivo con el apoyo de la computación.
- Analizar el grado de asociación entre dos variables a partir del cálculo del coeficiente de correlación de Spearman.

- Identificar y aplicar, usando la computación, de las técnicas estadísticas estudiadas, la más adecuada para darle solución a la problemática objeto de estudio.
- Utilizar paquetes estadísticos profesionales.
- Planear, desarrollar e informar sobre una investigación experimental.

CORRELACION CON OTRAS MATERIAS

Las asignaturas que más requieren de los conocimientos de la Estadística son la Taxonomía, recursos naturales y la Ecología. Sin embargo, cualquier materia que incluya en su contenido fenómenos que permitan realizar un experimento, donde las variables ha analizar se puedan medir en una escala nominal, ordinal, por intervalos o de razones o proporciones, la Estadística tiene relación con estas disciplinas.

METODOLOGÍA

Se impartirá la asignatura en el salón de clases y en el laboratorio de computación con un enfoque de carácter teórico-práctico, profundizando en las aplicaciones a ejemplos que tengan que ver con las ciencias biológicas, así como en la interpretación de los resultados.

Los alumnos desde el inicio del curso deberán conocer el programa, la forma de trabajo y la manera en que será evaluado y la bibliografía fundamental que utilizará.

A partir de un diagnóstico inicial, se sistematizará lo conocido por los alumnos del análisis descriptivo de datos, teoría del muestreo y análisis inferencial univariado y bivariado, para que ante nuevas interrogantes científicas relacionadas con fenómenos biológicos, que sirvan de motivación intrínseca y extrínseca, se introduzca la necesidad de estudiar los métodos de planeación de experimentos y otros métodos estadísticos inferenciales, que serán objeto de estudio. De esta manera el alumno tendrá una imagen de los contenidos nuevos a asimilar y se sentirá interesado por conocerlos. Sería muy elocuente que los ejemplos introductorios se escogieran de los que los estudiantes defendieron en el informe final de investigación de la asignatura Matemática II, en los cuales debe haberse esbozado interrogantes e hipótesis sólo solubles con los métodos que en esta asignatura se impartirán.

Los estudiantes deben desarrollar los ejercicios haciendo uso de calculadoras de bolsillo y corriendo programas de computadora contemplados en diferentes paquetes estadísticos.

Debe quedar claro que el análisis de varianza (ANOVA) es una técnica inferencial para comparar medias de varias poblaciones a través de los valores obtenidos de varias muestras independientes, a pesar de que se nombra análisis de varianza. Esta frase se debe a que este método se apoya en la descomposición de la varianza de todas las observaciones en varias partes, midiendo cada parte la variabilidad atribuible a alguna causa determinada, por ejemplo, a la variación interna en las diversas poblaciones, a variaciones de una población a otra, etc.

Tiempo debe dedicarse a la comprensión de las premisas que debe cumplir una ANOVA: Aleatorización, independencia de errores experimentales, homogeneidad de varianzas y normalidad, todo esto permitirá sistematizar conocimientos ya impartidos en Matemáticas I y II. Para probar la homogeneidad de varianzas se utilizará el test de Bartlett. También es importante que se dominen las posibles alternativas para cuando las premisas no se cumplen, y en este sentido es que se introducirán algunas transformaciones de escala: logarítmica, raíz cuadrada, angular y recíproca. Se sugiere aquí orientar un trabajo extraclase donde el alumno transforme la escala de un conjunto de datos y pruebe mediante el análisis descriptivo e inferencial, la hipótesis de normalidad.

Finalmente, el método estadístico como tal de análisis de varianzas se explicará en el diseño completamente aleatorizado.

Cuando se desea hacer una investigación que involucra el tratamiento estadístico de los datos que se obtengan, debe planearse con cuidado cómo se van a obtener éstos. Es decir, resulta necesario, antes de obtener los datos, contar con un buen diseño del muestreo y del experimento a realizar. En particular, la forma del muestreo o del diseño experimental que se utilice determina el tratamiento estadístico que se puede utilizar para el análisis de datos. La aplicación de los métodos de diseño experimental permite extraer la máxima información, en condiciones de un riguroso control o con un costo mínimo.

Se debe señalar que en la planeación de un trabajo de investigación intervienen cuestiones de tipo metodológico que todos no serán tratados en este curso, pero sí se insistirá en la necesidad de transformar el problema objeto de estudio en análisis de relaciones entre variables, así como en la forma de medir estas y de controlar algunas que intervienen (variables ajenas) y que pueden provocar confusión, con sus efectos sobre la variable de interés. Se introducirá la variable independiente como aquella que provoca un efecto en otra variable que es la dependiente.

Es necesario que a través de un debate se analicen las vías para controlar el error experimental, que es una medida de la variabilidad que existe entre las observaciones en las unidades experimentales.

A partir de ejemplos ya estudiados se introducirán los diseños clásicos de dos grupos y un grupo para una variable dependiente contra una independiente. Aquí en la ejercitación se sistematizará el análisis descriptivo e inferencial univariado y bivariado. A partir de este momento, siguiendo la lógica de situar problemas de las ciencias biológicas que para ser resueltos exijan el uso de determinado diseño experimental se introducirán el diseño completamente aleatorio, el diseño de bloques al azar, diseño en cuadrado latino, los diseños factoriales, diseño de parcelas divididas y diseño jerárquico. Siempre se debatirán las ventajas y desventajas de cada uno de ellos comparativamente, para que el alumno llegue a asimilar cuándo debe usar uno y no otro.

Siempre que el análisis de varianzas conduce a rechazar la hipótesis nula de no diferencia entre las medias de las poblaciones, surge la pregunta respecto a qué par de medias son diferentes. El test t dado en Matemáticas II no es válido para examinar las diferencias entre todos los pares de medias posibles, para ello se estudiarán los test de rangos múltiples: Duncan, LAD, SNK y Tukey.

Los cálculos pueden hacerse primero con calculadora de bolsillo con un tamaño de cada muestra pequeño, suponiendo el cumplimiento de las premisas, y después utilizando un paquete estadístico profesional.

Para el tema de análisis de varianza ANOVA de clasificación simple y ANOVA de clasificación doble después de exponer los elementos teóricos y de haber hecho ejercicios en el aula, en el laboratorio, se resolverán problemas con el uso de la computación, haciendo los cálculos requeridos e interpretando los resultados. El tiempo que ahorra la computación se debe utilizar para analizar el problema, cuál es el diseño de experimento que mejor lo describe, qué ANOVA debe ser realizado y para interpretar y verificar los resultados en los términos del fenómeno biológico o social objeto de estudio.

Para el ANOVA factorial, tiene que quedar claro que los diseños factoriales pueden construirse a partir de los datos obtenidos de diseños completamente aleatorizados, diseños de bloques al azar o diseño cuadrado latino, pero añadiéndoles a estos la característica esencial que los identifica, a saber, en un experimento factorial tienen que existir datos replicados por casilla, es decir se toma en cuenta la interacción fila x columna, o con mayor rigor científico, en los diseños factoriales no sólo es posible estudiar los efectos de dos o más factores (variables) individuales, sino también, la interacción entre dichos factores, o sea la dependencia del efecto de un factor en el nivel de otro factor.

A partir de lo conocido de regresión lineal simple y a través de un ejemplo relacionado con las ciencias biológicas se introducirá el análisis de regresión múltiple. Por ejemplo, supongamos que se desea establecer relación entre la densidad de jútías en una región determinada con respecto a múltiples variables independientes como puede ser la altura de los árboles, la cobertura, el índice de heterogeneidad, el por ciento de rocosidad, etc., con fines de prevenir bajos niveles de densidad y tomar a tiempo medidas de conservación. Se debe buscar entonces una dependencia de la forma: $Y = f(x_1, x_2, \dots, x_n)$ donde Y es la densidad de jútías en dicha región y x_1, x_2, \dots, x_n son las variables independientes mencionadas.

El ejemplo evidencia dos dificultades esenciales: la de no conocer a priori la forma de la función buscada y la de existir posibles correlaciones entre las variables independientes que no hagan tan claro cuáles variables son las que deben intervenir definitivamente en la ecuación y la posibilidad de haber más de una ecuación entre las cuales se debe escoger la que más apropiadamente satisfaga nuestras pretensiones. El análisis de regresión múltiple, y dentro de este, el método *stepwise* posibilitará la solución de la problemática presentada.

En la prueba de Chi-cuadrada de independencia se hará énfasis en su posibilidad de aplicación para las pruebas de bondad de ajuste, independencia de dos variables o grupos y homogeneidad o diferencia entre 2 grupos. Cuando las variables de interés estén incluso en escala nominal.

Para el tema de estadística no paramétrica se debe hacer primero una introducción donde se marquen las ventajas y desventajas en comparación con las pruebas paramétricas, con el fin de que el alumno valore la utilización de las no paramétricas sólo en los casos necesarios.

Se deben exponer las pruebas no paramétricas correspondientes a las paramétricas vistas anteriormente. Una vez expuesto el sustento teórico y visto ejercicios con pocos datos en el salón de clases se debe usar el paquete Statistica; para las pruebas no paramétricas, con el número de datos deseado.

Se insistirá en que el alumno entienda cuándo debe decidirse por una prueba paramétrica y cuándo debe aplicar una no paramétrica, valorando las fortalezas de las primeras frente a las últimas pero reconociendo que cuando las variables son cualitativas y no siguen una distribución normal, así como tamaño de muestras pequeñas existe para cada prueba paramétrica la no paramétrica correspondiente.

Se le pide al alumno que aplique la prueba no paramétrica equivalente a la prueba paramétrica en su investigación. Si el estudiante consideró variables que se presten para hacer pruebas no paramétricas podrá aplicar éstas.

También se estudiarán las alternativas no paramétricas para cuando los datos disponibles para el análisis no satisfacen las suposiciones del ANOVA de clasificación simple y del ANOVA de clasificación doble sin replicación, nos referimos al test de Kruskal Wallis y el análisis de Friedman, respectivamente. Se debe destacar, además, que el primero es el caso correspondiente al test U de Mann Whitney, pero ahora para comparar medias de varios grupos independientes y Friedman se corresponde con el test de Wilcoxon, lo que en función de comparar en tendencia central varios grupos o muestras relacionadas.

A los alumnos desde el inicio del curso hay que exigirles que a partir de las prácticas de campo comiencen a diseñar un experimento que vincule varias variables, ejecuten el muestreo, el análisis descriptivo y el análisis inferencial más adecuado, con el auxilio de los paquetes estadísticos profesionales que se han utilizado. Con estos resultados hará un informe de investigación que presentará al final del curso como parte de su evaluación.

En lo que respecta a la bibliografía, existe poca bibliografía en español, recomendamos los cuatro libros que se citan como bibliografía básica.

Los paquetes estadístico propuesto para este curso es el Statistic o el SPSS.

EVALUACIÓN

La evaluación estará centrada básicamente en los tres siguientes aspectos: evaluación frecuente, tareas, pruebas parciales y trabajo de investigación.

La evaluación frecuente, se cuantificará por la participación del estudiante en el salón de clases, en el laboratorio y por el cumplimiento de las actividades planeadas por el profesor.

Las tareas, consisten en una serie de ejercicios que el alumno debe resolver cada vez que se termine un capítulo y entregarlos por escrito al profesor para su revisión.

Pruebas parciales. Se sugiere realizarán tres exámenes parciales, el primero al terminar la primera unidad. El segundo examen al terminar el tercer capítulo. El último examen al terminar el capítulo V. Se sugiere que para todos los exámenes parciales se les permita a los estudiantes utilizar la computadora.

Trabajo de investigación. Se sugiere se aplique al final del curso, para medir el manejo del paquete estadístico haciendo énfasis en la interpretación.

Se recomienda que el trabajo de investigación junto con el manejo del paquete estadístico se le de el mismo valor porcentual que al promedio de los exámenes parciales.

Es fundamental que al alumno le quede claro cada uno de los instrumentos que van a servir para ser evaluado. Es importante que el profesor dedique tiempo suficiente para explicar cada una de las partes del informe del trabajo de investigación.

BIBLIOGRAFIA

- Cochran William G. y Gertrude M. Cox. "Diseño experimental". Trillas. México. 1997.
- Cuadras C. M. "Métodos de Análisis Multivariante " , Editorial Universitaria de Barcelona.
- Flury B. " A first Course in Multivariate Statistics" , Springer.
- Kreyszig Erwin. "Introducción a la Estadística Matemática Principios y métodos". Limusa. México. 1994.
- Lipschutz Seymour. "Probabilidad". MacGraw-Hill. México. 1971.
- Montgomery, Douglas C. "Diseño y Análisis de Experimentos", Grupo editorial iberoamericana.
- Ostle Bernard. "Estadística Aplicada". Limusa. México. 1996.
- Peña, Daniel. "Análisis de datos multivariantes", Mc Graw Hill.
- Reyes Pedro. "Bioestadística". Limusa. México.
- Scheffler William C. "Bioestadística". Iberoamérica. México. 1981.
- Snedcor W. George y Cochran G. William. "Métodos Estadísticos". Cecsa. México. 1979.
- Sokal R. y Rohlf F. J. "Biometría Principios y Métodos Estadísticos en la investigación biológica". H Blume ediciones. España. 1979.

- Steel G. D. Robert y James H. Torrie. Bioestadística Principios y procedimientos. McGraw-Hill. México, 1996.
- Spiegel Murray R.." Estadística Teoría y Problemas Resueltos". McGraw-Hill. México. 1979. 357 págs.
- Ter Braak C J. F. "Data analysis in Community". Cambridge University Press.
- UPN. "Introducción a los métodos estadísticos" tomo 2 y 3, Sep. 1987.
- Wayne W. Daniel. "Bioestadística base para el análisis de las ciencias de la salud". Limusa. México. 1997.
- Wayne W. Daniel. " Estadística con Aplicaciones a las ciencias Sociales y la Educación". Limusa. México. 1986.