**1. INFORMACIÓN GENERAL**

|  |  |
| --- | --- |
| **PROGRAMA** | **Ciencia de Datos** |
| **ÁREA** | **Matemáticas** |
| **ASIGNATURA** | **Inferencia Estadística 1** |
| **CRÉDITOS** |  |
| **SEMESTRE** |  |
| **HORAS PRESENCIALES** | **96 horas semestrales** |
| **HORAS DE TRABAJO AUTÓNOMO** | **96 horas semestrales** |
| **PROFESOR** | **Pedro Agustín Monterrey Gutiérrez** |
| **CORREO** | **pedro.monterrey@uexternado.edu.co** |

**2. PRESENTACIÓN**

En **Inferencia Estadística 1** los futuros profesionales de la Universidad Externado de Colombia desarrollarán las competencias cognitivas y comunicativas necesarias para pensar probabilísticamente y analizar datos utilizando métodos de perfil probabilístico; identificando el papel de la incertidumbre, es decir de la variación muestral, en los procesos de análisis. Para lograr esto, el curso introducirá las nociones fundamentales de la teoría de probabilidades y las utilizará para presentar los métodos básicos de la inferencia estadística y del análisis de datos exploratorio.

**3. COMPETENCIAS**

1. **Competencia cognitiva**: Capacidad para construir y sustentar el saber de un cuerpo disciplinario específico y la apropiación del conocimiento especializado.

* Formular argumentos usando las nociones básicas de la teoría de probabilidades.
* Identificar el papel del azar y su reflejo en la variación muestral en los métodos estadísticos para el análisis de datos.
* Establecer los alcances del conocimiento estadístico.

1. **Competencia comunicativa**: entendida como la capacidad para interactuar por medio del lenguaje, ubicarse en situaciones específicas y extender a contextos amplios los procesos de comprensión y producción de discursos diversos. Las habilidades de “hablar, escuchar, leer y escribir” se fundamentan en esta competencia.

* Comprender la modelación probabilística como la base de los métodos del análisis estadístico de los datos.
* Reconocer la estadística como una ciencia para aprender de los datos, cuya función es medir, controlar y comunicar la incertidumbre.

**4. RESULTADOS DE APRENDIZAJE**

1. **Resultados de la competencia cognitiva**:

* Identificar los diferentes enfoques (frecuentista y subjetivo) para definir la probabilidad
* Estructurar los componentes de los modelos probabilísticos para caracterizar el comportamiento de las mediciones en las poblaciones
* Identificar patrones de comportamiento de los datos y los modelos probabilísticos sobre los que este comportamiento se sustenta
* Estructurar los diferentes procedimientos del análisis de datos según las funciones exploratoria o confirmativa

1. **Resultados de la competencia comunicativa**:

* Construir y aplicar modelos probabilísticos sobre los que se sustentan los procedimientos de la inferencia estadística.
* Identificar alcances y limitaciones de los diferentes procedimientos del análisis de los datos.

**5. TEMÁTICAS O CONTENIDOS**

1. Probabilidad. Probabilidad Condicional e Independencia. El teorema de Bayes.
2. Variables Aleatorias, Esperanzas, Momentos, Función de Distribución. Ley de los Grandes Números. Teorema del Límite Central.
3. Distribuciones Discretas Univariadas. Distribuciones Continuas Univariadas. Distribuciones Multivariadas.
4. Población, muestra. Criterios de muestreo. Introducción al análisis de datos exploratorio. Las distribuciones muestrales
5. Inferencia Paramétrica: Verosimilitud e Información. Estimación Puntual. Comparación de Estimadores. Intervalos de Confianza. Pruebas de hipótesis.

**6. METODOLOGÍA**

En un mundo marcado por el cambio y la incertidumbre, los conocimientos se desactualizan aceleradamente, la formación académica tiene que fomentar la apropiación y desarrollo de competencias para la formación permanente, condición de la adaptabilidad y eficacia de los profesionales e investigadores del mañana.

Por ello, los programas de la Facultad sientan sus bases en un modelo pedagógico y académico que, por una parte, privilegia la participación de los estudiantes en la construcción de su propio saber, y por otra, enseña una metodología para la construcción del razonamiento. La metodología de la Facultad es eminentemente participativa y enseña a los estudiantes a dominar los recursos de la lógica y del razonamiento para desarrollar su capacidad de análisis crítico y de síntesis, a mejorar su efectividad argumentativa y comunicacional, y a practicar la indagación sistemática como recurso fundamental de la investigación.

Por esto las clases se realizarán en forma sincrónica, con el complemento de las tecnologías correspondientes que se trabajaran en forma asincrónica para la apropiación de los temas, espacios donde el estudiante contaran con apoyo del docente.

El Curso será orientado a la solución de problemas; los computadores serán utilizados intensivamente en las clases.

Las clases tendrán los siguientes momentos:

* Contextualización de los temas que se requiere para desarrollar los contenidos y objetivos pertinentes a la clase.
* Breve exposición de los conceptos, fórmulas, algoritmos por parte del docente; los estudiantes participaran activamente proporcionando ejemplos para aplicar los temas correspondientes.
* Al finalizar los estudiantes responderán inquietudes relacionadas con las aplicaciones y describirán los temas relevantes de la clase.

Talleres y tareas: corresponden al trabajo realizado por los estudiantes de forma individual y en grupos, sincrónico o asincrónico.

**7. EVALUACIÓN**

La evaluación es un proceso continuo y permanente que se desarrolla a lo largo de todo el periodo, valorando todos los aspectos de acuerdo con criterios previamente establecidos, que permite identificar los progresos alcanzados por los estudiantes y las dificultades experimentadas, estableciendo rutas de mejoramiento.

La evaluación sumativa y formativa del curso de Inferencia Estadística 1 se llevará a cabo usando diferentes instrumentos de evaluación definidos a continuación:

* Ejercicios: son problemas propuestos en clase para que los estudiantes los resuelvan usando los métodos enseñados.
* Parciales: son pruebas individuales dónde el estudiante demuestra saber los conceptos.
* Quices: son pruebas cortas y mucho más sencillas que las evaluaciones en las que los estudiantes demuestran que saben los conceptos vistos en clase.
* Talleres: Son pruebas que se desarrollan individual o grupalmente en las que los estudiantes demuestran que saben los contenidos del curso.
* Participación en discusiones y en el aula virtual: son oportunidades que tienen los estudiantes para demostrar su compromiso con el curso y su aprendizaje al usar el material disponible, aportar sus ideas y exponer sus dudas.

Porcentajes de las evaluaciones sumativas:

|  |  |
| --- | --- |
| Concepto | Porcentaje |
| Parcial 1. | 20% |
| Parcial 2. | 20% |
| Parcial 3. | 20% |
| Talleres, quices y participación | 20% |
| Parcial Final. | 20% |

**CONTENIDO TEMÁTICO**:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **semana** | **sesión** | **mes** | **día** | **Contenidos** |
| 1 | 1 | enero | 16 | Experimentos aleatorios. Sucesos aleatorios. Sucesos elementales. El espacio muestral. Operaciones con sucesos |
| 2 | 18 | Ejercitación sucesos aleatorios |
| 3 | 20 | Probabilidad. Propiedades. La definición axiomática de probabilidad. La definición clásica y geométrica. La probabilidad subjetiva |
| 2 | 1 | 23 | Ejercitación Probabilidades |
| 2 | 25 | Ejercitación probabilidades |
| 3 | 27 | Probabilidad condicional e independencia. El teorema de Bayes |
| 3 | 1 | 30 | Ejercitación probabilidad condicional, independencia y teorema de Bayes |
| 2 | febrero | 1 | Resumen tema de probabilidades |
| 3 | 3 | **1er parcial** |
| 4 | 1 | 6 | Variables aleatorias y su distribución de probabilidad. La clasificación de las variables en discretas y continuas |
| 2 | 8 | Ejercitación, determinación de distribuciones en el caso discreto. |
| 3 | 10 | Características numéricas: Medidas de posición y de dispersión; media y varianza de las distribuciones. Los percentiles |
| 5 | 1 | 13 | Ejercitación general. Variables aleatorias |
| 2 | 15 | Variables aleatorias discretas. Distribuciones de Bernoulli, Binomial, Poisson |
| 3 | 17 | Variables aleatorias discretas. Distribuciones geométrica e hipergeométrica |
| 6 | 1 | 20 | Variables aleatorias continuas: Las distribuciones uniforme, exponencial y normal |
| 2 | 22 | Ejercitación |
| 3 | 24 | Ejercitación |
| 7 | 1 | 27 | Población, muestra, criterios de muestreo |
| 2 | marzo | 1 | Las bases de datos, su estructura |
| 3 | 3 | Ejercitación resumen |
| 8 | 1 | 6 | **2do parcial** |
| 2 | 8 | Análisis de Datos Exploratorio: Descripción de las variables cualitativas. Representación gráfica de las variables continuas |
| 3 | 10 | Características numéricas para describir el comportamiento de las variables cuantitativas: Medidas de posición central y no central, las medias de dispersión |
| 9 | 1 | 13 | Los diagramas de cajas. Las medidas de forma. La distribución de los datos. La identificación de datos atípicos. |
| 2 | 15 | Ejercitación resumen |
| 3 | 17 | Ejercitación resumen |
| 10 | 1 | 20 | Festivo |
| 2 | 22 | Aproximación de las distribuciones muestrales utilizando el Método de Monte Carlo |
| 3 | 24 | Ejercitación |
| 11 | 1 | 27 | Estimadores puntuales. Propiedades fundamentales. Sesgo, precisión y error de estimación |
| 2 | 29 | El método de estimación Bootstrap |
| 3 | 31 | Ejercitación |
|  | | abril | 3-7 | RECESO |
| 12 | 1 | 10 | Estimador por intervalos de confianza. La precisión de un intervalo |
| 2 | 12 | Estimadores por intervalo más utilizados. El método del Boostrap para construir intervalos de confianza |
| 3 | 14 | Ejercitación |
| 13 | 1 | 17 | **3er parcial** |
| 2 | 19 | Pruebas de hipótesis: Modelos paramétricos y no paramétricos. Nociones fundamentales: Hipótesis estadísticas, región crítica, función de potencia. Los errores en el proceso de las pruebas de hipótesis |
| 3 | 21 | Los valores-p como medidas de evidencia, su interpretación y usos. Los fundamentos de las pruebas de significación. Alcances y limitaciones de los procedimientos de pruebas de hipótesis y de significación |
| 14 | 1 | 24 | Aplicaciones en situaciones prácticas |
| 2 | 26 | Los casos paramétricos más importantes: caso de una población |
| 3 | 28 | Los casos paramétricos más importantes: caso de dos poblaciones |
| 15 | 1 | mayo | 1 | Festivo |
| 2 | 3 | Estadística no paramétrica, sus fundamentos |
| 3 | 5 | Pruebas no paramétricas. |
| 16 | 1 | 8 | Ejercitación |
| 2 | 10 | Ejercitación resumen |
| 3 | 12 | Ejercitación resumen |
| 17 |  | 16 | **Examen Final** |

**8. REFERENCIAS BIBLIOGRÁFICAS**

Ramachandran, K. M., & Tsokos, C. P. (2014). Mathematical statistics with applications in R. Elsevier.

Tattar, P. N., Ramaiah, S., & Manjunath, B. G. (2016). A Course in Statistics with R. John Wiley & Sons.

Chihara, L., & Hesterberg, T. (2011). Mathematical statistics with resampling and R. Hoboken, NJ: Wiley.

Ruppert, D. (2011). Statistics and data analysis for financial engineering (Vol. 13). New York: Springer.