

Nombre de la asignatura: **PROBABILIDAD Y ESTADÍSTICA**

Línea de Trabajo: **Ingeniería de Software/Inteligencia Artificial**

Tiempo de dedicación del estudiante a las actividades de:

DOC	TIS	TPS	Horas Totales	Créditos
48	40	80	168	6

**DOC:** Docencia; **TIS:** Trabajo independiente significativo; **TPS:** Trabajo profesional supervisado

### 1. Historial de la asignatura

Lugar y Fecha de Elaboración o Revisión	Participantes	Observaciones (cambios y justificaciones)
27/Enero/2025	Dr. Mario Alberto Gómez Rodríguez Dr. Jesús Carlos Carmona Frausto	Versión inicial, incluyendo el uso de la IA

### 2. Pre-requisitos y Correquisitos

Ninguno

### 3. Objetivo de la asignatura

Aplicar modelos probabilísticos y técnicas estadísticas en el desarrollo de aplicaciones y experimentos computacionales, los cuales permitirán realizar aproximaciones de diversos fenómenos del estudio científico.

### 4. Aportación al perfil del graduado

Permite abordar problemas complejos mediante el uso de modelos probabilísticos, optimización de procesos y análisis de grandes volúmenes de información. La comprensión de variables aleatorias y sus distribuciones, junto con el ajuste de modelos y la aplicación de pruebas de hipótesis, dota al egresado de la capacidad para desarrollar sistemas inteligentes, optimizar algoritmos y mejorar el procesamiento de datos en diversas áreas como inteligencia artificial, seguridad informática y gestión de bases de datos. Además, el manejo de intervalos de confianza, comparación de poblaciones y experimentos multifactoriales fortalece su habilidad para evaluar el rendimiento de sistemas computacionales y tomar decisiones fundamentadas en evidencia.

### 5. Contenido temático

Unidad	Temas	Subtemas
1	Conceptos básicos de estadística	1.1 ¿Por qué la Estadística? 1.2 Poblaciones y Muestras 1.3 Algunos Conceptos de Muestreo 1.3.1 Muestras Representativas 1.3.2 Muestreo Aleatorio Simple y Muestreo Estratificado 1.3.3 Muestreo con y sin Reemplazo 1.3.4 Muestreo No Representativo 1.4 Variables Aleatorias y Poblaciones Estadísticas 1.5 Proporciones, Promedios y Varianzas 1.5.1 Proporción de la Población y Proporción de la Muestra 1.5.2 Promedio de la Población y Promedio de la Muestra 1.5.3 Varianza de la Población y Varianza de la Muestra 1.6 Medianas, Percentiles y Diagramas de Caja 1.7 La Inteligencia Artificial en el Análisis Estadístico y Muestreo

2	Introducción a probabilidad	<ul style="list-style-type: none"> <li>2.1 Espacios Muestrales, Eventos y Operaciones con Conjuntos</li> <li>2.2 Experimentos con Resultados Iguales Probables <ul style="list-style-type: none"> <li>2.2.1 Definición e Interpretación de la Probabilidad</li> <li>2.2.2 Técnicas de Conteo</li> <li>2.2.3 Funciones de Masa de Probabilidad y Simulaciones</li> </ul> </li> <li>2.3 Axiomas y Propiedades de las Probabilidades</li> <li>2.4 Probabilidad Condicional <ul style="list-style-type: none"> <li>2.4.1 La Regla de Multiplicación y Diagramas de Árbol</li> <li>2.4.2 Ley de la Probabilidad Total y el Teorema de Bayes</li> </ul> </li> <li>2.5 Eventos Independientes <ul style="list-style-type: none"> <li>2.5.1 Aplicaciones a la Fiabilidad de los Sistemas</li> </ul> </li> <li>2.6 Aplicaciones de la Teoría de Probabilidad en la Inteligencia Artificial</li> </ul>
3	Variables aleatorias y sus distribuciones	<ul style="list-style-type: none"> <li>3.1 Introducción</li> <li>3.2 Descripción de una Distribución de Probabilidad <ul style="list-style-type: none"> <li>3.2.1 Variables Aleatorias, Revisitadas</li> <li>3.2.2 La Función de Distribución Acumulada</li> <li>3.2.3 La Función de Densidad de una Distribución Continua</li> </ul> </li> <li>3.3 Parámetros de las Distribuciones de Probabilidad <ul style="list-style-type: none"> <li>3.3.1 Valor Esperado</li> <li>3.3.2 Varianza y Desviación Estándar</li> <li>3.3.3 Percentiles de la Población</li> </ul> </li> <li>3.4 Modelos para Variables Aleatorias Discretas <ul style="list-style-type: none"> <li>3.4.1 Las Distribuciones de Bernoulli y Binomial</li> <li>3.4.2 La Distribución Hipergeométrica</li> <li>3.4.3 Las Distribuciones Geométrica y Binomial Negativa</li> <li>3.4.4 La Distribución de Poisson</li> </ul> </li> <li>3.5 Modelos para Variables Aleatorias Continuas <ul style="list-style-type: none"> <li>3.5.1 La Distribución Exponencial</li> <li>3.5.2 La Distribución Normal</li> </ul> </li> <li>3.6 Distribuciones de Probabilidad e Inteligencia Artificial</li> </ul>
4	Variables aleatorias distribuidas conjuntamente	<ul style="list-style-type: none"> <li>4.1 Introducción</li> <li>4.2 Descripción de las Distribuciones de Probabilidad Conjunta <ul style="list-style-type: none"> <li>4.2.1 La PMF Conjunta y Marginal</li> <li>4.2.2 La PDF Conjunta y Marginal</li> </ul> </li> <li>4.3 Distribuciones Condicionales <ul style="list-style-type: none"> <li>4.3.1 Funciones de Masa de Probabilidad Condicional</li> <li>4.3.2 Funciones de Densidad de Probabilidad Condicional</li> <li>4.3.3 La Función de Regresión</li> <li>4.3.4 Independencia</li> </ul> </li> <li>4.4 Valor Medio de Funciones de Variables Aleatorias <ul style="list-style-type: none"> <li>4.4.1 El Resultado Básico</li> <li>4.4.2 Valor Esperado de Sumas</li> <li>4.4.3 La Covarianza y la Varianza de Sumas</li> </ul> </li> <li>4.5 Cuantificación de la Dependencia <ul style="list-style-type: none"> <li>4.5.1 Dependencia Positiva y Negativa</li> <li>4.5.2 Coeficiente de Correlación de Pearson (o Lineal)</li> </ul> </li> </ul>

		<p>4.6 Modelos para Distribuciones Conjuntas</p> <p>4.6.1 Modelos Jerárquicos</p> <p>4.6.2 Modelos de Regresión</p> <p>4.6.3 La Distribución Normal Bivariada</p> <p>4.6.4 La Distribución Multinomial</p> <p>4.7 Aplicaciones de las Distribuciones Conjuntas en IA</p>
5	Ajustando modelos a los datos	<p>5.1 Introducción</p> <p>5.2 Algunos Conceptos de Estimación</p> <p>5.2.1 Estimación No Sesgada</p> <p>5.2.2 Estimación Sin Modelo vs Estimación Basada en Modelo</p> <p>5.3 Métodos para Ajustar Modelos a los Datos</p> <p>5.3.1 El Método de los Momentos</p> <p>5.3.2 El Método de Máxima Verosimilitud</p> <p>5.3.3 El Método de Mínimos Cuadrados</p> <p>5.4 Comparación de Estimadores: El Criterio MSE</p> <p>5.5 Inteligencia Artificial: Redes Neuronales y Estimación Estadística</p>
6	Intervalos de confianza y predicción	<p>6.1 Introducción a los Intervalos de Confianza</p> <p>6.1.1 Construcción de Intervalos de Confianza</p> <p>6.1.2 Intervalos de Confianza Z</p> <p>6.1.3 La Distribución T y los Intervalos de Confianza T</p> <p>6.2 Semántica de los Intervalos de Confianza: El Significado de "Confianza"</p> <p>6.3 Tipos de Intervalos de Confianza</p> <p>6.3.1 Intervalos de Confianza T para la Media</p> <p>6.3.2 Intervalos de Confianza Z para Proporciones</p> <p>6.3.3 Intervalos de Confianza T para los Parámetros de Regresión</p> <p>6.3.4 El Intervalo de Confianza por Signo para la Mediana</p> <p>6.3.5 Intervalos de Confianza <math>\chi^2</math> para la Varianza Normal y la Desviación Estándar</p> <p>6.4 El Problema de la Precisión</p> <p>6.5 Intervalos de Predicción</p> <p>6.5.1 Conceptos Básicos</p> <p>6.5.2 Predicción de una Variable Aleatoria Normal</p> <p>6.5.3 Predicción en Regresión Lineal Simple Normal</p> <p>6.6 Predicción de Variables con Inteligencia Artificial: Redes Neuronales Artificiales y Árboles de Decisión</p>
7	Pruebas de hipótesis	<p>7.1 Introducción</p> <p>7.2 Establecimiento de un Procedimiento de Prueba</p> <p>7.2.1 Las Hipótesis Nula y Alternativa</p> <p>7.2.2 Estadísticos de Prueba y Reglas de Rechazo</p> <p>7.2.3 Pruebas Z y Pruebas T</p> <p>7.2.4 Valores P</p> <p>7.3 Tipos de Pruebas</p> <p>7.3.1 Pruebas T para la Media</p> <p>7.3.2 Pruebas Z para Proporciones</p> <p>7.3.3 Pruebas T sobre los Parámetros de Regresión</p> <p>7.3.4 La Prueba F de ANOVA en Regresión</p> <p>7.3.5 La Prueba de Signos para la Mediana</p> <p>7.3.6 Pruebas <math>\chi^2</math> para una Varianza Normal</p> <p>7.4 Precisión en las Pruebas de Hipótesis</p> <p>7.4.1 Errores Tipo I y Tipo II</p> <p>7.4.2 Cálculos de Potencia y Tamaño de Muestra</p> <p>7.5 Hypotheses testing in Artificial Intelligence: Machine Learning</p>

8	Comparando dos poblaciones	8.1 Introducción 8.2 Pruebas de dos muestras e intervalos de confianza para medias 8.2.1 Algunos resultados básicos 8.2.2 Intervalos de confianza 8.2.3 Pruebas de hipótesis 8.3 El procedimiento de la prueba de suma de rangos 8.4 Comparación de dos varianzas 8.4.1 La prueba de Levene 8.4.2 La prueba F bajo normalidad 8.5 Datos emparejados 8.5.1 Definición y ejemplos de datos emparejados 8.5.2 La prueba T para datos emparejados 8.5.3 La prueba T para proporciones emparejadas 8.5.4 La prueba de rangos con signo de 8.6 Comparación de poblaciones con IA: Aprendizaje Automático
9	Experimentos multifactoriales	9.1 Introducción 9.2 Diseños de Dos Factores 9.2.1 Pruebas F para Efectos Principales e Interacciones 9.2.2 Prueba de la Validez de Supuestos 9.2.3 Una Observación por Celda 9.3 Diseños de Tres Factores 9.4 Experimentos Factoriales 2r 9.4.1 Bloqueo y Confusión 9.4.2 Diseños Factoriales Fraccionados 9.5 Inteligencia Artificial y Diseños Experimentales: reducción de la dimensionalidad de los datos
10	Regresión polinomial y múltiple	10.1 Introducción 10.2 El Modelo de Regresión Lineal Múltiple 10.3 Estimación, Pruebas y Predicción 10.3.1 Los Estimadores de Mínimos Cuadrados 10.3.2 Prueba de Utilidad del Modelo 10.3.3 Prueba de Significancia de los Coeficientes de Regresión 10.3.4 Intervalos de Confianza y Predicción 10.4 Temas Adicionales 10.4.1 Mínimos Cuadrados Ponderados 10.4.2 Aplicaciones a Diseños Factoriales 10.4.3 Selección de Variables 10.4.4 Observaciones Influyentes 10.4.5 Multicolinealidad 10.4.6 Regresión Logística 10.5 Inteligencia Artificial y Aprendizaje automático: Regresión Polinomial y Múltiple

## 6. Metodología de desarrollo del curso

El profesor expondrá algunos conceptos teóricos del curso del curso y el estudiante mediante investigaciones conducidas por el docente generará el conocimiento necesario para poder usar diferentes métodos estadísticos y probabilísticos con diferentes conjuntos de datos. Adicionalmente, se usarán casos reales donde el alumno deberá implementar alguna de las técnicas vistas en el curso.

## 7. Sugerencias de evaluación

Se recomienda la siguiente ponderación:

Examen de medio término	20%
Tareas	40%
Proyecto Final	40%

## 8. Bibliografía y software de Apoyo

- Gupta, B. C., Guttman, I., & Jayalath, K. P. (2020). *Statistics and probability with applications for engineers and scientists using MINITAB, R and JMP*. John Wiley & Sons.
- Aggarwal, C. C. (2024). *Probability and statistics for machine learning: a textbook*. Springer Nature.
- Navidi, W. (2006). Estadística para ingenieros y científicos.
- Devore, J. L. (2000). Probability and statistics. *Pacific Grove: Brooks/Cole*.

## 9. Actividades propuestas

Unidad	Actividades
1	Análisis estadístico de logs del sistema: Los estudiantes procesan archivos de registro (logs) de un sistema, analizando métricas como tiempo medio entre fallos, distribución de tiempos de respuesta y tendencias.
2	Simulación de colisiones en una red: Se modela la probabilidad de colisiones en un protocolo de comunicación, como CSMA/CD, simulando paquetes en una red y calculando eventos de colisión.
3	Generación de tráfico en redes: Se modelan paquetes en una red como una variable aleatoria de distribución Poisson y se simula la llegada de paquetes en un servidor.
4	Análisis de correlación en grandes volúmenes de datos: Se analiza la relación entre el uso de CPU y la latencia de respuesta en un sistema distribuido, empleando matrices de correlación y análisis conjunto.
5	Regresión sobre datos de sensores IoT: Se utilizan datos de sensores inteligentes (temperatura, humedad) y se ajusta un modelo predictivo para optimizar el consumo energético.
6	Estimación de tiempos de ejecución en algoritmos: Se ejecutan algoritmos de ordenamiento en diferentes tamaños de entrada y se calculan intervalos de confianza sobre los tiempos de ejecución promedio.
7	Evaluación de eficiencia de algoritmos: Se comparan dos algoritmos de clasificación (ejemplo: KNN vs. SVM) para determinar si hay una diferencia estadísticamente significativa en su precisión.
8	Análisis de rendimiento entre dos arquitecturas: Se evalúa el rendimiento de una aplicación ejecutándose en dos plataformas distintas (ejemplo: ARM vs. x86) mediante pruebas t de Student.
9	Optimización de hiperparámetros en Machine Learning: Se entrena un modelo de redes neuronales variando múltiples hiperparámetros y se analizan sus efectos con un diseño factorial completo.
10	Predicción de fallos en servidores: Se desarrolla un modelo de regresión múltiple para predecir fallos en servidores en función de métricas como temperatura, uso de CPU y RAM.

**10. Catedrático (s) responsable (s)**

Handwritten signature of Mario Alberto Gómez Rodríguez in black ink.

Dr. Mario Alberto Gómez Rodríguez

Handwritten signature of Jesús Carlos Carmona Frausto in black ink.

Dr. Jesús Carlos Carmona Frausto