



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

---

DIVISIÓN ACADÉMICA DE CIENCIAS ECONÓMICO

## “PORTAFOLIO FINAL DE EVIDENCIA DE ESTADÍSTICA INFERENCIAL” LICENCIATURA EN CONTADURIA PUBLICA

**ALUMNO:**

Luis Enrique Muñoz Paulino

**MATRICULA:**

241b38018

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**CICLO ESCOLAR:**

Febrero 2025-Agosto 2025

Villahermosa, Tabasco, 31 de mayo de 2025

## Tabla de contenido

INTRODUCCIÓN .....	4
PROGRAMA DE ESTUDIO .....	5
UNIDAD 1 TEOREMA DEL LIMITE CENTRAL Y ESTIMACIÓN .....	18
“ENSAYO” .....	19
“MAPAS MENTALES” .....	32
“DOCUMENTO EXTENSO” .....	35
“PRACTICAS DE LABORATORIO DE.....	53
COMPUTO” .....	53
“Examen Primer Parcial” .....	65
CONCLUSIÓN SOBRE EL PARCIAL .....	72
UNIDAD 2 PRUEBA DE HIPOTESIS Y CONTROL DE CALIDAD .....	73
“ENSAYO” .....	74
“MAPAS MENTALES” .....	98
“DOCUMENTO EXTENSO” .....	100
“PRACTICAS DE LABORATORIO DE COMPUTO” .....	118
EXAMEN 2do. Parcial .....	127
CONCLUSIÓN SOBRE EL PARCIAL .....	130
UNIDAD 3 Y 4 ANOVA, CORRELACIÓN, REGRESIÓN LINEAL .....	131
“ENSAYO” .....	132
“MAPAS MENTALES” .....	151
“DOCUMENTO EXTENSO” .....	155
CONCLUSIÓN SOBRE EL PARCIAL .....	184
EVIDENCIAS .....	185
.....	185
CONCLUSIÓN FINAL .....	187
Referencias .....	188

## INTRODUCCIÓN

En este portafolio final de evidencias de la materia de *Estadística Inferencial*, presento de manera organizada los trabajos, actividades, ejercicios y prácticas que realicé a lo largo del curso. El objetivo principal de este portafolio es mostrar todo lo que he aprendido durante el semestre, así como evidenciar el desarrollo de mis habilidades para aplicar los conceptos estadísticos en diferentes situaciones.

Durante la materia, aprendí que la Estadística Inferencial es una herramienta muy útil porque nos permite sacar conclusiones sobre una población a partir de datos obtenidos de una muestra. A lo largo del curso, trabajé con temas como estimaciones, intervalos de confianza, pruebas de hipótesis, regresión lineal, entre otros, que me ayudaron a entender cómo se analizan los datos de manera más profunda y precisa.

Este portafolio también representa mi esfuerzo, compromiso y crecimiento académico. Al reunir todas estas evidencias, puedo reflexionar sobre mi proceso de aprendizaje y reconocer tanto los logros como las áreas en las que aún puedo seguir mejorando.





Nombre de la asignatura									Estadística Inferencial	Clave de la asignatura
Área de formación:	Docencia frente a grupo según SATCA				Trabajo de Campo Supervisado según SATCA				Carácter de la asignatura	
	HCS	HP	TH	C	HTCS	TH	C	TC	(X) Obligatoria	( ) Optativa
Sustantivas profesional	1	4	5	5					(X) Obligatoria	( ) Optativa

SERIACIÓN		
Explícita		Implícita
Asignaturas antecedentes	Asignaturas subsecuentes	Conocimientos previos
Introducción a la estadística	Investigación de Operaciones	Análisis de datos numéricos, agrupación de datos de tipo discretos y continuos, diagramación de estadística descriptiva, distribuciones probabilísticas discretas y continuas Herramientas de computación



PROPÓSITO DE LA ASIGNATURA	
<p>Aplicar las técnicas básicas de la Estadística Inferencial para la realización de diversos estudios de índole económico, social, político y otros indagando y analizando el comportamiento de fenómenos aleatorios mediante el estudio de diversos parámetros y estimaciones de las poblaciones conocidas y objeto de estudio a través de muestras representativas de las mismas y aplicando distribuciones teóricas de probabilidad, permitiendo la inferencia para realizar conclusiones a partir la información que se puede obtener de las pruebas de hipótesis y explicar fenómenos involucrados con los procesos de toma de decisiones racionales y sustentadas.</p>	
COMPETENCIAS A DESARROLLAR	
Genéricas	Específicas
<p><i>Capacidad de análisis y síntesis</i> <i>Capacidad de aplicar los conocimientos en la práctica</i> <i>Uso de las TIC</i></p>	<p>Tomada de decisiones oportunas racionales y sustentadas. Análisis de alternativas para la toma de decisiones. Uso de herramientas generadoras de información confiable y de calidad.</p>



UNIDAD No. I	TEOREMA DEL LÍMITE CENTRAL Y ESTIMACIÓN		Horas estimadas para cada unidad
			15
CONTENIDOS			
Conceptuales	Aprendizaje esperado	Evidencias de aprendizaje	
<p>1.1 Distribución muestral de la media. Teorema del límite central.</p> <p>1.2 Distribución de la proporción.</p> <p>1.3 Propiedades de los estimadores</p> <p>Estimación puntual y por intervalo para parámetros con población infinita y tamaño grande de muestra para la:</p> <ul style="list-style-type: none"> <li>• Media</li> <li>• Proporción</li> <li>• Diferencia de medias</li> <li>• Diferencia de proporciones dos poblaciones.</li> </ul> <p>1.4 Estimación puntual y por intervalo para parámetros con población finita y/o tamaño pequeño de muestra.</p>	<p><i>Interpreta, evalúa y sintetiza los problemas que requieren de la estadística inferencial</i></p> <p><i>Realiza estimaciones puntuales y por intervalo para parámetros estadísticos con poblaciones finitas e infinitas con el uso de software especializado en estadística.</i></p> <p><i>Aplica técnicas y herramientas tecnológicas para el planteamiento y solución de problemas mediante la prueba de hipótesis</i></p>	<p><i>Mapas conceptuales.</i></p> <p><i>Diseño de experimentos que permitan la aplicación del concepto de estimación puntual y estimación por intervalo.</i></p> <p><i>Trabajo extra clase de investigación, revisión e identificación de conceptos.</i></p> <p><i>Planteamiento de una investigación estadística a un problema aplicado a la carrera</i></p> <p><i>Solución de caso práctico.</i></p>	



1.5 Aplicación utilizando software disponible:



UNIDAD No. II	PRUEBA DE HIPÓTESIS Y CONTROL DE CALIDAD		Horas estimadas para cada unidad
			25
CONTENIDOS			
Conceptuales	Aprendizaje esperado	Evidencias de aprendizaje	
2.1 Conceptos básicos 2.2 Tipos de errores 2.3 Prueba de hipótesis para media y proporción poblacional (población infinita -muestra grande) 2.4 Prueba de hipótesis para diferencia de dos medias y proporciones poblacionales 2.5 Casos de población finita o tamaño pequeño de muestra. 2.6 Conceptos básicos de Control de Calidad 2.7 Diagrama de control de media (X) 2.8 Diagrama de control de (p) 2.9 Diagrama de control de (c) 2.10 Muestreo de aceptación. 2.11 Aplicación utilizando software disponible:	Aprender a utilizar las técnicas estadísticas para la formulación y contrastación de una hipótesis estadística para los principales parámetros de una distribución de probabilidad utilizando los resultados para la toma de decisiones.  Comprender los fundamentos básicos de los modelos de control de calidad, proponiendo esquemas de implantación en empresas e instituciones.	<i>Identificar y utilizar los métodos de toma de decisiones.</i>  <i>Resolución de problemas prácticos relacionados con la carrera.</i>  <i>Simulación de casos que generen hipótesis con interés en media, diferencia de medias, proporciones, diferencia de proporciones, varianza y relación de varianza.</i>  <i>Prueba de hipótesis documentada de los casos de estudio.</i>  <i>Determinación del impacto del cambio del tamaño de la muestra en la prueba de hipótesis.</i>  <i>Portafolio de evidencias con problemas resueltos.</i>  <i>Simulación de casos para el control de calidad.</i>	



UNIDAD No. III	<b>ANOVA y CORRELACIÓN, REGRESIÓN LINEAL</b>		Horas estimadas para cada unidad
			25
CONTENIDOS			
Conceptuales	Aprendizaje esperado	Evidencias de aprendizaje	
3.1 Introducción a la técnica del análisis de varianza (ANOVA) 3.2 Fundamentos del ANOVA 3.3 Aplicaciones del ANOVA de un solo factor 3.4 Datos bivariados 3.5 Correlación. 3.6 Análisis e interpretación del coeficiente de Pearson ( $r$ ) y de determinación ( $r^2$ ). 3.7 Regresión lineal simple. 3.8 Método de los mínimos cuadrados. 3.9 Análisis e interpretación de los coeficientes $b_0$ y $b_1$ . 3.10 Prueba de hipótesis de significatividad de la ecuación de regresión. 3.11 Estimación por intervalo para los coeficientes $b_0$ y $b_1$ . 3.12 Aplicación utilizando software disponible:	<p>Aplicar el análisis correlación para conocer el grado de relación entre dos variables, que permita validar la modelización del método de regresión lineal que pueda explicar el comportamiento de las variables y hacer predicciones</p> <p>Comprender y aplicar el análisis de varianza para contrastar hipótesis acerca de varias medias poblacionales.</p>	<p><i>Identificar y utilizar el análisis de correlación entre variables.</i></p> <p><i>Resolución de problemas prácticos relacionados con la carrera.</i></p> <p><i>Simulación de casos de análisis de correlación, regresión lineal y análisis de varianza.</i></p> <p><i>Portafolio de evidencias con problemas resueltos.</i></p>	



UNIDAD No. IV	<b>SERIES DE TIEMPO Y ESTADÍSTICA NO PARAMÉTRICA</b>		Horas estimadas para cada unidad
			15
CONTENIDOS			
Conceptuales	Aprendizaje esperado	Evidencias de aprendizaje	
4.1 Componentes de una serie de tiempo 4.2 Métodos de alisamiento promedios móviles y alisamiento exponencial 4.3 Análisis de la tendencia por regresión lineal simple 4.4 Análisis de la componente estacional e índices estacionales. 4.5 Tablas de contingencia 4.6 La prueba Ji-cuadrada 4.7 Prueba de independencia 4.8 Prueba de homogeneidad. 4.9 La prueba del signo. 4.10 Aplicación utilizando software disponible	<p>Aplicar las técnicas básicas para el análisis del conjunto de dos variables cualitativas, para determinar si entre ellas hay relación o asociación o por el contrario son independientes,</p> <p>Aplicar la metodología del análisis de series de tiempo para realizar estimaciones de los productos</p>	<p><i>Identificación de los diferentes elementos de una serie de tiempo y de alisamiento.</i></p> <p><i>Identificar y utilizar el análisis de correlación entre variables tomando al tiempo como factor.</i></p> <p><i>Resolución de problemas prácticos relacionados con la carrera con énfasis en las series de tiempo</i></p> <p><i>Simulación de casos de análisis de series de tiempo.</i></p> <p><i>Análisis de variables categóricas y pruebas estadísticas para estas variables</i></p> <p><i>Portafolio de evidencias con problemas resueltos.</i></p>	



Contenidos procedimentales	Contenidos actitudinales
<p><i>Aplicación de los principales conceptos y métodos más importantes y generales de la Estadística Inferencial.</i></p> <p><i>Análisis de situaciones reales cotidianas Exploración de problemas y fenómenos del entorno</i></p> <p><i>Vinculación del cálculo en contenidos de la inferencia estadística y los tipos de distribuciones con variables continuas.</i></p> <p><i>Selección del tipo de muestreo que sea necesario para dar tratamiento a un experimento estadístico.</i></p> <p><i>Trabajo en equipo. Cálculo del tamaño de la muestra.</i></p> <p><i>Solución de problemas más complejos en las distribuciones Muestrales.</i></p>	<p><i>Se debe manifestar una amplia confianza en el momento de plantear y resolver situaciones problemáticas.</i></p> <p><i>La disciplina, el esfuerzo y la perseverancia se constituyen en una actitud fundamental ante la búsqueda de resultados.</i></p> <p><i>La valoración del intercambio de ideas como fuente del aprendizaje y como herramienta fundamental ante situaciones problemáticas.</i></p> <p><i>La valoración de la estadística como parte de la construcción de la persona no solo como ser humano sino también como profesional.</i></p> <p><i>El respeto por el pensamiento ajeno que induce a formar profesionales con una amplia capacidad de auto proceso.</i></p> <p><i>El sentido de la crítica frente a los resultados obtenidos en la resolución y visualización de un problema.</i></p> <p><i>La curiosidad, apertura y duda como la base que ha de permitir una reestructuración permanente en los procedimientos estadísticos.</i></p> <p><i>La valoración de un lenguaje claro y preciso como expresión y organización del pensamiento.</i></p>



<p><i>Desarrollar la metodología de prueba de hipótesis como una técnica para analizar diferencias y tomar decisiones; determinar los riesgos implicados al basarnos únicamente en la información de muestra; y estudiar la interrelación de estos riesgos con el tamaño de la muestra utilizada.</i></p> <p><i>Extender los principios básicos de la metodología de prueba de hipótesis hacia las pruebas de una muestra más comúnmente utilizada, que implican datos numéricos.</i></p> <p><i>Utilizar las pruebas de hipótesis para la media, para la varianza o la desviación estándar.</i></p> <p><i>Desarrollar los principios básicos de la prueba de hipótesis a pruebas de dos muestras que impliquen variables numéricas.</i></p> <p><i>Se consideran tanto procedimientos de muestras independientes como relacionadas.</i></p> <p><i>Fijación de límites y tolerancias</i></p> <p><i>Utilizar diferentes tipos de gráficos para el control de la calidad</i></p>	<p><i>Distinguir entre la hipótesis nula y alternativa, los riesgos de cometer un error del tipo I y un error del tipo II.</i></p> <p><i>Entender el concepto de potencia de una prueba y distinguir entre prueba de un extremo y de dos extremos.</i></p> <p><i>Entender el planteamiento del valor p de la prueba de hipótesis.</i></p> <p><i>Apreciar la conexión entre intervalos de confianza y pruebas de hipótesis.</i></p> <p><i>Ventajas y desventajas del tipo de muestreo en el control de la calidad.</i></p> <p><i>Cálculo de indicadores de la calidad de entrada y de salida.</i></p> <p><i>Identifica diferentes diagramas como Pareto, Causa – efecto, Diagrama de flujo</i></p>
<p><i>Desarrollar tanto descriptiva como inferencialmente los modelos de regresión lineal simple y de correlación, como un medio de utilizar una variable para predecir otra, y medir la intensidad de la asociación entre dos variables.</i></p> <p><i>Utilizar el análisis residual para determinar si el modelo adecuado ha sido ajustado a los datos y el análisis de influencia para determinar si algunas observaciones están influyendo indebidamente el modelo de regresión.</i></p> <p><i>Fortalece el concepto del error estándar de la estimación. Elabora e interpreta la tabla del ANOVA</i></p>	<p><i>Aprender a determinar una ecuación que pueda utilizarse en pronósticos.</i></p> <p><i>Interpretar los coeficientes de regresión obtenidos mediante el uso del método de mínimos cuadrados de la regresión.</i></p> <p><i>Comprender el significado de la relación entre una variable explicada y otra variable explicativa y su aplicación con fines de pronóstico.</i></p> <p><i>Interpretar los coeficientes de determinación y de correlación y describir la intensidad de la relación entre las variables.</i></p> <p><i>Diferenciar entre varias medidas de variación en el análisis de regresión.</i></p> <p><i>Familiarizarse con las suposiciones del análisis de regresión.</i></p> <p><i>Hacer inferencias sobre los coeficientes de regresión y valor predicho de una</i></p>



*variable de respuesta.*



	<p><i>Deducir y obtener una apreciación del valor de utilizar los recursos informáticos para la presentación y análisis de la relación lineal y de correlación de las variables y de lo adecuado del modelo.</i></p> <p><i>Realiza un análisis de varianza (ANOVA)</i></p>
<p><i>Establece la relación entre dos variables, considerando la variable independiente como variable tiempo</i></p> <p><i>Determina los cuatro componentes de la serie de tiempo Diseña la ecuación lineal, identificando los parámetros</i></p> <p><i>Aplica las técnicas de pronóstico usando información cronológica.</i></p> <p><i>Obtiene pronósticos</i></p> <p><i>Desarrolla la valoración de una clasificación cruzada de datos con base en varias variables categóricas en Tablas de contingencia, e interpreta objetiva y plenamente el significado de los resultados</i></p>	<p><i>Valora el uso de las tendencias para realizar estimaciones en función al tiempo.</i></p> <p><i>Muestra interés en los pronósticos lineales con relación a variables económicas</i></p> <p><i>Deduce y obtiene una apreciación del valor de usar recursos informáticos para la presentación de tablas de interpretación numérica de los mismos.</i></p> <p><i>Identifica y reconoce el uso de la prueba chi cuadrada para la bondad de ajuste e independencia</i></p>



<b>Metodología para la construcción del conocimiento</b>	
<b>Actividades de aprendizaje con el docente</b>	<b>Actividades de aprendizaje autónomo</b>
<p><i>Exposición de términos conceptuales</i></p> <p><i>Debate de conceptos básicos de la probabilidad aplicados a la inferencia.</i></p> <p><i>Aprendizaje orientado a solución de problemas</i></p> <p><i>Uso de software y calculadora como herramientas de la aplicación para facilitar la comprensión de conceptos, solución de problemas e interpretación de resultados.</i></p> <p><i>Organización de talleres de resolución de problemas prácticos Análisis y reflexión.</i></p> <p><i>Lecturas y exposición de los tipos de errores estadísticos y conceptos de hipótesis</i></p> <p><i>Elaboración de mapas conceptuales</i></p> <p><i>Colección de datos de una muestra para su análisis.</i></p> <p><i>Realizar pruebas de hipótesis, determinación de error tipo I, tipo II.</i></p> <p><i>Realizar pruebas de bondad del ajuste.</i></p>	<p><i>Investigar conceptos de la inferencia estadística. Definir los métodos de recolección de datos.</i></p> <p><i>Definir, comprender y diferenciar el concepto de error estándar y error de muestreo.</i></p> <p><i>Calcular el tamaño de una muestra.</i></p> <p><i>Definir mediante un mapa conceptual la distribución muestral de medias y proporciones.</i></p> <p><i>Identificación y aplicación de las técnicas de estimación de intervalos de confianza a la solución de problemas de aplicación práctica.</i></p> <p><i>Formular y resolver ejercicios aplicando la metodología de prueba de hipótesis para: la media, diferencia de medias, proporción, diferencia de proporciones, varianza y relación de varianzas</i></p> <p><i>Obtener el tamaño de la muestra para diferentes situaciones del error tipo I, error tipo II y para la potencia de la prueba.</i></p> <p><i>Simular un caso en donde:</i></p> <p><i>Se genere una hipótesis para una situación en donde el interés pueda ser, la media, diferencia de medias, proporción, diferencia de proporciones, varianza y relación de varianzas.</i></p>



<p><i>Lecturas y exposición de la correlación entre variables. Elaboración de mapas conceptuales</i></p> <p><i>Talleres para la solución de problemas de correlación, regresión lineal y análisis de varianza.</i></p> <p><i>Lecturas y exposición de series de tiempo y estadística no paramétrica</i></p> <p><i>Elaboración de mapas conceptuales</i></p> <p><i>Talleres para la solución de problemas series de tiempo y estadística no paramétrica</i></p>	<p><i>Generar datos del caso Probar la hipótesis del caso Obtener conclusiones</i></p> <p><i>Cambiar el tamaño de muestra y mostrar su impacto.</i></p> <p><i>Formular y resolver ejercicios aplicando los conceptos de correlación, regresión y análisis de varianza</i></p> <p><i>Simular un caso en donde: Se plantee la correlación entre dos variables, se desarrolle un modelo estadístico de relación entre variables y se genere un análisis de varianza</i></p> <p><i>Obtener conclusiones</i></p> <p><i>Formular y resolver ejercicios aplicando los conceptos estadísticos de series de tiempo y estadística no paramétrica</i></p> <p><i>Simular un caso en donde:</i></p> <p><i>Se planteen los conceptos estadísticos de series de tiempo y estadística no paramétrica</i></p> <p><i>Obtener conclusiones</i></p>
--	---



Evidencias de desempeño		
Acreditación	Evaluación	Calificación
<p>Asistir cuando menos al 90% de las sesiones.</p> <p>Entregar y aprobar el 100% de los trabajos prácticos para tener derecho a una evaluación parcial</p> <p>Calificación mínima de 6 (seis).</p>	<p>Examen de diagnóstico.</p> <p>Revisión y exposición de ejercicios extraclase. Análisis y revisión de las actividades de investigación.</p> <p>Solución e interpretación de problemas resueltos con apoyo del software.</p> <p>Exposición de temas relacionados con la materia.</p> <p>Participación en talleres de resolución de problemas. Entrega de trabajos de investigación en equipo.</p> <p>Realizar investigación de campo.</p> <p>Resolver problemas prácticos en dinámicas grupales.</p> <p>Portafolio de evidencias.</p>	<p><b>UNIDAD I</b></p> <p>Elaboración de Mapas Mentales 0.25%</p> <p>Interpretación y Solución de Problemas 0.5%</p> <p>Trabajo de Investigación en Equipo 0.75%</p> <p>Investigación de Campo 1%</p>
	<p>Revisión y exposición de ejercicios extraclase. Análisis y revisión de las actividades de investigación.</p> <p>Solución e interpretación de problemas resueltos con apoyo del software.</p> <p>Participación en talleres de resolución de problemas. Entrega de trabajos de investigación en equipo.</p> <p>Realizar investigación de campo.</p> <p>Resolver problemas prácticos en dinámicas grupales.</p> <p>Compilación de apuntes por unidades.</p> <p>Cumplir en tiempo y forma con las actividades</p>	<p><b>UNIDAD II</b></p> <p>Elaboración de Mapas Mentales 0.25%</p> <p>Interpretación y Solución de Problemas 0.5%</p> <p>Trabajo de Investigación en Equipo 0.75%</p> <p>Investigación de Campo 1%</p>



	<i>encomendadas. Portafolio de evidencias.</i>	
	<p><i>Revisión y exposición de ejercicios extraclase. Análisis y revisión de las actividades de investigación.</i></p> <p><i>Solución e interpretación de problemas resueltos con apoyo del software.</i></p> <p><i>Participación en talleres de resolución de problemas. Entrega de trabajos de investigación en equipo.</i></p> <p><i>Realizar investigación de campo.</i></p> <p><i>Resolver problemas prácticos en dinámicas grupales.</i></p> <p><i>Compilación de apuntes por unidades.</i></p> <p><i>Cumplir en tiempo y forma con las actividades encomendadas.</i></p> <p><i>Portafolio de evidencias.</i></p>	<p><i>UNIDAD III</i></p> <p><i>Trabajos de Investigación en Equipo 0.75%</i></p> <p><i>Interpretación y Solución de Problemas 0.75%</i></p> <p><i>40% Investigación de Campo 1%</i></p>
	<p><i>Revisión y exposición de ejercicios extraclase. Análisis y revisión de las actividades de investigación.</i></p> <p><i>Solución e interpretación de problemas resueltos con apoyo del software.</i></p>	<p><i>UNIDAD IV</i></p> <p><i>Trabajos de Investigación en Equipo 0.75%</i></p> <p><i>Interpretación y Solución de Problemas 0.75%</i></p>



	<p><i>Participación en talleres de resolución de problemas. Entrega de trabajos de investigación en equipo.</i></p> <p><i>Realizar investigación de campo.</i></p> <p><i>Resolver problemas prácticos en dinámicas grupales.</i></p> <p><i>Compilación de apuntes por unidades.</i></p> <p><i>Cumplir en tiempo y forma con las actividades encomendadas.</i></p> <p><i>Portafolio de evidencias</i></p>	<p><i>Investigación de Campo, presentación de un Caso</i></p>	<p>1%</p>
--	--	---	-----------



**FUENTES DE APOYO Y CONSULTA**

**BÁSICA**

1. VelascoSotomayor, Gabriel (2005). Estadística con Excel. México: Editorial Trillas.
2. Aguilar A. Introducción a la inferencia Estadística. Primera Edición. Editorial Pearson México. 2010
3. Anderson, D. R., Sweeney, D. J. y Williams, T. A. Estadística para Administración y Economía. Décima edición. International Thomson Editores. México. 2008.
4. Gutiérrez, P. H. y De la Vara, S. R. Análisis y diseño de experimentos. Ed. McGraw Hill. 2014.
5. Infante, S. G. y Zárate de L. G. Métodos Estadísticos. Ed. Trillas. México. 2000.
6. Ipiña S. Inferencia Estadística y Análisis de datos. Primera edición. España. 2008.
7. Levine, D. M., Krehbiel, T. C. y Berenson, M. L. Estadística para administración. Cuarta edición. Pearson. México. 2006.
8. Lind, D. A. Estadística aplicada a los negocios y economía. MCGRAW HILL HIGHER EDUCATION. 15ª edición. México. 2012
9. Montgomery, D., Peck., E. A. y Vining, G. G. Introducción al Análisis de Regresión Lineal. Ed. CECSA. 2006.
10. Montgomery, D. Diseño y Análisis de experimentos. Segunda edición. Ed. Limusa Wiley. 2004.
11. O'Mahony, M. Sensory measurement and evaluation of foods. Ed. CRC. 2007.
12. Pardo, M. A. y Ruiz, D. M. A. Análisis de datos con SPSS 13 Base. Ed. McGraw Hill. 2005.
13. Sancho, J., Bota, E., y De Castro, J. J. Introducción al Análisis Sensorial de los Alimentos. Alfaomega Grupo Editor. 2002.
14. Scheaffer, R. L., Ott, L. R. y Mendenhall, W. Elementos de muestreo. Ed. Paraninfo Thomson. 2006
15. Landero Hernández, René y González Ramírez, Mónica T. (2011). Estadística con SPSS y Metodología de la investigación. México: Editorial Trillas.

**COMPLEMENTARIA**

1. Johnson, Robert y Kuby, Patricia (2003). Estadística Elemental. Lo esencial. México: Thomson editores.
2. Berenson, Mark L.; Levine, David M. y Krehbiel, Timothy C. (2001). *Estadística para administración*. México: Editorial Pearson Educación.



<b>RESPONSABLE DEL DISEÑO</b>	
Elaborado por	<p><i>Dra. Aida Armenta Ramírez</i>  <i>M. en C. Alfredo Manzur Bocanegra</i>  <i>M. A. Carlos Alberto Paz Gómez</i>  <i>M. A. Cesar Enrique Barrera Ovando</i>  <i>M. en M. Daniel Chablé Hernández</i>  <i>M. A y M. D Gladys Elena Mateos Gutiérrez</i>  <i>M. A José Domingo Pérez Pérez</i>  <i>M. A José Guadalupe Sánchez Ruíz</i>  <i>M. A. Leonardo Hernández Triano</i>  <i>M. A. Moisés Notario Franco</i>  <i>M. A Octavio Metelín Zarrazaga</i>  <i>I. C. Omar Armando Canto López</i>  <i>M. A. Rubén Pérez Salvador</i>  <i>M. A. Wilber Sarao Pérez</i></p>
Fecha actualización	12 de diciembre de 2016

*Nota:*

*Lo más importante en los programas de estudio es la congruencia entre sus distintos elementos o apartados. Es decir, si el propósito es “formar una empresa de agro negocios” . La competencia es hacer una empresa en todo lo que esta implica. Y los aprendizajes esperados, son los distintos pasos o etapas para su conformación. Los contenidos deben posibilitar la creación de la empresa y se calificará con la instalación de la empresa, con su existencia real.*

*Los programas de estudios por competencias llevan otros componentes, como el de los INDICADORES DE DESEMPEÑO, pero para una IES que inicia su “aventura” en este enfoque curricular, conviene ir por pasos, dado que implica procesos de formación docente. Y también de acompañamiento pedagógico y trabajo colegiado.*



### Nomenclatura

HCS- Horas Clase a la semana.

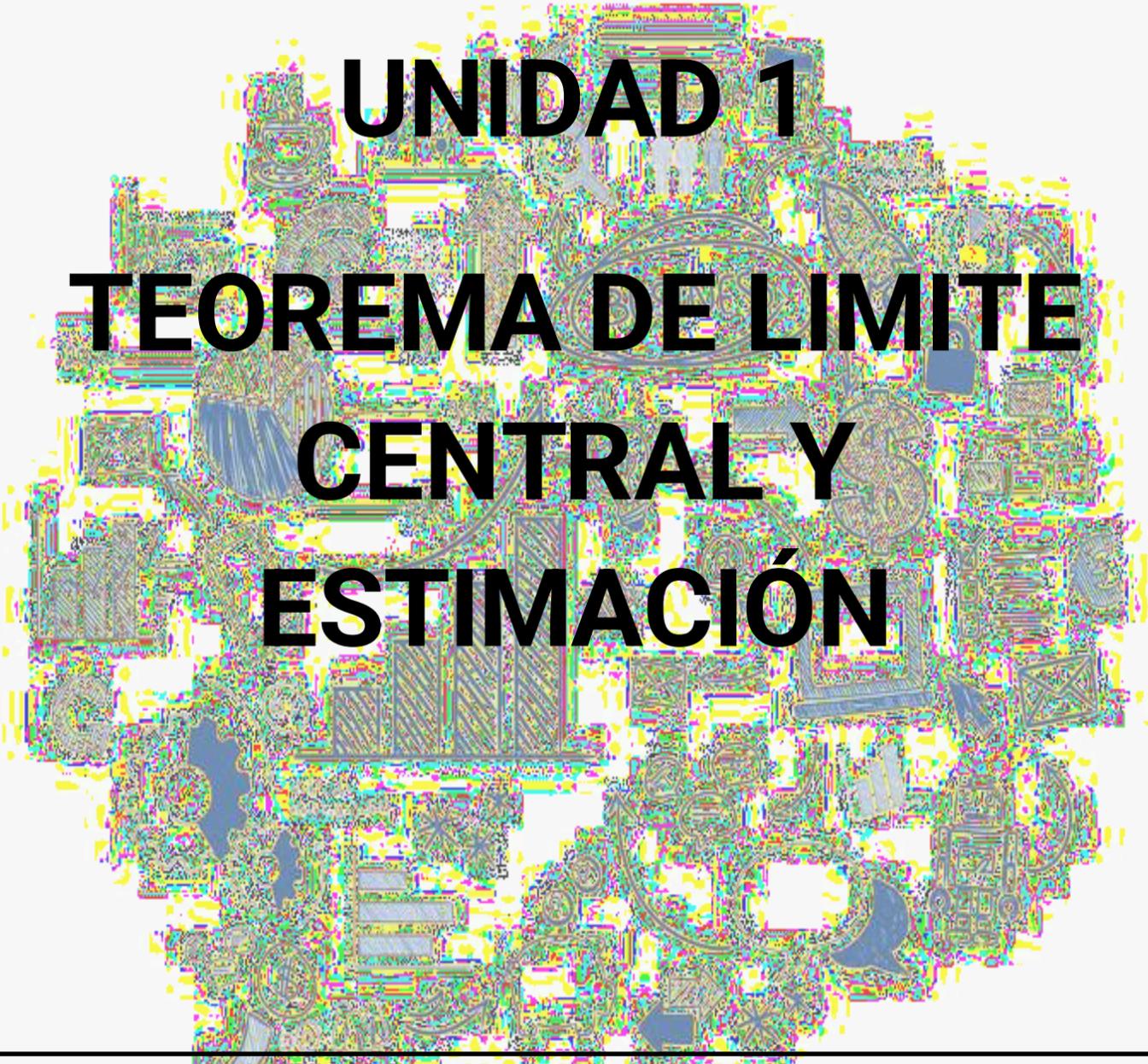
HPS- Horas Prácticas a la semana (laboratorio, seminarios o talleres).

HTCS- Hora de Trabajo de Campo Supervisado a la semana (Servicio Social, Práctica Profesional, internado, estancias, ayudantías). TH-

Total de Horas.

C- Créditos.

TC- Total de créditos.



# **UNIDAD 1**

# **TEOREMA DE LIMITE**

# **CENTRAL Y**

# **ESTIMACIÓN**

## **CONTENIDO:**

**ENSAYO**  
**MAPA MENTAL**  
**DOCUMENTO EXTENSO**  
**PRACTICAS DEL LABORATORIO DE COMPUTO**  
**EXAMEN DE LA UNIDAD**  
**CONCLUSIÓN DE LA UNIDAD**



# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



## **“ENSAYO”**

### **ALUMNO:**

Luis Enrique Muñoz Paulino

## **LICENCIATURA EN CONTADURIA PÚBLICA**

### **ASIGNATURA:**

Estadística Inferencial

### **DOCENTE:**

Cesar Enrique Barrera Ovando

VILLAHERMOSA, TABASCO, 07 DE MARZO DEL  
2025

## INTRODUCCIÓN

El siguiente ensayo es un análisis de nuestro aprendizaje de los de la unidad uno teorema de limite central y estimación de la materia de estadística inferencial, en los cuales redactare los aprendizajes que obtuvimos, las dificultades a la hora de aprender el software con el cual desarrollamos la primera práctica.

El estudio de la estadística nos permite comprender el comportamiento de los datos y hacer inferencias sobre poblaciones a partir de muestras. En este ensayo, analizaré dos conceptos fundamentales: el Teorema del Límite Central y la estimación. El TLC es una de las piedras angulares de la probabilidad, ya que establece que, bajo ciertas condiciones, la distribución de la media muestral se aproxima a una distribución normal, sin importar la distribución original de los datos. Esto tiene implicaciones cruciales en la inferencia estadística, particularmente en la estimación de parámetros.

A lo largo de este ensayo, explicaré cómo el TLC facilita la aplicación de métodos estadísticos y la construcción de estimadores confiables. También reflexionaré sobre la importancia de la estimación en la toma de decisiones y en la interpretación de datos en diversas áreas del conocimiento.

El primer día de la clase de **Estadística Inferencial** fue una mezcla de emoción y nervios. Al entrar al aula, sentí la expectativa de enfrentar una materia que, según había escuchado, podía ser un reto, pero también muy útil para el análisis de datos

En mi caso yo ya había cursado antes este de la universidad, pero por

cuestiones más es TD nunca llegué a presentar la materia de estadística inferencial por lo tanto nunca supe o tenía conocimiento o de que temas o más o menos una idea de cómo era la materia así que al igual que mis compañeros igual iba a ser mi primera vez usando esta materia.

El profesor comenzó con una presentación que no me había tocado hasta ese momento, el profesor se refirió de una forma muy amable, atento, pero a la vez en lo que a mi corresponde esos son los maestros más difíciles. Porque digo esto pues hubo una regla que dijo “si no gustan estar en la clase con gusto se pueden salir” , y pues yo diga que me puse a ver, claro ya no estamos en preparatoria aquí cada uno se hace responsable de sus acciones y pues eso es lo que si me quedo grabado.

El profesor comenzó con una presentación sobre la importancia de la estadística en la vida real, mencionando aplicaciones en ciencia, economía y hasta en redes sociales. Eso me llamó la atención porque, hasta ahora, solo veía la estadística como un montón de fórmulas y números, pero él la planteó como una herramienta para tomar decisiones basadas en datos.

A medida que hablaba de conceptos como muestras, estimaciones y probabilidades, me di cuenta de que iba a necesitar concentración y práctica para entender bien los temas, anteriormente ya habíamos llevado probabilidad y estadística

y pues no estaba tan perdido Al mismo tiempo, me dio cierta tranquilidad saber que el profesor me explicó las cosas de manera clara y con ejemplos cotidianos.

Otra cosa que igual me llamó la atención qué el profesor pues nos dio permiso de que nosotros pudiéramos elegir la ponderación de cuánto iba a valer tanto el examen las prácticas, hoy recuerdo que nos dio quizás 2 clases y como grupo no nos poníamos de acuerdo porque pues algunos querían una cosa otros otra y no pudimos llegar a un acuerdo y literal tomamos hoy la decisión prácticamente en el último día y a

la hora del salón con la ponderación que tiene ahora siento que puedo salir un poco bien pero pues quién sabe.

Al final de la clase, salí con una mezcla de curiosidad y respeto por la materia. Sé que no será un camino fácil, pero también creo que si me esfuerzo, podré sacar provecho y entender mejor cómo funcionan los datos en el mundo real.

Desde que comencé a estudiar Contaduría Pública, me he dado cuenta de que los números van mucho más allá de simples cálculos. En este proceso, la Estadística Inferencial ha llamado particularmente mi atención porque nos permite tomar decisiones informadas basadas en el análisis de datos. En esta disciplina, no solo se estudian muestras de datos, sino que también se hacen predicciones y estimaciones sobre una población más amplia. Esta capacidad de interpretar datos y hacer inferencias resulta fundamental en el ejercicio contable y en la toma de decisiones empresariales.

En primer lugar, la Estadística Inferencial nos ayuda a analizar tendencias financieras y a predecir comportamientos económicos. Como contadores, muchas veces tenemos que trabajar con datos históricos para proyectar ingresos, costos y presupuestos futuros. Gracias a herramientas como la estimación de intervalos de confianza y las pruebas de hipótesis, podemos evaluar con precisión si un determinado patrón financiero se mantendrá en el futuro o si es necesario tomar medidas correctivas. Esta capacidad es clave para asesorar a empresas y clientes en la planificación financiera y la gestión de riesgos.

Además, la Estadística Inferencial es esencial en la auditoría contable. En una auditoría, generalmente no es posible revisar todas las transacciones de una empresa debido a la gran cantidad de registros. En su lugar, se toman muestras representativas de los datos y, a partir de ellas, se hacen inferencias sobre el conjunto total. Técnicas como el muestreo estadístico permiten a los auditores detectar anomalías o fraudes sin necesidad de revisar cada transacción individualmente. Esto hace que el proceso sea más eficiente y confiable.

Otro aspecto importante en el que la Estadística Inferencial contribuye a la

Contaduría Pública es en el análisis de mercados y la toma de decisiones empresariales. Por ejemplo, al evaluar la viabilidad de una inversión o la aceptación de un nuevo producto, las empresas recopilan datos de clientes y mercados. A través del análisis estadístico, podemos interpretar esos datos para identificar tendencias y estimar la probabilidad de éxito de una estrategia comercial. Esto permite que las empresas tomen decisiones fundamentadas en evidencia numérica y no solo en suposiciones.

También es importante mencionar la relevancia de la Estadística Inferencial en el cumplimiento de normativas y regulaciones. Muchas entidades gubernamentales requieren informes financieros y análisis de riesgos basados en datos estadísticos. Como contadores, debemos estar preparados para aplicar métodos inferenciales que nos permitan cumplir con estos requisitos y garantizar la transparencia en la información financiera.

La Estadística Inferencial es una herramienta indispensable en la Contaduría Pública. Nos ayuda a analizar tendencias, optimizar auditorías, evaluar mercados y cumplir con regulaciones. A medida que avanzo en mi formación, me doy cuenta de que su aplicación va más allá de los cálculos y que realmente nos permite tomar decisiones estratégicas basadas en datos concretos. Sin duda, dominar esta disciplina será clave para mi desarrollo profesional y para contribuir al éxito financiero de las organizaciones con las que trabaje.

uno de los conceptos que más ha llamado mi atención es la **media aritmética**. Al principio, parecía simplemente una operación matemática básica, pero pronto me di cuenta de que su aplicación va mucho más allá de solo sumar y dividir números.

La media es una herramienta fundamental para interpretar datos en diferentes contextos y tomar decisiones basadas en información objetiva.

---

La **media aritmética**, también conocida como **promedio**, es un valor representativo que se obtiene sumando todos los datos de un conjunto y dividiéndolos entre la cantidad total de elementos. Por ejemplo, si en un examen cinco estudiantes

obtienen calificaciones de 70, 80, 90, 85 y 75, la media se calcula sumando estos valores ( $70 + 80 + 90 + 85 + 75 = 400$ ) y dividiendo por el número de estudiantes ( $400 / 5 = 80$ ). Esto significa que, en promedio, los estudiantes obtuvieron una calificación de 80 en el examen.

La media no solo nos ayuda a resumir conjuntos de datos en un solo valor representativo, sino que también es útil en muchas áreas de la vida cotidiana. En economía, por

ejemplo, se usa para calcular el ingreso promedio de una población, lo que permite evaluar el nivel de vida en una región. En educación, los profesores la utilizan para analizar el desempeño de los estudiantes y detectar si es necesario reforzar ciertos temas. Incluso en el deporte, la media ayuda a calcular el rendimiento promedio de un equipo o un atleta en varias competencias.

Una de las cosas más importantes que he aprendido sobre la media es que, aunque es un indicador útil, no siempre cuenta toda la historia de los datos. Si hay valores extremadamente altos o bajos en un conjunto de datos, la media puede verse afectada y no representar con precisión la realidad. Por ejemplo, si en un grupo de trabajadores la mayoría gana entre 1,000 y 2,000 dólares al mes, pero uno de ellos gana 50,000 dólares, la media salarial será mucho más alta de lo que la mayoría realmente percibe. En estos casos, otros valores estadísticos como la mediana pueden ser más adecuados para interpretar la información.

A pesar de sus limitaciones, la media sigue siendo una herramienta esencial para analizar datos de manera rápida y efectiva. Su importancia radica en que nos permite obtener una visión general de un conjunto de valores y compararlos entre sí. Gracias a la media, podemos identificar tendencias, predecir resultados y tomar mejores decisiones en distintos ámbitos.

La media aritmética es un concepto básico pero fundamental en estadística y en la vida diaria. Su capacidad para resumir datos y facilitar el análisis la convierte en una herramienta indispensable en numerosos campos. A medida que continúe mi aprendizaje en estadística, estoy seguro de que seguiré encontrando nuevas formas en las que la media nos ayuda a comprender mejor el mundo que nos rodea.

Al igual que la media la mediana me llamo la atención desde que llevé

probabilidad y estadística aprendí de que hay muchas maneras de analizar un conjunto de datos. Uno de los conceptos que más me ha llamado la atención es la **mediana**, una medida de tendencia central que muchas veces resulta más útil que el promedio.

La mediana es el valor que ocupa la posición central en un conjunto de datos ordenados

de menor a mayor. Si el número de datos es impar, la mediana es el valor

central. Si el número de datos es par, la mediana se obtiene promediando los dos valores centrales. A diferencia de la media aritmética, que puede verse afectada por valores extremos o atípicos, la mediana ofrece una representación más fiel del comportamiento típico de un conjunto de datos.

Podemos utilizar la mediana en muchas situaciones de la vida cotidiana. Por ejemplo, en economía y estudios de ingresos, la mediana salarial es una mejor medida para describir el ingreso típico de una población que la media, ya que evita que valores extremadamente altos o bajos distorsionen la percepción de la realidad. También se usa en análisis de precios de viviendas, en estudios sobre tiempos de espera en servicios y en mediciones de rendimiento académico, entre otros.

En mi experiencia estudiando estadística, me ha parecido interesante cómo la mediana nos ayuda a entender la distribución de los datos sin ser influenciada por valores atípicos. Por ejemplo, en una clase con 30 estudiantes, si la mayoría obtiene calificaciones entre 6 y 8, pero un par de estudiantes sacan un 2 y un 10, la media podría verse afectada y dar una impresión errónea sobre el desempeño del grupo. En cambio, la mediana reflejaría mejor el rendimiento central de los estudiantes.

La importancia de la mediana radica en su capacidad para representar datos de manera justa y equilibrada. En muchos casos, es más confiable que la media, especialmente cuando trabajamos con datos que pueden tener valores extremos. Gracias a la mediana, podemos tomar decisiones más informadas y analizar situaciones de manera más precisa. Por esta razón, considero que es una herramienta fundamental en la estadística y en la interpretación de datos en general.

La moda, para mí, es mucho más que un simple conjunto de prendas o

tendencias. Es una forma de expresión, una manera de mostrarle al mundo quién soy, cómo me siento y qué pienso en determinado momento. A lo largo de la historia, la moda ha sido un reflejo de las sociedades, de los movimientos culturales y de los cambios políticos. Cada época tiene sus

propios estilos y, a través de ellos, podemos leer y entender mucho sobre los contextos en los que vivieron las personas.

La moda no solo tiene que ver con lo que está "in" o con lo que se ve en las pasarelas. Tiene un componente social importante; está vinculada a la identidad y la pertenencia. A través de la ropa, los accesorios, y los colores, muchas veces buscamos encajar en un grupo o, por el contrario, diferenciarnos de los demás. Es un medio para comunicar ideas, valores y creencias sin necesidad de palabras. Puede ser una forma de protesta, como sucedió en varias décadas del siglo XX, donde la moda se convirtió en un símbolo de rebeldía y liberación.

Para mí, la importancia de la moda radica en su capacidad de adaptación y su flexibilidad. No solo se trata de lo que llevamos puesto, sino de cómo nos sentimos al hacerlo. Es una herramienta poderosa para aumentar la confianza en uno mismo, para transformar nuestra percepción y la de los demás. También tiene un impacto en la economía global, generando empleos en el diseño, la manufactura, la distribución y el marketing, lo que hace que sea una de las industrias más grandes del mundo.

La moda tiene un papel fundamental en el empoderamiento personal y colectivo. Nos permite redefinir nuestra imagen constantemente, adaptarnos a nuevas etapas de la vida y experimentar con lo que somos en cada momento. Además, nos conecta con otras personas que comparten gustos, intereses y visiones, creando comunidades que trascienden más allá de lo superficial. En resumen, la moda no es solo un tema de vestimenta, es un reflejo de la sociedad, de nuestra historia, y de la forma en que buscamos vivir, expresarnos y ser vistos.

El teorema del límite central es uno de los pilares fundamentales de la estadística y la probabilidad. A lo largo de mi aprendizaje, llegué a entender que, en

términos sencillos, este teorema establece que cuando tomamos una muestra suficientemente grande de una población, la distribución de las medias muestrales de estas muestras se aproximará a una distribución normal, sin importar la forma de la distribución original de la población.

Es decir, aunque los datos de la población no firman una distribución normal, al promediar los valores de muestras aleatorias y tomar más de ellas, la distribución de estos

promedios tiende a tener una forma de campana, característica de la distribución normal. Este fenómeno ocurre mientras las muestras son grandes, generalmente mayores de 30 elementos, lo que asegura que la distribución se aproxime a la norma.

Lo interesante de este teorema es que no solo se aplica a las distribuciones que conocemos como normales, sino que proporciona una herramienta útil para la inferencia estadística. Es increíble cómo algo tan sencillo, como calcular el promedio de una muestra, puede proporcionarnos una base sólida para hacer predicciones y análisis en una gran variedad de campos, desde la economía hasta la biología. A medida que me adentro más en el tema, me doy cuenta de que el teorema del límite central no solo simplifica el trabajo con datos complejos, sino que también abre la puerta a una mayor comprensión de cómo se comportan los conjuntos de datos cuando se manipulan y analizan a gran escala.

El teorema del límite central es fascinante porque nos demuestra que, independientemente de las características particulares de una población, las muestras grandes tienen un comportamiento predecible, lo que facilita muchísimo el análisis y la toma.

Hoy estoy aprendiendo sobre la distribución de proporción, y aunque al principio me resultó un concepto abstracto, ahora empieza a tomar sentido. La distribución de proporción se refiere a cómo se distribuyen los valores de una proporción dentro de un conjunto de datos. En términos simples, es una forma de estudiar las frecuencias 

---

relativas de un evento en un grupo más grande.

Lo que me ha llamado la atención es que se basa en el concepto de proporciones o probabilidades. Es como analizar la proporción de personas que cumplen con ciertas características dentro de una población. Por ejemplo, si en

una encuesta el 60% de los participantes prefiere un tipo de música, esa proporción se considera una muestra de la población total. Lo interesante es que, cuando hablamos de distribución, estamos observando cómo esa proporción podría comportarse en diferentes muestras si realizáramos el mismo experimento muchas veces.

A medida que avanzo, entiendo que esta distribución tiene ciertas características que

me ayudan a prever cómo podrían distribuirse los resultados en el futuro. Utilizando una distribución normal, puedo predecir con mayor precisión cómo se comportará una proporción en muestras grandes, lo cual tiene una gran utilidad en la estadística inferencial.

Lo más importante que he aprendido hoy es que, a pesar de que no siempre tenemos acceso a toda la población, podemos usar la distribución de proporciones para hacer estimaciones confiables sobre los parámetros de esa población. Es como tomar una pequeña parte y, basándome en esa parte, hacer suposiciones sobre todo el conjunto. Esto me hace sentir más seguro sobre cómo hacer inferencias y tomar decisiones basadas en datos.

La distribución de proporción me ha abierto los ojos sobre cómo se puede trabajar con datos categóricos y probabilidades, y aunque al principio parecía complicado, me doy cuenta de que es una herramienta esencial para entender patrones y tendencias en diferentes contextos.

Hoy me encuentro aprendiendo sobre un tema que, hasta ahora, había oído mencionar de manera superficial, pero que realmente está comenzando a captar mi atención: las propiedades de los estimadores. Nunca imaginé que los estimadores pudieran ser tan fundamentales en estadística y en la inferencia de datos. Al principio, me sentí un poco perdido entre los términos técnicos, pero conforme fui avanzando, empecé a entender por qué son tan importantes y cómo impactan en los resultados de un análisis.

Los estimadores son básicamente fórmulas o reglas que nos permiten obtener una aproximación de un parámetro desconocido a partir de una muestra de datos.

Como ejemplo, cuando estamos estimando la media de una población a partir de una

muestra, el estimador que usamos es la media muestral. Sin embargo, no todos los estimadores son igualmente buenos, y aquí es donde entran las propiedades que los definen como buenos o malos.

Una de las propiedades clave de un buen estimador es **la imparcialidad**. Esto significa que, en promedio, el estimador no debería sobrestimar ni subestimar el valor real del parámetro. Lo que aprendí es que un estimador imparcial es uno cuyo valor esperado coincide con el valor verdadero del parámetro que estamos estimando. Esto es algo crucial, porque si el estimador siempre se desvía de la verdad, cualquier conclusión que saquemos podría ser

incorrecta.

Otra propiedad esencial es **la eficiencia**. Aquí entendí que un estimador eficiente es aquel que tiene la menor varianza entre todos los estimadores imparciales posibles. En otras palabras, un estimador eficiente proporciona estimaciones que, en promedio, están más cerca del valor real del parámetro, lo cual es una ventaja enorme cuando trabajamos con grandes cantidades de datos. No solo nos interesa que el estimador sea imparcial, sino que también queremos que sea lo más preciso posible.

La **consistencia** es otra propiedad que se volvió clara para mí. Esta se refiere a que, a medida que aumenta el tamaño de la muestra, el estimador debería acercarse al valor real del parámetro. Por ejemplo, si tengo una muestra muy pequeña, la estimación podría ser muy imprecisa, pero a medida que recojo más datos, la estimación debería volverse más precisa. Esta propiedad me pareció interesante porque resalta cómo las muestras grandes tienden a ofrecer mejores estimaciones.

Finalmente, la **suficiencia** es una propiedad que me costó un poco más entender. Un estimador es suficiente si recoge toda la información relevante del parámetro que está estimando. Es decir, no se puede obtener más información sobre el parámetro a partir de los datos si ya tenemos este estimador. Esta propiedad me parece interesante porque implica que no necesitamos más datos o procedimientos complejos para estimar el parámetro de manera adecuada, lo cual puede ser muy útil en situaciones prácticas.

Aunque todavía estoy en proceso de profundizar en estos conceptos, ahora tengo una visión más clara sobre cómo se evalúa un estimador y por qué son tan cruciales en el análisis de datos. Comprender estas propiedades me ayudará, sin duda, a tomar decisiones más informadas cuando se trate de elegir un estimador

adecuado para mis propios proyectos estadísticos.

Hoy me he adentrado en el uso del software SPSS para trabajar con estadísticas, y debo decir que la experiencia ha sido reveladora, aunque también un poco desafiante. Al principio, el programa parecía abrumador con su interfaz repleta de opciones y herramientas, pero al

familiarizarme con las funciones básicas, me fui sintiendo más cómodo. Lo primero que aprendí fue cómo importar datos; tener claro cómo organizar mi información en las hojas de datos fue clave para que todo funcionara correctamente. Además, la opción de etiquetas para variables me pareció muy útil, ya que facilita la interpretación de los datos, especialmente cuando se manejan grandes volúmenes de información.

Pero lo que realmente me sorprendió fue la posibilidad de realizar pruebas estadísticas más complejas. No voy a mentir, al principio las opciones avanzadas me parecieron confusas.

A pesar de las dificultades iniciales, me he dado cuenta de lo poderoso que es SPSS. Cada herramienta y cada función tienen un propósito claro, y a medida que lo voy explorando, me doy cuenta de lo esencial que puede llegar a ser para cualquier investigación o trabajo que implique análisis de datos. No me queda más que seguir practicando y aprendiendo nuevas técnicas. Con el tiempo, estoy seguro de que podré aprovechar al máximo todo el potencial que ofrece este software.

---

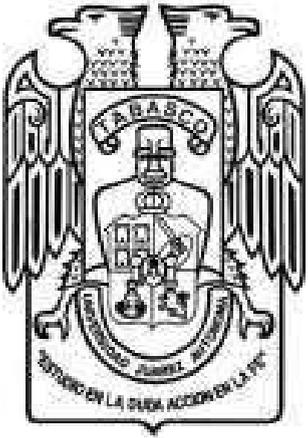
## CONCLUSIÓN

En conclusión, al estudiar el Teorema del Límite Central y realizar simulaciones, me he dado cuenta de la importancia crucial que tiene este teorema en el campo de la estadística.

Comprendí que, independientemente de la distribución original de los datos, cuando se realizan suficientes muestras, la distribución de la media de esas muestras tiende a ser normal. Esto tiene implicaciones profundas, ya que nos permite hacer inferencias sobre poblaciones más grandes a partir de muestras relativamente pequeñas.

Las simulaciones fueron clave para internalizar este concepto, ya que me permitieron visualizar cómo la distribución de las medias se vuelve cada vez más cercana a una normal a medida que aumentaba el tamaño de la muestra. A través de este proceso, también pude ver cómo el tamaño de la muestra y la varianza afectan el comportamiento de la distribución de la media, lo que me ayudó a comprender mejor los aspectos prácticos del Teorema del Límite Central en investigaciones y análisis de datos reales. Sin duda, esta experiencia ha fortalecido mi entendimiento sobre la teoría estadística y su aplicación en diversas situaciones.

---



**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



***“MAPAS MENTALES”***

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

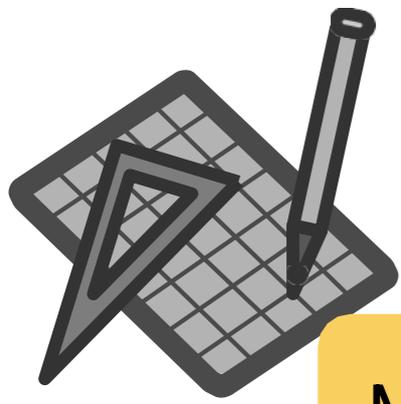
**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 2025**



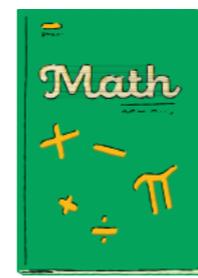
## MEDIANA

es un conjunto es un valor que se encuentra a la mitad de los otros valores



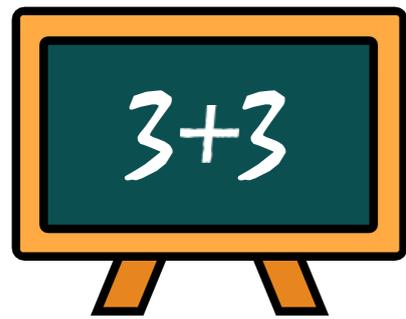
## MEDIA

es el valor que se obtiene al dividir la suma de un conglomerado de números entre la cantidad de ellos.

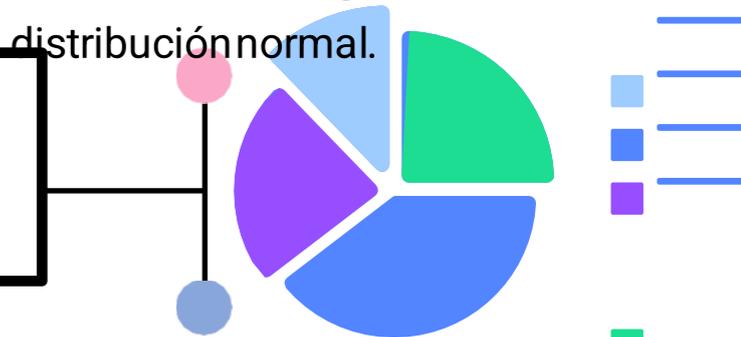


## TEOREMA DEL LÍMITE CENTRAL

es una teoría estadística que establece que, dada una muestra aleatoria suficientemente grande de la población, la distribución de las medias muestrales seguirá una distribución normal.



## TEOREMA DEL LÍMITE CENTRAL Y ESTIMACIÓN



## MOD

es el valor que aparece más dentro de un conjunto de datos. A diferencia de la media y la mediana, la moda no requiere valores numéricos y puede utilizarse con datos categóricos o discretos.

## LAS MEDIDAS DE TENDENCIA CENTRAL

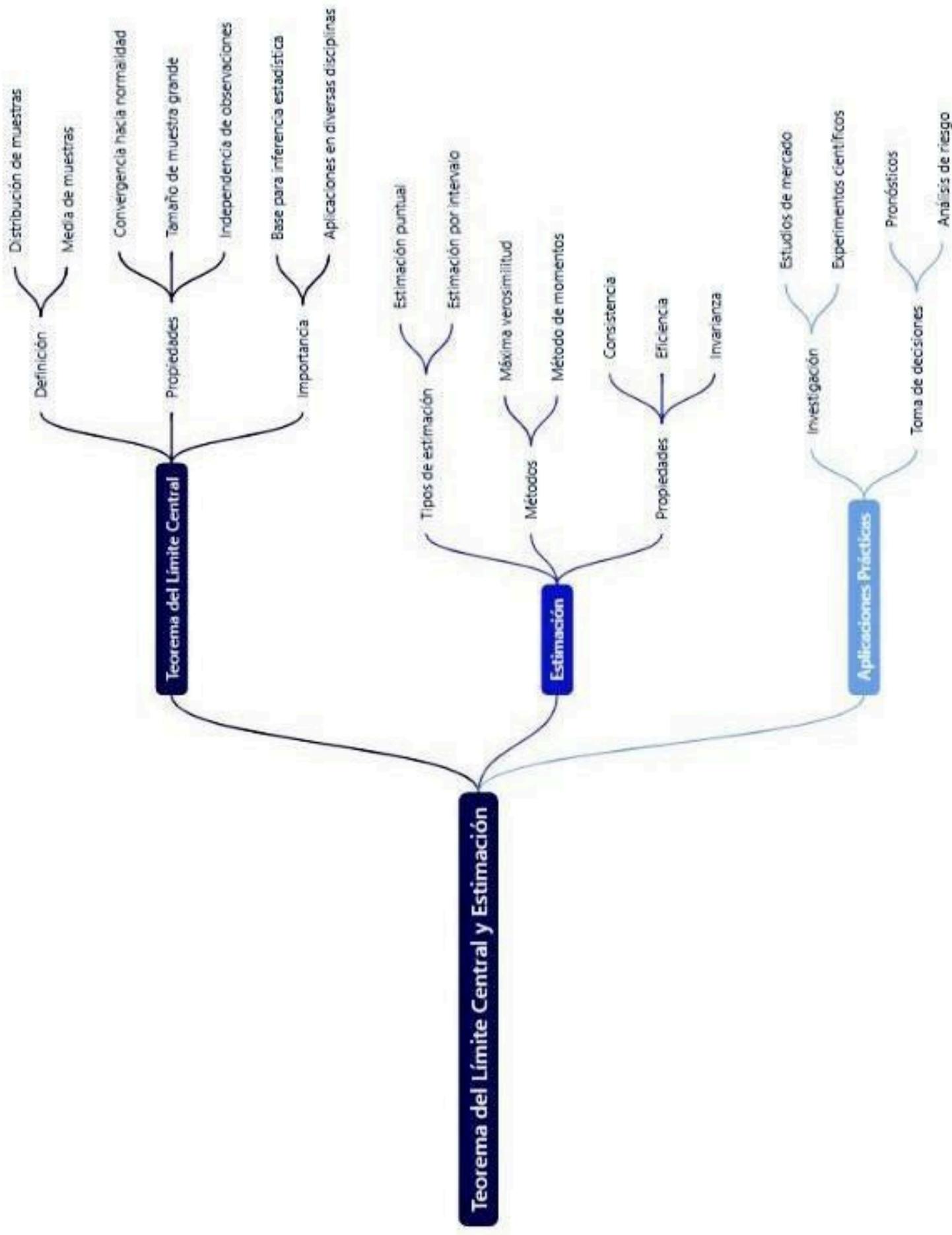
son herramientas fundamentales en estadística que permiten resumir un conjunto de datos en un solo valor representativo.

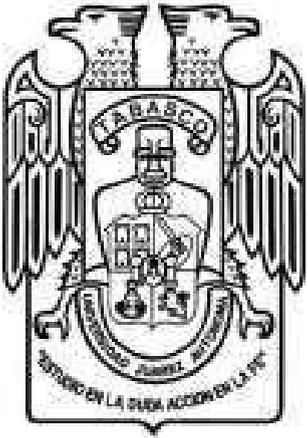


## DISTRIBUCIÓN DE LA PROPORCIÓN

es un concepto fundamental en estadística, utilizado para describir la variabilidad de una proporción muestral en relación con la población.







**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**  
**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



**“DOCUMENTO EXTENSO”**

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 2025**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



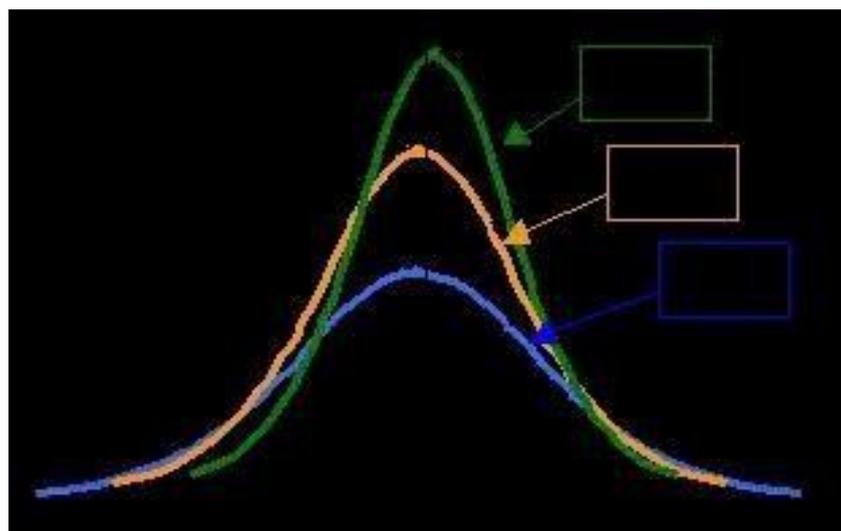
### 1.1 Distribución muestral de la media. “Teorema del límite central.”

El teorema central del límite (TCL) es una teoría estadística que establece que, dada una muestra aleatoria suficientemente grande de la población, la distribución de las medias muestrales seguirá una distribución normal.

Además, el TCL afirma que a medida que el tamaño de la muestra se incrementa, la media muestral se acercará a la media de la población de la proporción.

El teorema central del límite es un resultado matemático que garantiza que, si sumamos variables cualesquiera (no necesariamente normales), la variable suma también seguirá una distribución normal (esto siempre que se cumplan algunas condiciones básicas).

Así, cuando un dato o resultado es la suma de contribuciones independientes, de igual magnitud y “con un tamaño típico”, este resultado corresponderá a una distribución Gaussiana siempre que el número de contribuciones (el número de sumandos) sea un número considerable (no pequeño).





# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Con un tamaño típico se quiere garantizar que las contribuciones tienen que “estar controladas” , esto es, las contribuciones extremas tienen que estar controladas por una probabilidad muy pequeña (En jerga matemática las contribuciones tiene que tener varianza finita).

Este teorema asegura, de manera esquemática, que, cuando sumamos un número grande de variables, la variable resultante sigue una distribución normal.

De manera general, si  $X_1, X_2, \dots, X_n$  son variables de media o esperanza  $\mu_i = E(X_i)$  y varianza  $\sigma_i^2 = \text{Var}(X_i)$ ,  $i = 1, \dots, n$ , se verifica que la variable suma  $Y = X_1 + X_2 + \dots + X_n$  (sin es un número tendiendo a infinito) se puede aproximar por una variable normal, de media la suma de las medias y varianza la suma de varianzas (desviación típica raíz de la suma de varianzas), es decir

### Principales propiedades del teorema central del límite

El teorema central del límite tiene una serie de propiedades de gran utilidad en el ámbito estadístico y probabilístico. Las principales son:

Si el tamaño de la muestra es suficientemente grande, la distribución de las medias muestrales seguirá aproximadamente una distribución normal. El TCL considera una muestra como grande cuando el tamaño de la misma es superior a 30. Por tanto, si la muestra es superior a 30, la media muestral tendrá una función de distribución

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**

próxima a una normal. Y esto se



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



independientemente de la forma de la distribución con la que estamos trabajando.

La media poblacional y la media muestral serán iguales. Es decir, la media de la distribución de todas las medias muestrales será igual a la media del total de la población.

La varianza de la distribución de las medias muestrales será  $\sigma^2/n$ .

Que es la varianza de la población dividido entre el tamaño de la muestra.

Que la distribución de las medias muestrales se parezca a una normal es tremendamente útil, Porque la distribución normal es muy fácil de aplicar para realizar contrastes de hipótesis y construcción de intervalos de confianza. En estadística que una distribución sea normal es bastante importante, dado que muchos estadísticos requieren este tipo de distribución. Además, el TCL nos permitirá hacer inferencia sobre la media poblacional a través de la media muestral. Y esto es de gran utilidad cuando por falta de medios no podemos recolectar datos de toda una población.

### IMPORTANCIA Y CONSECUENCIAS DEL TEOREMA DEL LÍMITE CENTRAL

Se trata de una propiedad muy importante en la estadística porque permite justificar y resolver diversos problemas claves. En este documento se presenta primero el enunciado del Teorema y se analizan a continuación algunas de sus

# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

consecuencias.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### ¿Qué es la media, la mediana y la moda?

La media es la media aritmética de un conjunto de valores numéricos. La mediana es el valor medio de un conjunto de datos cuando los valores se ordenan de forma ascendente o descendente. La moda representa el valor o categoría más común dentro del conjunto de datos.

La media, la mediana y la moda son las tres medidas de tendencia central más usadas para poblaciones que no cuentan con demasiados datos, es decir, que no necesitan agruparse.

Al hablar de medidas de tendencia central, nos referimos a medidas estadísticas que pretenden resumir en un único valor a un conjunto de valores.

La media, mediana y moda se expresan en la misma unidad que los datos originales. Estas medidas proporcionan información sobre el valor central o típico de un conjunto de datos, ayudándonos a analizar y comparar diferentes puntos de datos.

### ¿Qué es la media?

La media, también conocida como promedio, es el valor que se obtiene al dividir la suma de un conglomerado de números entre la cantidad de ellos.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



La media representa el punto de equilibrio de la distribución y está influida por los valores extremos. Proporciona una medida de la tendencia general o valor medio de los datos.

Se calcula sumando todos los valores del conjunto de datos y dividiendo la suma por el número total de puntos de datos.

### **Algunas características de la media son:**

- Considera todas las puntuaciones
- El numerador de la fórmula es la cantidad de valores
- Cuando hay puntuaciones extremas, no tiene una representación exacta de la muestra

Cómo sacar la Media

**Para obtener la Media de un conjunto solo tienes que seguir estos**

**sencillos pasos:**

1. Determina el conjunto de valores que buscas promediar.
2. Suma los valores para obtener el total
3. Haz el conteo de la cantidad de valores en el conjunto.

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

4. Divide la suma del conjunto entre la cantidad de números.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA



## ¿Qué es la mediana?

La mediana es un conjunto es un valor que se encuentra a la mitad de los otros valores, es decir, que al ordenar los números de menor a mayor, éste se encuentra justamente en medio entre los que están por arriba.

Se calcula sumando todos los valores del conjunto de datos y dividiendo la suma por el número total de puntos de datos.

La media representa el punto de equilibrio de la distribución y está influida por los valores extremos. Proporciona una medida de la tendencia general o valor medio de los datos.

## Algunas características de la media son:

- Las operaciones para calcular el valor son muy sencillas de realizar.
- La medida no depende de los valores de las variables, solamente de su orden.
- Generalmente, los valores son enteros.
- Se puede calcular aunque los números que se encuentren arriba y abajo no tengan límites.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### Como sacar la Mediana

Los pasos para sacar la mediana son:

1. Ordena todos los números del más pequeño al más grande.
  2. Encuentra el número del medio del conjunto.
- Si tienes una cantidad impar: Tacha el número al final de la izquierda, después el primero a la derecha, y repite el proceso hasta quedarte con un número, que será la mediana.
  - Si tienes una cantidad par, al final quedarás con dos números en el centro. Súmalos y divídelos entre 2 para obtener la mediana.

### Ejemplo de Mediana

- La cantidad de valores es impar

Si se tienen los valores: **9,5,4,2,7**, se ordenan: **2, 4, 5, 7, 9**. El elemento de en medio es el **5**, ya que se encuentra dos valores por encima y dos valores por debajo.

- La cantidad de valores es par

Si se tienen los valores **9,5,4,2**, se ordenan: **2,4,5,9**. En este caso se toman los dos valores centrales **5 y 4**, la mediana es el promedio de ambos: **9**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### ¿Qué es la moda?

La moda es el valor que aparece más dentro de un conjunto de datos. A diferencia de la media y la mediana, la moda no requiere valores numéricos y puede utilizarse con datos categóricos o discretos.

Un conjunto de datos puede tener un modo, conocido como unimodal, o varios modos, denominados bimodal o multimodal. Se llama amodal cuando en un conglomerado no se repiten los valores.

### Las principales características de la moda son:

- Es una muestra muy clara
- Las operaciones para determinar el resultado son muy fáciles de elaborar
- Los valores que se presentan pueden ser [cualitativos y cuantitativos](#)

### Como sacar la Moda

Los pasos para obtener la moda de un conjunto son:

- Escribe todos los números del conjunto.
- Encuentra el número o los números (en los casos bimodales o multimodales) que aparezcan más veces.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



## Ejemplo de Moda

Moda	2,5,5,7,9,10	$\hat{M} = 5$
Bimodal	2,3,3,5,7,8,9,9	$\hat{M} = 3,9$
Multimodal	2,3,3,5,7,7,8,9,9	$\hat{M} = 3,7,9$
Amodal	2,4,5,7,9	$\hat{M} = \text{No}$
Adyacente	2, $\frac{3+3+5+5}{4}$ , 7, 8	$\hat{M} = 4$

## ¿Cuándo se utiliza la media, la mediana y la moda?

La media suele ser la medida de tendencia central más usada por su alto grado de utilidad para múltiples contextos.

Sin embargo, cuando en una población se presentan casos que cuentan con datos muy por arriba o por debajo de los que presenta el resto del grupo, se recomienda usar la mediana o la moda, pues la media es más afectada por los valores extremos.

**La importancia de las medidas de tendencia central y sus usos en la vida diaria**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Las medidas de tendencia central son herramientas fundamentales en estadística que permiten resumir un conjunto de datos en un solo valor representativo. Entre las más utilizadas se encuentran la media, la mediana y la moda. Cada una de estas medidas tiene características específicas que las hacen útiles en distintos contextos. Su importancia radica en su capacidad para facilitar el análisis de datos, la toma de decisiones y la identificación de patrones en diversas áreas de la vida cotidiana.

En primer lugar, la media aritmética es la medida más comúnmente utilizada. Se obtiene sumando todos los valores de un conjunto de datos y dividiendo el resultado entre el número total de observaciones. Su principal ventaja es que proporciona un valor equilibrado, pero puede verse afectada por valores atípicos que distorsionen el resultado. En la vida diaria, la media se emplea en múltiples ámbitos, como el cálculo del promedio de calificaciones escolares, el análisis de salarios en una empresa o la evaluación del rendimiento de un deportista a lo largo de una temporada.

Por otro lado, la mediana es el valor central de un conjunto de datos ordenados de menor a mayor. Es especialmente útil cuando existen valores extremos que pueden sesgar la media. En la vida cotidiana, la mediana se emplea en la economía y en el análisis de ingresos, ya que permite obtener una

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

representación más precisa de la distribución de los salarios dentro de una



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



población. También se utiliza en la evaluación del precio de viviendas, ayudando a determinar un valor más representativo del mercado.

La moda, en cambio, es el valor que aparece con mayor frecuencia en un conjunto de datos. Su utilidad radica en su capacidad para identificar tendencias y preferencias. En la vida diaria, la moda se usa en el ámbito del comercio y el marketing para analizar qué productos o servicios son más demandados por los consumidores. También es empleada en estudios de opinión pública para conocer la preferencia de la población sobre un tema específico.

Las medidas de tendencia central no solo permiten resumir información, sino que también facilitan la toma de decisiones fundamentadas. En sectores como la educación, la salud, la economía y la investigación científica, estas herramientas ayudan a interpretar datos y generar conclusiones valiosas. Por ejemplo, en el sector de la salud, los promedios de presión arterial y niveles de glucosa ayudan a diagnosticar y monitorear enfermedades. En el ámbito empresarial, conocer el salario promedio de los empleados permite establecer estrategias salariales más equitativas.

En conclusión, las medidas de tendencia central son fundamentales para el análisis de datos y la toma de decisiones en diversas áreas de la vida cotidiana.

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

Su correcta aplicación permite comprender mejor la información,



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



identificar patrones y realizar comparaciones significativas. Al emplearlas de manera adecuada, se pueden optimizar procesos y mejorar la toma de decisiones tanto a nivel individual como colectivo.

### **Distribución de la Proporción**

La distribución de la proporción es un concepto fundamental en estadística, utilizado para describir la variabilidad de una proporción muestral en relación con la población total. En muchas aplicaciones de la estadística inferencial, se emplea para estimar la fracción de individuos en una población que poseen una determinada característica, basándose en una muestra aleatoria.

### **Concepto de Proporción y su Distribución**

Una proporción es una medida de frecuencia relativa que se obtiene dividiendo el número de elementos de una muestra que cumplen con una determinada condición entre el tamaño total de la muestra. Matemáticamente, se expresa como:

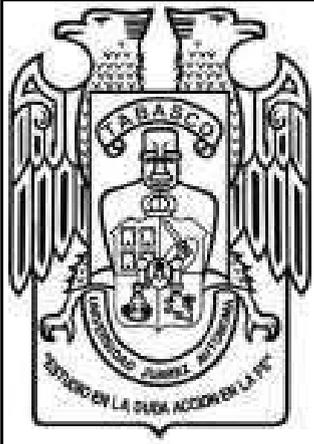
donde:

- es la proporción muestral,

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

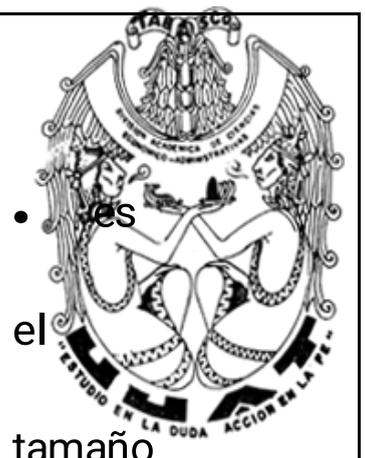
## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

- es el número de éxitos en la muestra,



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



el tamaño

de la

muestra.

Cuando se toma una muestra aleatoria de una población y se calcula la proporción de éxitos, esta proporción puede variar de una muestra a otra, lo que da lugar a la distribución de la proporción muestral.

### **Distribución Normal Aproximada de la Proporción**

Si el tamaño de la muestra es suficientemente grande, la distribución de la proporción muestral sigue aproximadamente una distribución normal debido al Teorema del Límite Central. La media y la desviación estándar de esta distribución se pueden calcular como:

donde  $p$  es la proporción de la población. Para que esta aproximación sea válida, es necesario que se cumpla la regla de normalidad:

Si estas condiciones no se cumplen, la distribución de la proporción muestral puede ser asimétrica y no se podría modelar con una distribución normal.

### **Aplicaciones de la Distribución de la Proporción**

La distribución de la proporción es ampliamente utilizada en estadística

# UNIVERSIDAD JUÁREZ AUTÓNOMA

para la toma de decisiones. Algunas aplicaciones incluyen:

## DE TABASCO



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



1. **Estimación de Proporciones Poblacionales:** Se utiliza para hacer inferencias sobre una población con base en una muestra. Por ejemplo, en encuestas electorales, se estima el porcentaje de votantes que prefieren a un candidato.
2. **Pruebas de Hipótesis:** Se emplea en pruebas de significancia para determinar si la proporción observada en una muestra difiere significativamente de una proporción teórica.
3. **Control de Calidad:** En la industria, se usa para verificar si la proporción de defectos en un lote de productos está dentro de los límites aceptables.

La distribución de la proporción es una herramienta clave en el análisis estadístico, ya que permite realizar inferencias sobre una población a partir de muestras. Su aplicación en diversas áreas, como la investigación social, la economía y la ingeniería, demuestra su importancia para la toma de decisiones basadas en datos. Comprender su comportamiento y los criterios que garantizan su normalidad es fundamental para obtener conclusiones válidas y precisas.

### Propiedades de los Estimadores

En estadística inferencial, los estimadores desempeñan un papel

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

fundamental en la obtención de información sobre una población a partir de una



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



muestra. Un estimador es una función de los datos muestrales utilizada para aproximar un parámetro desconocido de la población. Para evaluar la calidad de un estimador, es necesario considerar diversas propiedades que permiten determinar su eficacia y precisión. A continuación, se presentan las principales propiedades de los estimadores.

### 1. Insesgadez

Un estimador es insesgado si su esperanza matemática es igual al parámetro que se desea estimar. En términos formales, si es un estimador del parámetro poblacional, entonces el estimador es insesgado si:

Si el estimador no cumple esta propiedad, se dice que está sesgado. Sin embargo, en algunos casos, un estimador sesgado puede ser preferible si su sesgo es pequeño y su varianza es reducida.

### 2. Consistencia

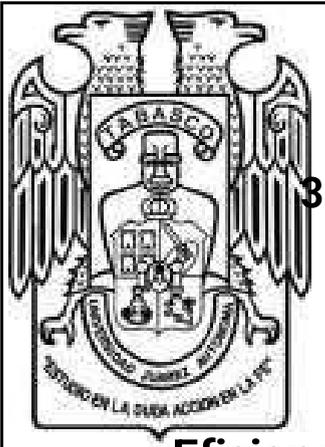
Un estimador es consistente si, al aumentar el tamaño de la muestra, la probabilidad de que el estimador se acerque al verdadero valor del parámetro tiende a 1. Es decir, matemáticamente se expresa como:

La consistencia es una propiedad deseable, ya que garantiza que con muestras suficientemente grandes, el estimador proporciona valores cercanos

**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**

al parámetro real de la población.



3.

# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### **Eficiencia**

Un estimador es eficiente si tiene la menor varianza posible entre todos los estimadores insesgados del parámetro. Es decir, si  $\hat{\theta}_1$  y  $\hat{\theta}_2$  son dos estimadores insesgados de  $\theta$ , entonces  $\hat{\theta}_1$  es más eficiente que  $\hat{\theta}_2$  si:

Un estimador eficiente proporciona estimaciones más precisas y reduce la incertidumbre en la inferencia estadística.

### **4. Suficiencia**

Un estimador es suficiente si contiene toda la información relevante sobre el parámetro que se desea estimar. En otras palabras, un estimador suficiente no desperdicia información contenida en la muestra. Formalmente, un estimador es suficiente para  $\theta$  si la distribución condicional de la muestra, dada  $\hat{\theta}$ , es independiente de  $\theta$ . El teorema de factorización de Neyman-Fisher proporciona un criterio para determinar si un estimador es suficiente.

### **5. Robustez**

Un estimador es robusto si su desempeño no se ve afectado significativamente por desviaciones de los supuestos teóricos, como la presencia de valores atípicos o distribuciones no normales. Los estimadores robustos son especialmente útiles en la práctica, donde los datos pueden no

**UNIVERSIDAD JUÁREZ AUTÓNOMA**

**DE TABASCO**

ajustarse perfectamente a los modelos teóricos.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Las propiedades de los estimadores son fundamentales para evaluar su calidad y utilidad en la estimación de parámetros poblacionales. Un buen estimador debe ser insesgado, consistente, eficiente, suficiente y robusto, dependiendo de la aplicación específica. En la práctica, a menudo se busca un balance entre estas propiedades para seleccionar el mejor estimador posible en función de las características de los datos y los objetivos del análisis estadístico.



**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



**“PRACTICAS DE LABORATORIO DE  
COMPUTO”**

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 202**

## INTRODUCCIÓN

El presente documento tiene como objetivo servir como evidencia de los temas abordados en la primera unidad, titulada "Teorema del Límite Central y Estimación", de la materia de Estadística Inferencial. Para ello, se llevó a cabo una práctica en el centro de cómputo de la universidad, la cual debía resolverse utilizando el software IBM SPSS Statistics.

El documento incluye una recopilación de los ejercicios realizados con variables de un conjunto de datos, así como capturas de pantalla que muestran los resultados obtenidos, incluyendo tablas y gráficos generados por el software. Además, se presenta un informe que contiene la interpretación de los resultados obtenidos durante el análisis.

SPSS Statistics es un software de análisis de datos que podemos utilizar para llevar a cabo tareas estadísticas y de investigación. Con él, podemos gestionar y analizar grandes volúmenes de datos, realizar análisis descriptivos, crear gráficos y aplicar modelos estadísticos complejos. Nos permite trabajar con diferentes tipos de datos, como encuestas, experimentos o bases de datos numéricas, y tiene herramientas que facilitan el análisis de regresión, ANOVA, correlaciones, y más. También tiene una interfaz gráfica que me permite realizar estos análisis sin necesidad de ser experto.



		Estadísticos											
		Numero de elemento	Edad del Encuestado	Estado civil del encuestado	Genero del encuestado	Ocupacion del encuestado	Peso del encuestado	Estatura del encuestado	Taller	Sin libro	Depresion	Nivel Socioeconómico	Puntaje Malestar Emocional
N	Válido	20	20	20	20	20	20	20	20	10	20	18	18
	Perdidos	0	0	0	0	0	0	0	0	10	0	2	2
Media		10.50	26.00	1.90	1.35	2.05	69.15	1.6350	1.50	2.00	3.80	2.00	27.72
Error estándar de la media		1.323	1.606	.216	.109	.285	2.912	.02377	.115	.000	.490	.198	.939
Mediana		10.50	24.00	2.00	1.00	1.50	64.50	1.6300	1.50	2.00	3.50	2.00	27.00
Moda		1 <sup>a</sup>	18 <sup>a</sup>	1 <sup>a</sup>	1	1	85	1.60	1 <sup>a</sup>	2	3	1 <sup>a</sup>	27 <sup>a</sup>
Desv. estándar		5.916	7.182	.968	.489	1.276	13.023	.10630	.513	.000	2.191	.840	3.982
Varianza		35.000	51.579	.937	.239	1.629	169.608	.011	.263	.000	4.800	.706	15.859
Asimetría		.000	.776	.991	.681	.741	.463	1.243	.000		-.117	.000	-.059
Error estándar de asimetría		.512	.512	.512	.512	.512	.512	.512	.687	.512	.536	.536	.536
Curtosis		-1.200	-.652	.335	-1.719	-1.218	-.926	3.339	-2.235		-.814	-1.594	-.473
Error estándar de curtosis		.992	.992	.992	.992	.992	.992	.992	1.334	.992	1.038	1.038	1.038
Rango		19	22	3	1	3	45	.47	1	0	7	2	15
Mínimo		1	18	1	1	1	50	1.48	1	2	0	1	20
Máximo		20	40	4	2	4	95	1.95	2	2	7	3	35
Suma		210	520	38	27	41	1383	32.70	30	20	76	36	499
Percentiles	25	5.25	20.00	1.00	1.00	1.00	59.25	1.6000	1.00	2.00	2.25	1.00	24.75
	50	10.50	24.00	2.00	1.00	1.50	64.50	1.6300	1.50	2.00	3.50	2.00	27.00
	75	15.75	31.50	2.00	2.00	3.75	83.50	1.6750	2.00	2.00	5.75	3.00	30.50

a. Existen múltiples modos. Se muestra el valor más pequeño.

Esta tabla presenta estadísticas descriptivas sobre diferentes variables en una muestra de 20 encuestados. La edad promedio de los encuestados es de 26 años, con una desviación estándar de 5.916. En cuanto al estado civil, género y ocupación, la moda sugiere que la mayoría de los encuestados comparten categorías similares en estas variables, El peso promedio es de 89.15 kg, y la estatura media es de 1.677 m entre los encuestados.

En los Aspectos Académicos y Psicológicos en la Asistencia a taller la media es 1.35, lo que sugiere que más personas han asistido superando a las que no, en cuanto los Acceso a libros la media de 1.50 indica una distribución equitativa entre quienes tienen y quienes no tienen libros. Y en la parte de Depresión la media de 1.80 nos sugiere que la mayoría de los encuestados no presentan depresión, aunque hay variabilidad en las respuestas debido a posibles casos en el Nivel socioeconómico el promedio es 2.80 en una escala donde 3 parece ser un valor modal, lo que implica una tendencia hacia la clase media. Con la variable de Malestar emocional la media de 27.2 sugiere un nivel moderado, con una variabilidad importante (desviación estándar de 8.57).

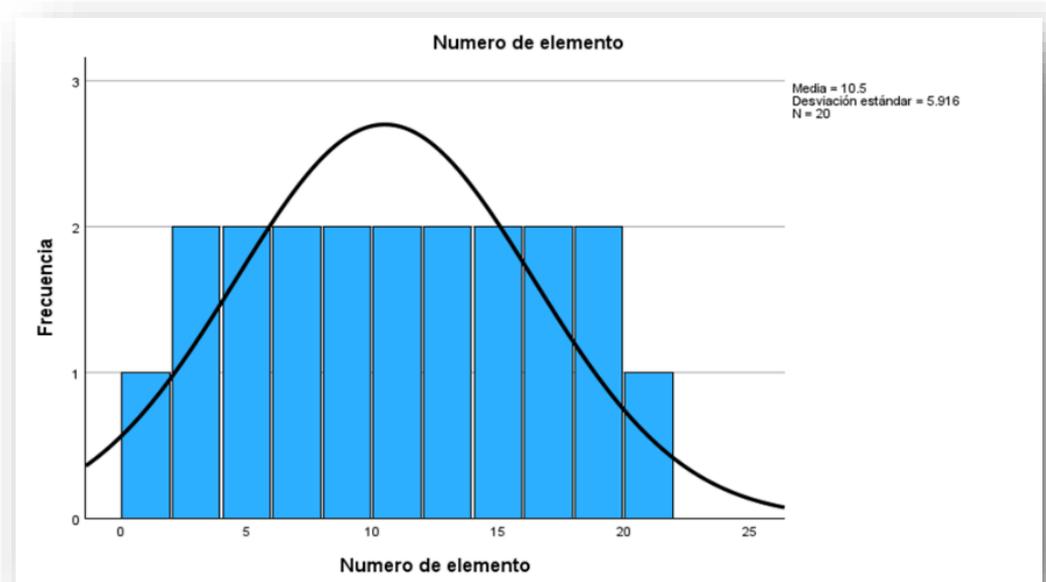
Si vemos los datos de Distribución y las Tendencias la asimetría y curtosis en varias variables indican que algunas distribuciones pueden no ser normales, el peso y depresión tienen valores más extremos. Los percentiles muestran que el 25% de los encuestados tiene un nivel de

malestar emocional de 20 o menos, mientras que el 75% lo tiene en 30 o menos, lo que indica que hay algunos valores más altos que elevan la media.

En general, esta tabla brinda una visión general del perfil de los encuestados, destacando patrones sociodemográficos y emocionales.

## 1. Numero de elemento (Tabla de frecuencia y histograma).

Numero de elemento				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	1	5.0	5.0	5.0
	2	5.0	5.0	10.0
	3	5.0	5.0	15.0
	4	5.0	5.0	20.0
	5	5.0	5.0	25.0
	6	5.0	5.0	30.0
	7	5.0	5.0	35.0
	8	5.0	5.0	40.0
	9	5.0	5.0	45.0
	10	5.0	5.0	50.0
	11	5.0	5.0	55.0
	12	5.0	5.0	60.0
	13	5.0	5.0	65.0
	14	5.0	5.0	70.0
	15	5.0	5.0	75.0
	16	5.0	5.0	80.0
	17	5.0	5.0	85.0
	18	5.0	5.0	90.0
	19	5.0	5.0	95.0
	20	5.0	5.0	100.0
Total	20	100.0	100.0	



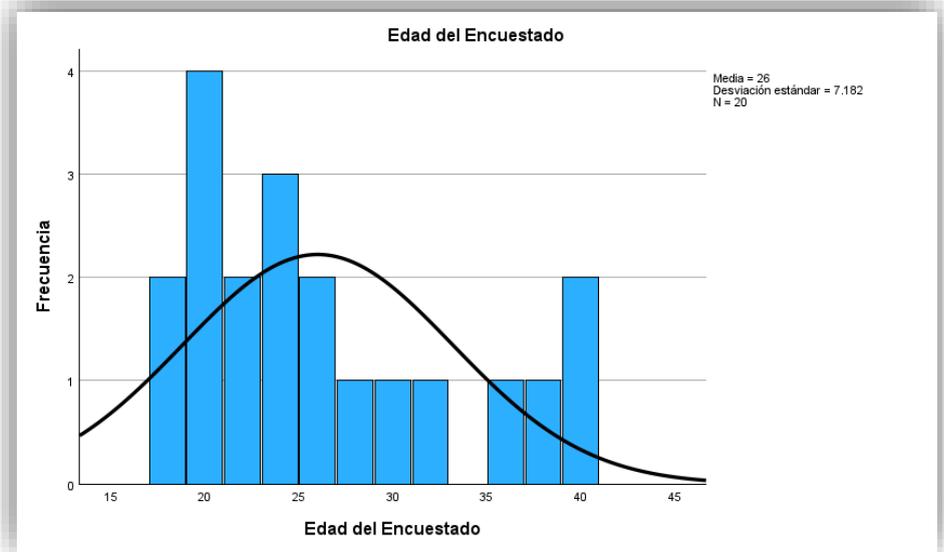
Esta tabla muestra la distribución de una variable categórica, donde cada número de elemento tiene la misma frecuencia de aparición. Como cada valor aparece exactamente una vez en la muestra podemos decir que su (frecuencia = 1). Al igual observamos que como hay 20 elementos en total, cada uno representa el 5% del total siendo este su porcentaje. El porcentaje válido también es del 5% para cada categoría, ya que no hay datos perdidos y el porcentaje acumulado muestra cómo se van sumando estos valores, alcanzando el 100% en el último elemento.

Con esto concluimos que, al no ver valores repetidos, esta variable es completamente uniforme y equitativa en la muestra. Dando la posibilidad de que se trate

de una identificación numérica o una variable que no presenta concentración en ningún valor específico.

## 2. Edad del encuestado (Tabla de frecuencia y histograma).

Edad del Encuestado				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	18	2	10.0	10.0
	19	2	10.0	20.0
	20	2	10.0	30.0
	21	1	5.0	35.0
	22	1	5.0	40.0
	23	1	5.0	45.0
	24	2	10.0	55.0
	25	1	5.0	60.0
	26	1	5.0	65.0
	28	1	5.0	70.0
	30	1	5.0	75.0
	32	1	5.0	80.0
	35	1	5.0	85.0
	37	1	5.0	90.0
	39	1	5.0	95.0
	40	1	5.0	100.0
Total	20	100.0	100.0	



### ITEM: NINGUNo

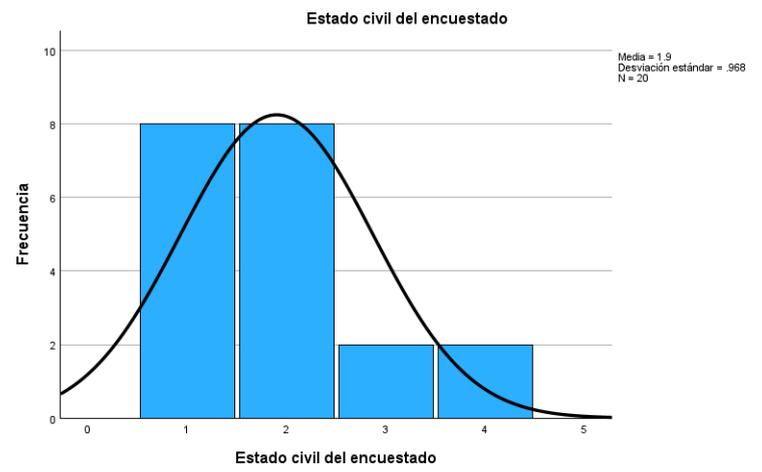
La tabla muestra la distribución de edades de los encuestados en el estudio.

Observó que las edades van desde los 18 hasta los 40 años, con una representación bastante variada. La mayoría de las edades aparecen solo una vez, lo que indica que no hay una concentración específica en un solo grupo etario. Solo la edad de 18 años se repite dos veces, siendo el 10% de la muestra, mientras que el resto de las edades tienen una frecuencia del 5%. Esto sugiere que, aunque la muestra es diversa, hay una ligera presencia de encuestados más jóvenes.

## 3. Estado civil (Tabla de frecuencia y histograma).

**Estado civil del encuestado**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Soltero	8	40.0	40.0	40.0
	Casado	8	40.0	40.0	80.0
	Union libre	2	10.0	10.0	90.0
	Viudo	2	10.0	10.0	100.0
	Total	20	100.0	100.0	

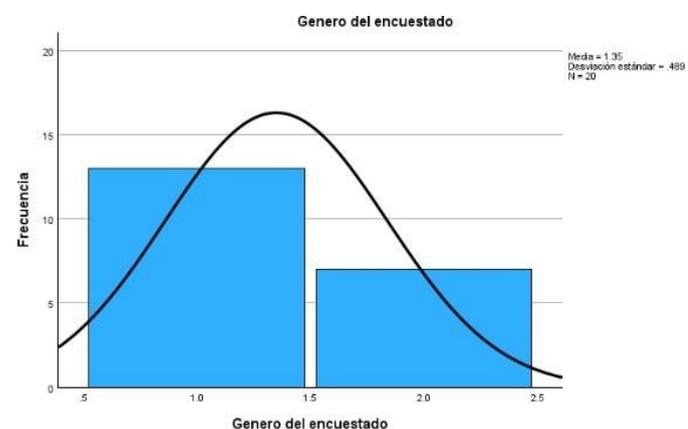


**ITEMS: 1) soltero 2) casado 3) unión libre 4) viudo**

Esta tabla nos muestra la distribución del estado civil de los encuestados en nuestro estudio. El 40% de los encuestados son solteros y otro 40% son casados. Esto indica una distribución bastante equitativa entre estos dos grupos. Hay una menor presencia de uniones libres y viudos ambos representan un 10% cada uno podemos decir que estos estados civiles son menos comunes en nuestra muestra.

#### 4. Estado civil (Tabla de frecuencia y histograma).

Genero del encuestado					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Mujer	13	65.0	65.0	65.0
	Hombre	7	35.0	35.0	100.0
	Total	20	100.0	100.0	

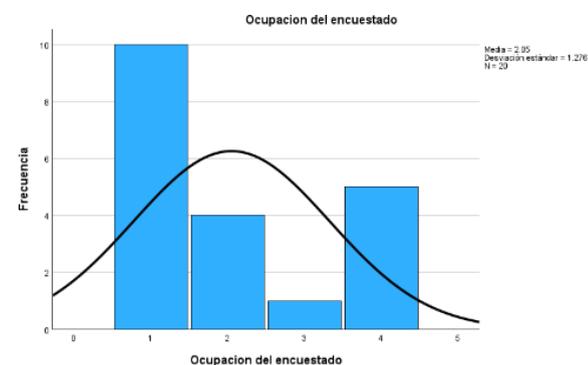


**ITEMS: 1) mujer 2) hombre**

Esta tabla vemos la distribución de género de los encuestados en nuestro estudio. Vemos que del total de los encuestados (son 20), 13 son mujeres, lo que representa un 65% de la muestra, mientras que 7 son hombres, siendo el 35% restante. Esto nos indica una clara predominancia de mujeres en nuestra muestra, casi el doble que la de hombres. Sabemos que las opiniones y experiencias pueden variar según el género y eso podría influir en los resultados generales de la encuesta.

## 5. Ocupación del encuestado (Tabla de frecuencia y histograma).

Ocupacion del encuestado					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Estudiante	10	50.0	50.0	50.0
	Comerciante	4	20.0	20.0	70.0
	Desempleado	1	5.0	5.0	75.0
	Empleado	5	25.0	25.0	100.0
	Total	20	100.0	100.0	



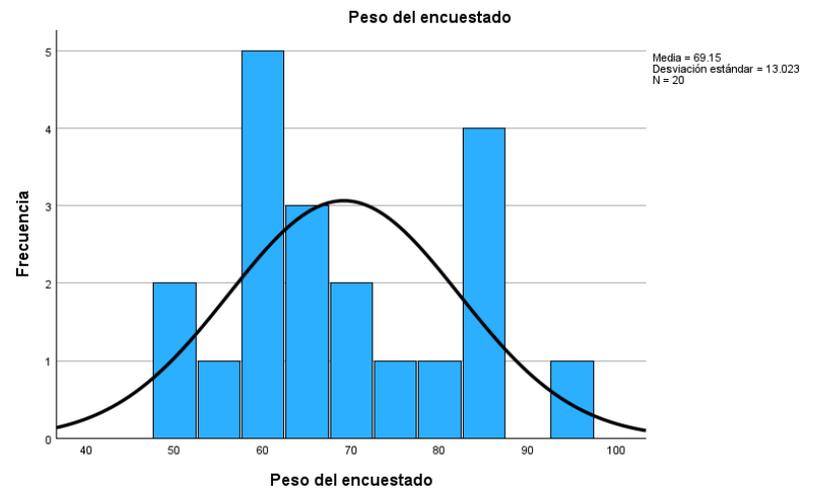
**ITEMS: 1) estudiante 2) comerciante 3) desempleado 4) empleado**

Esta tabla nos muestra la distribución ocupacional de los participantes en nuestra encuesta. La mayoría de los encuestados, un 50%, son estudiantes, lo cual sugiere que nuestra muestra está fuertemente inclinada hacia este grupo demográfico. Le siguen los empleados, representando el 25% de los encuestados, y los comerciantes con un 20%. Un pequeño porcentaje, el 5%, se identifica como desempleado. La distribución ocupacional es crucial para entender el contexto de las respuestas obtenidas, ya que las perspectivas y experiencias varían significativamente entre estudiantes, empleados, comerciantes y personas desempleadas. respuestas difieren entre los distintos grupos ocupacionales para obtener una visión más completa y matizada de los datos.

## 6. PESO (Tabla de frecuencia y histograma).

**Peso del encuestado**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	50	1	5.0	5.0	5.0
	52	1	5.0	5.0	10.0
	56	1	5.0	5.0	15.0
	58	1	5.0	5.0	20.0
	59	1	5.0	5.0	25.0
	60	1	5.0	5.0	30.0
	62	2	10.0	10.0	40.0
	63	1	5.0	5.0	45.0
	64	1	5.0	5.0	50.0
	65	1	5.0	5.0	55.0
	68	1	5.0	5.0	60.0
	72	1	5.0	5.0	65.0
	76	1	5.0	5.0	70.0
	79	1	5.0	5.0	75.0
	85	3	15.0	15.0	90.0
	87	1	5.0	5.0	95.0
	95	1	5.0	5.0	100.0
	Total	20	100.0	100.0	



## ITEM: NINGUNO

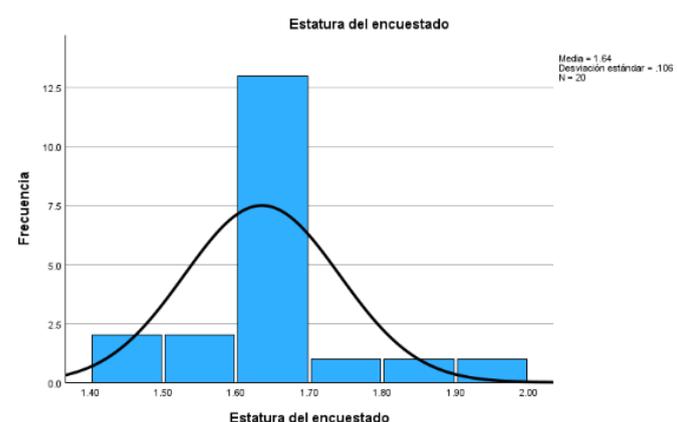
Esta tabla muestra la distribución del peso de los encuestados en el estudio. Se presentan cuatro columnas principales: Frecuencia, Porcentaje, Porcentaje válido y Porcentaje acumulado.

La frecuencia nos indica cuántas personas tienen un peso específico. mayormente cada peso aparece una sola vez (frecuencia = 1), excepto el peso de 62 kg, que aparece dos veces. El Porcentaje y Porcentaje válido reflejan la misma información, ya que no hay datos perdidos.

Cada peso representa el 5% del total, excepto 62 kg, que representa el 10% y Porcentaje acumulado muestra la suma progresiva de los porcentajes esto quiere decir que el peso de 50 kg tiene un porcentaje acumulado del 5%, el de 52 kg un 10% hasta llegar al 100%. Vemos peso de los encuestados está bien distribuido en un rango que va desde los 50 kg hasta los 95 kg.

## 7. Estatura (Tabla de frecuencia y histograma)

Estatura del encuestado				
	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	1.48	2	10.0	10.0
	1.50	1	5.0	15.0
	1.55	1	5.0	20.0
	1.60	4	20.0	40.0
	1.62	2	10.0	50.0
	1.64	2	10.0	60.0
	1.65	2	10.0	70.0
	1.66	1	5.0	75.0
	1.68	2	10.0	85.0
	1.70	1	5.0	90.0
	1.80	1	5.0	95.0
	1.95	1	5.0	100.0
Total	20	100.0	100.0	



ITEM: NINGUNO

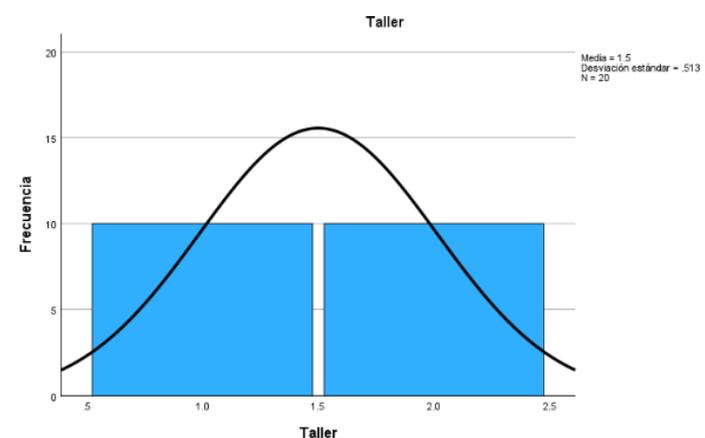
Esta tabla muestra la distribución de la estatura de los encuestados. Se evaluaron 20 personas, y se presentan cuatro columnas principales: Frecuencia, Porcentaje, Porcentaje válido y Porcentaje acumulado.

La estatura más común es 1.60 m, con 4 personas (20% de la muestra). La estatura 1.48 m tiene 2 personas (10%). Las demás alturas tienen una frecuencia de 1 persona (5% cada una).

El porcentaje acumulado muestra que el 50% de los encuestados miden 1.62 m o menos, y el 100% está por debajo de 1.95 m. La distribución de estaturas es variada, con una ligera concentración en torno a 1.60 m.

## 8. TALLER (Tabla de frecuencia y histograma).

		Taller			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	con libro	10	50.0	50.0	50.0
	sin libro	10	50.0	50.0	100.0
	Total	20	100.0	100.0	



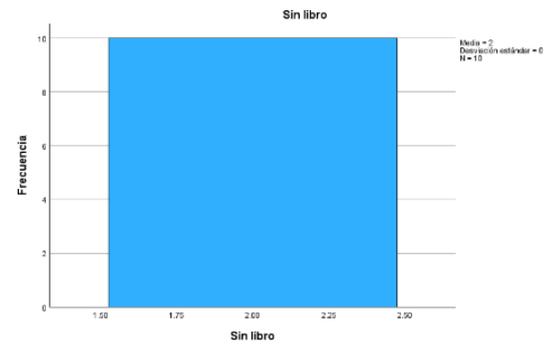
**ITEM: 1) con libro    2) sin libro**

Esta tabla muestra la distribución de los encuestados en dos grupos según la posesión de un libro dentro del taller. 10 personas (50%) tienen libro. 10 personas (50%) no tienen libro.

El porcentaje acumulado muestra que, al sumar ambos grupos, se alcanza el 100% de los encuestados. Esto indica una división equitativa en la muestra entre quienes tienen y quienes no tienen libro.

## 9. Sin libro (Tabla de frecuencia y histograma)

		Sin libro			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	2	10	50.0	100.0	100.0
Perdidos	Sistema	10	50.0		
Total		20	100.0		

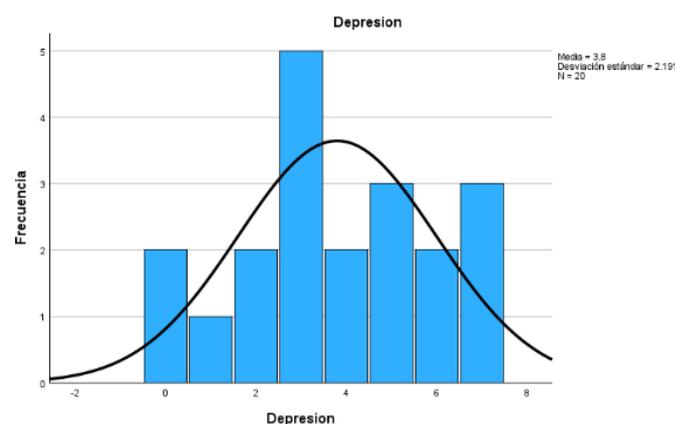


## ITEM: 1) Valido 2) Perdidos

Esta tabla detalla la distribución de los encuestados que no tienen libro y los datos perdidos. 10 personas (50%) se registraron en la categoría "2". 10 personas (50%) aparecen como datos perdidos en el sistema. El porcentaje acumulado llega al 100%. Esto sugiere que, dentro del grupo sin libro, hay una categoría específica ("2") y la mitad de los datos se considera perdida, lo que podría afectar el análisis.

### 10. Depresión (Tabla de frecuencia y histograma)

Depresion					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	0	2	10.0	10.0	10.0
	1	1	5.0	5.0	15.0
	2	2	10.0	10.0	25.0
	3	5	25.0	25.0	50.0
	4	2	10.0	10.0	60.0
	5	3	15.0	15.0	75.0
	6	2	10.0	10.0	85.0
	7	3	15.0	15.0	100.0
	Total	20	100.0	100.0	



## ITEM: 1) 2) 3) 4) 5) 6) 7)

El 25% de los participantes obtuvo un puntaje de 3, siendo la categoría más frecuente, mientras que un 15% alcanzó los niveles 5, 6 y 7. Además, el 10% de los encuestados no presentó síntomas de depresión, obteniendo un puntaje de 0. El porcentaje acumulado revela que el 50% de los encuestados tiene un puntaje de 3 o menor, mientras que el otro 50% presenta niveles superiores. Esto indica una distribución diversa de síntomas depresivos, con una concentración moderada en

niveles bajos y medios.

## CONCLUSIÓN

En conclusión, este documento cumple la función de evidenciar los conocimientos adquiridos en la primera unidad de Estadística Inferencial, específicamente en torno al Teorema del Límite Central y la Estimación. A través del uso de IBM SPSS Statistics, se llevaron a cabo distintos análisis de datos que permitieron aplicar conceptos teóricos en un entorno práctico. La inclusión de ejercicios, tablas, gráficos e interpretaciones facilita la comprensión de los resultados obtenidos. Además, el uso de SPSS demostró ser una herramienta valiosa para el análisis estadístico, permitiendo gestionar grandes volúmenes de datos y aplicar distintos modelos sin necesidad de un conocimiento avanzado en programación.

IBM SPSS Statistics es una herramienta fundamental que me puede ayudar en el análisis de datos de manera eficiente y precisa. Su capacidad para gestionar grandes volúmenes de información, realizar análisis descriptivos, generar gráficos y aplicar modelos estadísticos avanzados facilita la interpretación de resultados y la toma de decisiones basadas en datos.

Además, su interfaz gráfica intuitiva me permite realizar análisis sin necesidad de conocimientos avanzados en programación, lo que optimiza el tiempo y mejora la comprensión de los conceptos estadísticos. Gracias a SPSS, puedo aplicar métodos como regresión, ANOVA y correlaciones de manera sencilla, lo que resulta especialmente útil en investigaciones, estudios académicos y proyectos que requieran un análisis cuantitativo riguroso.

UNIVERSIDAD JUAREZ AUTONOMA DE TABASCO

División Académica de Ciencias Económica Administrativas

**“Examen Primer Parcial”**

Lic. En Contaduría pública

Alumno:

Luis Enrique Muñoz Paulino

Matricula:

241B38018

Semestre

: 3

Asignatura:

Estadística Diferencial

Impartida por:

Prof. Cesar Enrique Barrera Ovando

Ciclo largo febrero 2025-agosto 2025

Villahermosa, tabasco, México

# Introducción

El presente documento muestra la interpretación de algunas variables en específico de una base de datos, como parte del primer examen parcial de la materia de estadística diferencial.

En el cual verificamos las variables de forma correcta según si es cuantitativa o cualitativa, empleando lo aprendido con el programa SSPS, que nos ayudara a realizar descripciones de variables, gráficos, tablas, etc.

# Interpretación de las variables

## Estadísticos descriptivos

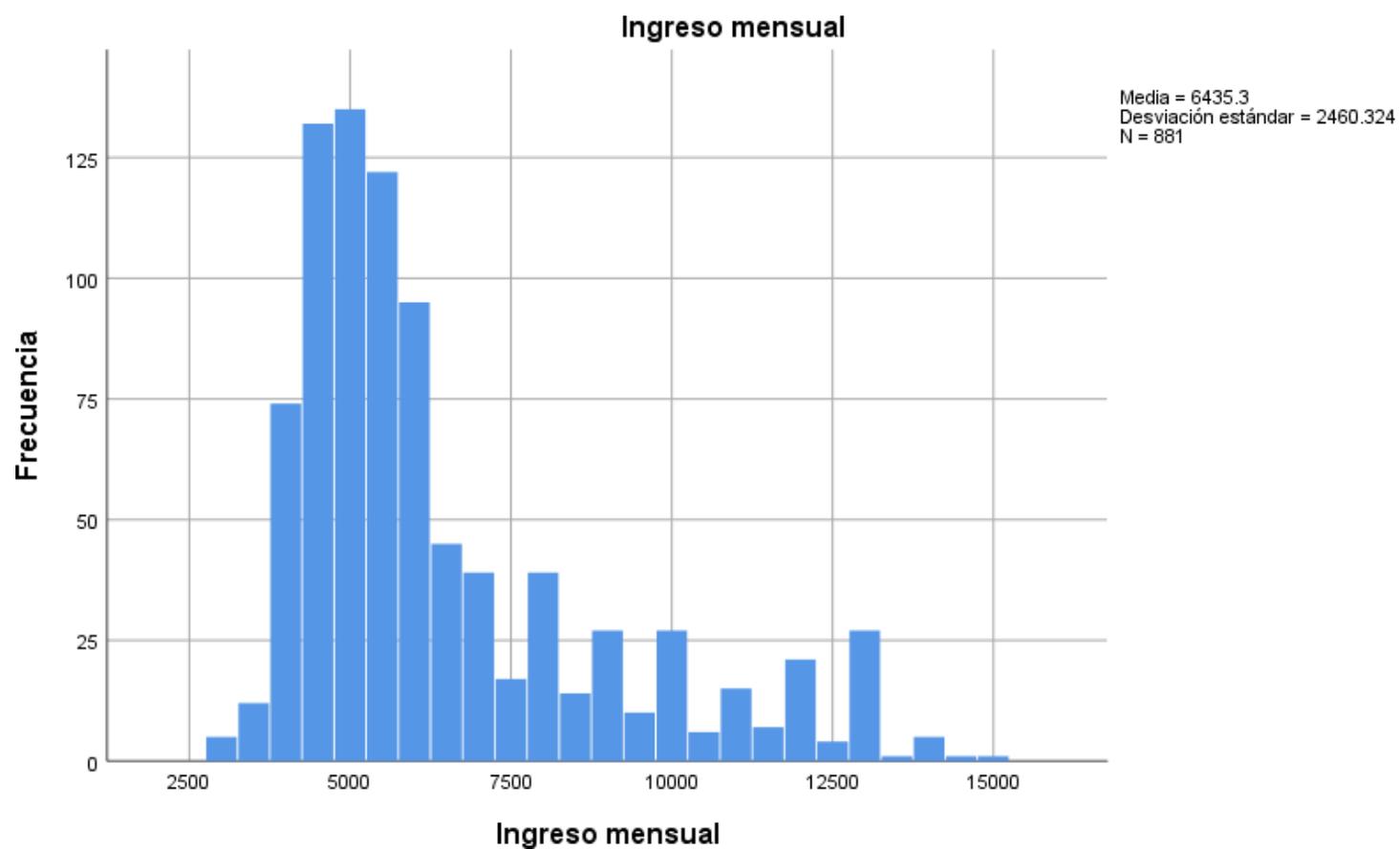
	N	Mínimo	Máximo	Media	Desv. Desviación
Ingreso mensual	881	3000	15000	6435.30	2460.324
Talla	881	1.50	2.10	1.7912	.11112
Grado de instruccion academico	881	1	4	3.20	.852
N válido (por lista)	881				

La tabla muestra un análisis estadístico(encuesta) que se tomó a 881 personas. En la cual el análisis muestra descriptivo de 3 variables: ingreso mensual, talla, grado de instrucción académico proporcionando en un resumen las medidas de tendencias centrales y la variabilidad de datos pues cada persona que participo presenta situaciones diferentes, haciendo que haya una variabilidad de datos.

## Ingreso mensual

**Ingreso mensual**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	3000	5	.5	.6	.6
	3500	12	1.2	1.4	1.9
	4000	74	7.4	8.4	10.3
	4500	132	13.2	15.0	25.3
	5000	135	13.5	15.3	40.6
	5500	122	12.2	13.8	54.5
	6000	95	9.5	10.8	65.3
	6500	45	4.5	5.1	70.4
	7000	39	3.9	4.4	74.8
	7500	17	1.7	1.9	76.7
	8000	39	3.9	4.4	81.2
	8500	14	1.4	1.6	82.7
	9000	27	2.7	3.1	85.8
	9500	10	1.0	1.1	86.9
	10000	27	2.7	3.1	90.0
	10500	6	.6	.7	90.7
	11000	15	1.5	1.7	92.4
	11500	7	.7	.8	93.2
	12000	21	2.1	2.4	95.6
	12500	4	.4	.5	96.0
13000	27	2.7	3.1	99.1	
13500	1	.1	.1	99.2	
14000	5	.5	.6	99.8	
14500	1	.1	.1	99.9	
15000	1	.1	.1	100.0	
	Total	881	88.2	100.0	
Perdidos	Sistema	118	11.8		
	Total	999	100.0		

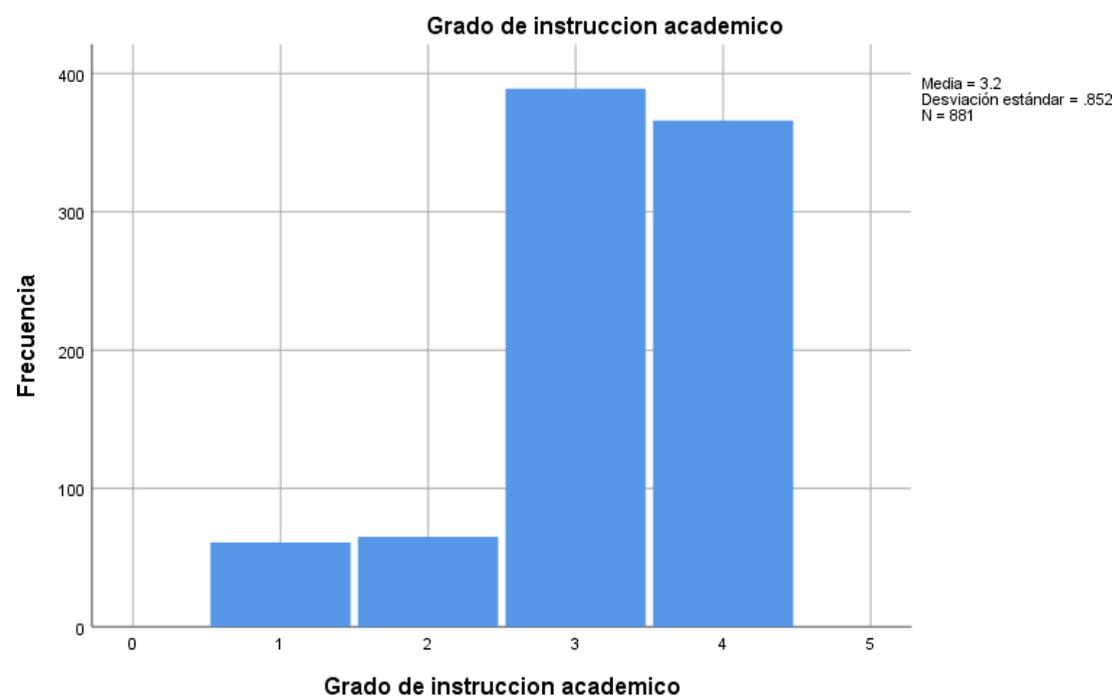


la tabla de frecuencia muestra la distribución del ingreso mensual de 881 personas. Los valores de ingreso varían desde \$3000 hasta \$15000. Se puede ver que los ingresos mas frecuentes son de \$4500 siendo un total de 132, \$5000 con un total de 135. Lo que indica que la mayoría de los encuestados esta en este rango y a medida que los ingresos aumentan la frecuencias de personas disminuye dando a entender que menos gente ganan esta cantidad.

# Grado de instrucción académica

Grado de instruccion academico

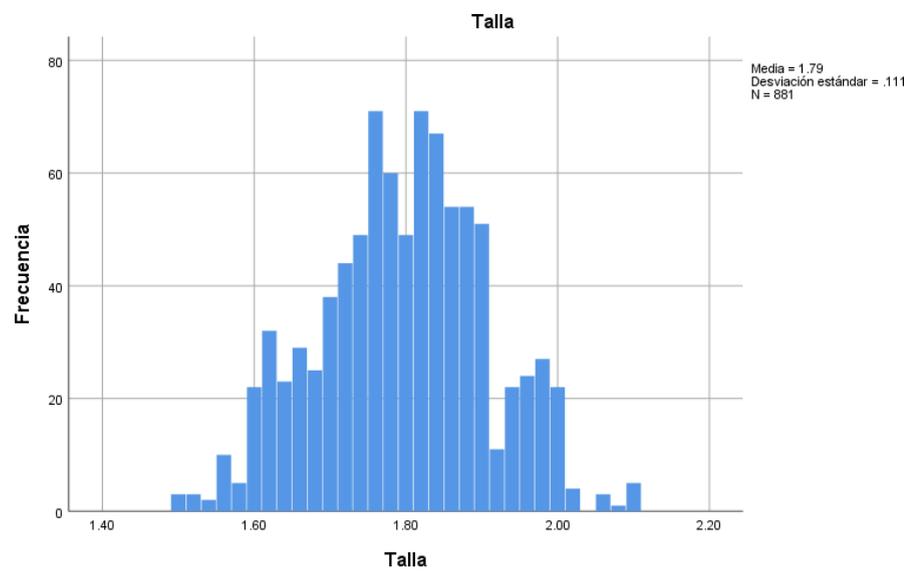
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	1	61	6.1	6.9	6.9
	2	65	6.5	7.4	14.3
	3	389	38.9	44.2	58.5
	4	366	36.6	41.5	100.0
	Total	881	88.2	100.0	
Perdidos	Sistema	118	11.8		
Total		999	100.0		



Se observa que la mayor parte de la muestra tiene un nivel de instrucción, de variables que van del 1 al 4. Nos dice que el nivel 1 es igual al 6.1% , el nivel 2 cuenta con 6.5% con 65 personas, el nivel 3 cuenta con 38.9% con 389 personas y el nivel 4 con 36.6% con 366 personas, indicando que la mayoría de los encuestado están capacitado con el nivel máximo.

# Grado de instrucción

		Talla			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	1.50	3	.3	.3	.3
	1.51	2	.2	.2	.6
	1.52	1	.1	.1	.7
	1.53	1	.1	.1	.8
	1.54	1	.1	.1	.9
	1.55	5	.5	.6	1.5
	1.56	5	.5	.6	2.0
	1.57	3	.3	.3	2.4
	1.58	2	.2	.2	2.6
	1.59	5	.5	.6	3.2
	1.60	17	1.7	1.9	5.1
	1.61	12	1.2	1.4	6.5
	1.62	20	2.0	2.3	8.7
	1.63	8	.8	.9	9.6
	1.64	15	1.5	1.7	11.4
	1.65	16	1.6	1.8	13.2
	1.66	13	1.3	1.5	14.6
	1.67	12	1.2	1.4	16.0
	1.68	13	1.3	1.5	17.5
	1.69	13	1.3	1.5	19.0
	1.70	25	2.5	2.8	21.8
	1.71	22	2.2	2.5	24.3
	1.72	22	2.2	2.5	26.8
	1.73	30	3.0	3.4	30.2
	1.74	19	1.9	2.2	32.3
	1.75	37	3.7	4.2	36.5
	1.76	34	3.4	3.9	40.4
	1.77	24	2.4	2.7	43.1
	1.78	36	3.6	4.1	47.2
	1.79	30	3.0	3.4	50.6
	1.80	19	1.9	2.2	52.8
	1.81	43	4.3	4.9	57.7
	1.82	28	2.8	3.2	60.8
	1.83	38	3.8	4.3	65.2
	1.84	29	2.9	3.3	68.4
	1.85	38	3.8	4.3	72.8
	1.86	16	1.6	1.8	74.6
	1.87	25	2.5	2.8	77.4
	1.88	29	2.9	3.3	80.7
	1.89	23	2.3	2.6	83.3
	1.90	28	2.8	3.2	86.5
	1.91	4	.4	.5	86.9
	1.92	7	.7	.8	87.7
	1.93	15	1.5	1.7	89.4
	1.94	7	.7	.8	90.2
	1.95	12	1.2	1.4	91.6
	1.96	12	1.2	1.4	93.0
	1.97	9	.9	1.0	94.0
	1.98	18	1.8	2.0	96.0
	1.99	13	1.3	1.5	97.5
	2.00	9	.9	1.0	98.5
	2.01	3	.3	.3	98.9
	2.02	1	.1	.1	99.0
	2.05	2	.2	.2	99.2
	2.06	1	.1	.1	99.3
	2.07	1	.1	.1	99.4
	2.09	3	.3	.3	99.8
	2.10	2	.2	.2	100.0
	Total	881	88.2	100.0	
Perdidos	Sistema	118	11.8		
	Total	999	100.0		



En este tabla se muestra la talla que esta entre los 1.50 m y los 2.10 m, la cual la media de la población es de 1.70 siendo la medida central

## PREGUNTAS

1. En la variable Grado de instrucción académica, cuantas personas cumplen con las condiciones de cada grupo (ejemplo cuantos son de licenciatura, cuantos de ) y su porcentaje.

nivel 1 es igual al 6.1% con un total de 61 personas  
 nivel 2 cuenta con 6.5% con 65 personas,  
 nivel 3 cuenta con 38.9% con 389 personas  
 nivel 4 con 36.6% con 366 personas,

**2. En la variable ingreso mensual cuantas personas tienen ingreso entre \$7,000 y \$12,000 y su porcentaje**

R: 166 personas tiene el ingreso entre \$7000 y \$12000 y el porcentaje es 18.8422%

**3. En la variable Talla, cuantas personas miden entre 1.69 y 1.80, y su porcentaje**

R: 311 personas 35.30%

## CONCLUSIÓN SOBRE EL PARCIAL

El Teorema del Límite Central me ayudó a entender cómo, al aumentar el tamaño de una muestra, la distribución de la media muestral tiende a ser normal, sin importar la forma de la población original. Esto es clave para justificar muchas de las técnicas inferenciales que usamos en estadística.

Además, al trabajar con estimaciones puntuales e intervalos de confianza, pude ver cómo es posible hacer inferencias sobre parámetros poblacionales a partir de datos muestrales, lo que considero una habilidad muy valiosa aplicación práctica de estos conceptos utilizando el software SPSS. Al desarrollar ejercicios y prácticas con esta herramienta, pude llevar la teoría a un contexto más real y aplicable, reforzando mis habilidades para el análisis de datos con tecnología especializada.

En general, esta unidad me permitió integrar conocimientos teóricos y prácticos, desarrollar mi pensamiento crítico y mejorar mis competencias en el uso de herramientas estadísticas. Me siento más preparado para enfrentar los siguientes temas del curso y aplicar estos aprendizajes en contextos académicos y profesionales.



# **UNIDAD 2**

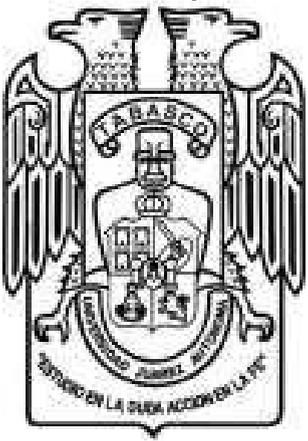
# **PRUEBA DE HIPOTESIS Y**

# **CONTROL DE CALIDAD**

## **CONTENIDO:**

**ENSAYO**  
**MAPA MENTAL**  
**DOCUMENTO EXTENSO**  
**PRACTICAS DEL LABORATORIO DE COMPUTO**  
**EXAMEN DE LA UNIDAD**

**CONCLUSIÓN FINAL DE LA UNIDAD**



**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



**“ENSAYO”**

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 2025**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### INTRODUCCIÓN

El siguiente ensayo es un análisis de nuestro aprendizaje durante la unidad uno, específicamente sobre el tema de **\*\*pruebas de hipótesis\*\*** de la materia de Estadística Inferencial. En este escrito redactaré los aprendizajes que obtuvimos, así como las dificultades que enfrentamos al aprender a utilizar el software con el que desarrollamos la primera práctica.

El estudio de la estadística nos permite comprender el comportamiento de los datos y hacer inferencias sobre poblaciones a partir de muestras. En este ensayo, analizaré un concepto clave en este proceso: **\*\*las pruebas de hipótesis\*\***. Estas constituyen una herramienta fundamental para la toma de decisiones informadas, y que permiten evaluar afirmaciones o supuestos sobre parámetros poblacionales a partir de información muestral.

Las pruebas de hipótesis nos ayudan a determinar si los resultados obtenidos en una muestra son suficientemente significativos como para generalizarlos a toda la población. Durante esta unidad, aprendimos a formular hipótesis nula y alternativa, elegir un nivel de significancia adecuado y aplicar distintos tipos de pruebas según el tipo de datos y el objetivo del análisis.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



A lo largo de este ensayo, explicaré cómo estas pruebas facilitan la validación de supuestos y la detección de diferencias significativas en contextos reales. También reflexionaré sobre la importancia de conocer los errores tipo I y tipo II, y cómo estos influyen en la interpretación correcta de los resultados. Por último, compartiré cómo fue nuestro proceso de aprendizaje en el uso del software estadístico, el cual nos presentó algunos retos iniciales, pero que finalmente nos permitió aplicar los conceptos aprendidos de manera práctica y efectiva..

En mi aprendizaje dentro de la estadística inferencial, uno de los temas que más me llamó la atención fueron las pruebas de hipótesis, ya que representan una herramienta muy poderosa para tomar decisiones con base en datos. Una prueba de hipótesis es un método que nos permite aceptar o rechazar una afirmación sobre un parámetro poblacional, basándonos en la información que obtenemos de una muestra.

Lo que más me quedó claro es que en estas pruebas no trabajamos con toda la población, sino que usamos una muestra representativa y, a partir de ahí, decidimos si la evidencia apoya o no una cierta hipótesis. Por ejemplo, si se sospecha que un valor medio ha cambiado o que un proceso ya no cumple con ciertos estándares, las pruebas de hipótesis nos ayudan a confirmar o rechazar ese tipo de ideas.

Durante el estudio entendí que todo comienza con formular dos hipótesis: la hipótesis nula ( $H_0$ ), que es la afirmación inicial que queremos poner a prueba, y la

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

hipótesis alternativa ( $H_1$ ), que representa una conclusión diferente si la evidencia no



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



apoya a  $H_0$ . Estas hipótesis se contrastan con un nivel de significancia (generalmente 0.05), que nos ayuda a definir qué tan tolerantes somos al error al momento de tomar la decisión.

En la práctica, por ejemplo, si una empresa fabricante de baterías afirma que su producto dura en promedio 48 meses, esa afirmación sería la hipótesis nula ( $H_0: \mu \geq 48$ ).

Si los clientes comienzan a quejarse de que las baterías duran menos, se puede tomar una muestra y realizar una prueba de hipótesis para ver si el promedio real es menor, lo que plantearía una hipótesis alternativa ( $H_1: \mu < 48$ ).

Esta misma lógica se puede aplicar directamente en el ámbito de la contabilidad, donde constantemente trabajamos con suposiciones y estimaciones. Por ejemplo, si en una empresa se estima que los gastos administrativos representan el 15% de los ingresos, podríamos plantear una hipótesis nula de que  $\mu = 15\%$  y, con base en los datos de varios periodos o sucursales, hacer una prueba de hipótesis para ver si realmente ese porcentaje se mantiene o no. De esa manera, se puede evaluar si hay un cambio significativo que requiera ajustar presupuestos o políticas contables.

Otro caso práctico sería analizar si el tiempo promedio de cobro de cuentas por cobrar se mantiene dentro del límite establecido por la empresa. Si se sospecha que ha aumentado, se puede hacer una prueba de hipótesis con los tiempos actuales y tomar decisiones más acertadas para optimizar la gestión financiera.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Las pruebas de hipótesis me han mostrado cómo es posible tomar decisiones objetivas en base a datos. Son muy útiles no solo en contextos académicos, sino también en el entorno profesional, especialmente en contabilidad, donde el análisis riguroso y la toma de decisiones fundamentadas son clave para el buen manejo financiero de cualquier organización.

Una de las primeras cosas que comprendí es que en estadística no se trata solamente de calcular promedios o porcentajes, sino de usar esa información para llegar a conclusiones. Las pruebas de hipótesis son un ejemplo claro de esto. Este método nos ayuda a determinar si una afirmación sobre una población es verdadera o falsa, basándonos en los datos obtenidos de una muestra representativa.

El procedimiento comienza con la formulación de dos hipótesis. La hipótesis nula ( $H_0$ ) es la afirmación que se pone a prueba, normalmente representa el “estado actual” o una suposición establecida. Por otro lado, la hipótesis alternativa ( $H_1$ ) es lo que creemos que podría ser cierto si  $H_0$  no lo es. Por ejemplo, si una empresa dice que su nuevo software de contabilidad reduce el tiempo de registro en un 30%, esa afirmación se puede someter a prueba para ver si realmente hay una mejora significativa.

Otro concepto muy importante que aprender es el del nivel de significancia ( $\alpha$ ), que normalmente se fija en 0.05. Esto quiere decir que estamos dispuestos a aceptar un 5% de probabilidad de cometer un error tipo I, que consiste en rechazar una



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



hipótesis verdadera. También está el error tipo II , que no es rechazar una hipótesis falsa. Conocer estos errores me ha ayudado a entender mejor el riesgo que conlleva tomar decisiones en base a datos.

En la prueba, calculamos una estadística de prueba. a partir de los datos de la muestra , como puede ser a partir de los datos de la muestra, como puede ser la media o la proporción. Luego, se compara ese valor con un valor crítico o se calcula el p-valor . Si el resultado cae en la región crítica o si el p-valor es menor que el nivel de significancia, entonces se rechaza la hipótesis nula.

Este proceso me parece muy útil para analizar, por ejemplo, si los gastos de una empresa están aumentando de manera significativa, si el tiempo de entrega de los productos está dentro del estándar o si los ingresos están cambiando por alguna estrategia implementada. Las pruebas de hipótesis permiten hacer este tipo de evaluaciones con mayor objetividad.

Una de las partes más interesantes para mí es cómo se puede aplicar todo esto a la contabilidad. Por ejemplo, si se estima que el margen de ganancia promedio en un producto es del 20%, y de repente se observan cambios, podemos hacer una prueba de hipótesis para ver si esa variación es estadísticamente significativa o si se trata de una fluctuación normal. Esto puede ser muy útil para evaluar decisiones como cambiar proveedores, ajustar precios o rediseñar procesos.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Otro ejemplo práctico en contabilidad es el análisis de los días promedio de cobranza . Si la empresa considera que, en promedio, se cobra a los 30 días, y últimamente los cobros se están tardando más, se puede tomar una muestra de facturas recientes y realizar una prueba de hipótesis para saber si realmente se ha incrementado ese tiempo. Si el resultado es significativo, es un indicio para revisar los procedimientos de crédito o la selección de clientes.

Además, en auditorías o en el análisis de errores contables, también es muy útil. Si se detectan errores en ciertas cuentas o partidas, se puede determinar si la proporción de errores es superior a la tolerada, utilizando herramientas estadísticas que apoyan una revisión más precisa.

El otro tema fundamental que hemos estudiado es el control de calidad , y este se relaciona directamente con las pruebas de hipótesis. El control de calidad se refiere a los procesos que aseguran que un producto o servicio cumpla con los estándares establecidos. Aquí es donde entra la estadística como herramienta de verificación y mejora continua.

Uno de los conceptos clave en control de calidad son las gráficas de control , que nos muestran cómo varía un proceso a lo largo del tiempo. Estas gráficas utilizan líneas de control superior e inferior, y si un punto cae fuera de esos límites, nos está indicando que algo está fuera de control y atención necesita.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Otra herramienta son los planos de muestreo , que nos permiten revisar una parte del total (una muestra) para determinar si todo el lote cumple con las especificaciones. Esto se relaciona mucho con la idea de hacer inferencias a partir de una muestra, como sucede en las pruebas de hipótesis.

En lo personal, me llamó atención la cómo el control de calidad no solo se aplica en la industria fabricante, sino también en servicios. Por ejemplo, en el área contable o administrativa, se puede aplicar control de calidad al revisar tiempos de entrega de informes financieros, cumplimiento de normas fiscales o exactitud en los registros contables.

Lo que más me ha gustado de aprender sobre las pruebas de hipótesis y el control de calidad es darme cuenta de que no son temas separados, sino que en realidad se complementan bastante . A veces en clase pensamos que cada concepto va por su lado, pero cuando empezamos a hacer ejercicios y ver casos reales, me di cuenta de que muchas veces se necesita aplicar los dos al mismo tiempo para analizar si un proceso realmente está funcionando como debería.

Por ejemplo, si una empresa dice que sus productos no deben tener más del 3% de defectos y se encuentran 10 defectos en una muestra de 200 piezas, ahí no basta con decir "hay muchos errores", sino que se puede usar una prueba de hipótesis para comprobar si ese número de defectos en realidad pasa el límite tolerado . Si el análisis



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



...muestra que sí, entonces se puede tomar una decisión bien fundamentada, como detener la producción, revisar los materiales o ajustar alguna parte del proceso.

Y en el área de contabilidad también me di cuenta de que esta combinación tiene mucho sentido. Imaginemos que el área financiera asegura que el 95% de los informes se entregarán puntualmente. En lugar de quedarnos con la palabra, podríamos aplicar un pequeño control de calidad y una prueba de hipótesis con datos reales para ver si eso se está cumpliendo de verdad o si solo se dice por compromiso. Si encontramos que el porcentaje es menor, eso nos da la señal de que hay que mejorar el proceso o revisar qué está fallando.

Como estudiante, esto me ha hecho ver que la estadística no solo sirve para pasar la materia, sino que realmente es una herramienta muy práctica para analizar, evaluar y tomar decisiones con base en datos, algo que voy a necesitar sí o sí cuando esté en el campo laboral.

De igual manera aprendí a como realizar una prueba de hipótesis basándome en los siguientes pasos:

### 1. Definir las hipótesis

Lo primero que hago cuando me enfrento a una prueba de hipótesis es establecer las dos hipótesis que voy a probar: la hipótesis nula ( $H_0$ ) y la hipótesis alternativa

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

muestra que sí, entonces se puede tomar una decisión bien fundamentada,  
(H1H\_1H1).



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### 2. Elegir el nivel de significancia

El siguiente paso es decidir el nivel de significancia ( $\alpha$ ), que es la probabilidad de cometer un error tipo I, es decir, rechazar la hipótesis nula cuando en realidad es verdadera. Generalmente, selecciona un valor de  $\alpha = 0.05$ , lo que significa que acepta una probabilidad del 5% de cometer un error.

### 3. Seleccionar la Prueba Adecuada

Dependiendo de los datos con los que cuento, elijo la prueba adecuada. Si conozco la desviación estándar de la población y tengo una muestra suficientemente grande, utilizaré la prueba z para los medios. Si no conozco la desviación estándar o si estoy trabajando con una muestra pequeña, utilizaré la prueba t de Student.

### 4. Calcular el Estadístico de Prueba

Con los datos que tengo, calcula el estadístico de prueba. Este estadístico me dice cuán lejos está la media muestral de la media que propuso la hipótesis nula, en términos de la variabilidad de la muestra.

Si, por ejemplo, tengo una muestra de estudiantes con un tiempo promedio de

### 5. Determinar la Región de Rechazo

# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA

### **2. Elegir el nivel de significancia**

estudio de 8 horas, y mi hipótesis nula era que la media es 10, calcularé el valor de  $z$  (ot)

para ver si mi muestra difiere significativamente de lo que se esperaba bajo la hipótesis

nula.

### **5. Determinar la Región de Rechazo**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Una vez que tengo el estadístico de prueba, lo comparo con el valor crítico de la distribución correspondiente ( $z_{\alpha}$ ). El valor crítico depende del nivel de significancia  $\alpha$  que elegí al principio. Si la estadística calculada está en la región de rechazo, rechazo la hipótesis nula.

Para una prueba de dos colas (cuando  $H_1: \mu \neq \mu_0$ ), tengo dos valores críticos: uno para el extremo superior y otro para el extremo inferior de la distribución. Si el valor calculado del estadístico  $z_{ot}$  es más grande en valor absoluto que los valores críticos, rechazo  $H_0$ .

### 6. Tomar la decisión

Finalmente, con base en el valor del estadístico de prueba y el nivel de significancia, tomo una decisión:

- Si rechaza la hipótesis nula, esto significa que los datos proporcionan suficiente evidencia para aceptar que la hipótesis alternativa es verdadera (por ejemplo, que la media no es 10).
- Si no rechazo la hipótesis nula, significa que no tengo suficiente

### 7. Interpretación de resultados

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

Una vez que tengo el estadístico de prueba, lo comparo con el valor crítico de la evidencia para afirmar que la media es diferente de lo que se propuso

inicialmente.

### **7. Interpretación de resultados**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Una vez que tomó una decisión, lo último que hago es interpretar los resultados de la prueba de manera clara y concisa. Si rechazo  $H_0$ , debo concluir que los datos son lo suficientemente fuertes como para indicar que la media poblacional es diferente del valor que planteé en la hipótesis nula. Si no rechazo  $H_0$ , puedo decir que no tengo evidencia suficiente para afirmar que la media es diferente

- Hipótesis nula ( $H_0$ ): Es la suposición inicial, que generalmente establece que no hay efecto o que no hay diferencia. Por ejemplo, si quiero probar si el promedio de horas de estudio de mis compañeros es 10, entonces mi hipótesis nula sería que la media es 10.
- Hipótesis alternativa ( $H_1$ ): Es lo contrario de la hipótesis nula. En este caso, si quiero verificar si la media no es 10, entonces la hipótesis alternativa sería que la media es diferente de 10.

Uno de los temas que más me llamó la atención mientras estudiaba pruebas de hipótesis fue la posibilidad de cometer errores al tomar decisiones, incluso al usar datos y métodos estadísticos. Me di cuenta de que las decisiones en estadística no son siempre 100% seguras, ya que dependen de muestras y no de toda la población. Por eso, hablamos de dos tipos de errores que se pueden cometer: el error tipo I y el error tipo II.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Cuando hacemos pruebas de hipótesis, comenzamos con una hipótesis nula ( $H_0$ ) y una alternativa ( $H_1$ ). Luego, basándonos en una muestra de datos, tomamos la decisión de rechazar  $H_0$  o no rechazarla. Aquí está el detalle: siempre existe una probabilidad de fallar, porque los datos provienen de una muestra, lo que trae incertidumbre.

Así que entendí que hay dos formas en que podemos equivocarnos:

1. Rechazar una hipótesis que en realidad es verdadera.
2. No rechazar una hipótesis que en realidad es falsa.

### Error Tipo I ( $\alpha$ )

El error tipo I se produce cuando rechazamos la hipótesis nula ( $H_0$ ), aunque sea verdadera. En otras palabras, estamos diciendo que hay evidencia para rechazar una afirmación que, de hecho, es cierta.

Este error está relacionado con el nivel de significancia ( $\alpha$ ) que fijamos antes de realizar la prueba. Por ejemplo, si optamos por un nivel de significancia del 5% ( $\alpha = 0.05$ ), estamos aceptando un 5% de probabilidad de cometer un error tipo I.

### Ejemplo práctico:

Imaginemos que una empresa afirma que sus focos tienen una duración promedio de 1000 horas. Si al hacer una prueba de hipótesis decidimos rechazar esa



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



afirmación, pero en realidad los focos sí duran 1000 horas, habríamos cometido un error tipo I.

¿Por qué es importante?

En muchos campos, como la contabilidad, este tipo de error puede llevarnos a desestimar un proceso o un resultado que realmente es correcto, lo que podría resultar en decisiones erróneas, como cambiar una estrategia que funcionaba o realizar auditorías innecesarias.

Error Tipo II ( $\beta$ )

El error tipo II se presenta cuando no rechazamos la hipótesis nula, incluso cuando es falsa. En otras palabras, aceptamos algo como válido cuando en realidad no lo es.

Este error está relacionado con el poder de la prueba, que se refiere a la capacidad de detectar un efecto real cuando existe. Cuanto mayor sea el tamaño de la muestra o más pronunciado sea el efecto, menor será la probabilidad de que cometamos un error tipo II.

Ejemplo práctico:

duran menos de 1000 horas, pero al hacer la prueba no logramos rechazar  $H_0$ , estaríamos cometiendo un error tipo II. Es decir, creemos que todo está en orden cuando no lo está.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



¿Y en contabilidad?

Pongamos que en un control financiero, se sospecha que hay retrasos en la entrega de informes. Si realizamos una prueba y no encontramos evidencia suficiente para afirmar que el 95% se entregan a tiempo, pero en realidad solo el 85% se entrega a tiempo, estamos cometiendo un error tipo II. Esto puede implicar no detectar un problema real a tiempo.

Como estudiante de contabilidad, es fundamental tener en cuenta estos errores, porque aunque trabajemos con números y fórmulas, las decisiones que tomamos a partir de análisis estadístico pueden tener repercusiones reales en los negocios. Por ejemplo, rechazar un informe sin motivo puede perjudicar la imagen de un área; y no detectar un problema real puede resultar en pérdidas económicas o afectar la toma de decisiones.

He aprendido también que disminuir la probabilidad de un tipo de error normalmente aumenta la del otro. Por eso es esencial encontrar un balance según la situación. En algunos casos, cometer un error tipo I puede ser más grave, y en otros, un tipo II puede ser lo más perjudicial.

Entender los errores tipo I y tipo II me ha permitido ver la estadística con una perspectiva más responsable. Ya no se trata solo de obtener un p-valor y ver si es menor que 0.05. Se trata de reconocer que tomamos decisiones con un cierto nivel de riesgo

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

yque, como futuros profesionales, debemos saber cuándo es aceptable correr



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



ese riesgo y cuándo no. Por todo esto, creo que este tema es esencial, no solo en estadística, sino en todas las decisiones basadas en datos.

La de aprender sobre las pruebas de hipótesis, que son una herramienta crucial en la estadística inferencial. En particular, quiero compartir cómo entiendo la **prueba de hipótesis para los medios**, que me permite tomar decisiones sobre el valor promedio de una población basándome en una muestra de datos. A continuación, explico cómo funciona esta prueba, qué pasos seguir para realizarla y cómo. Como estudiante universitario, tuve la oportunidad de aprender sobre las pruebas de hipótesis, que son una herramienta crucial en la estadística inferencial. En particular, quiero compartir cómo entiendo la **prueba de hipótesis para los medios**, que me permite tomar decisiones sobre el valor promedio de una población basándome en una muestra de datos.

La prueba de hipótesis es un procedimiento estadístico que nos permite tomar decisiones o sacar conclusiones sobre una población, basándonos en una muestra de datos. En el caso de la **prueba de hipótesis para los medios**, nos enfocamos en determinar si los medios de una población es igual a un valor específico que planteamos en la hipótesis nula. En otras palabras, la prueba nos ayuda a saber si el valor de la media que creemos es correcto o si los datos sugieren que debería ser diferente.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Por ejemplo, supongamos que un fabricante afirma que las bombillas de su producción tienen una vida útil promedio de 1,000 horas. Para comprobar si esta afirmación es cierta, realice una prueba de hipótesis en la que compare la media muestral de la vida útil de algunas bombillas con la media poblacional de 1,000 horas.

La de estudiar los conceptos fundamentales de la estadística, y uno de los más importantes y útiles, en mi opinión, es la **proporción poblacional**. En términos simples, una proporción poblacional es la fracción de la población que tiene una característica específica, y entender cómo calcularla y usarla me ha permitido hacer análisis más precisos sobre grupos y tomar decisiones informadas. Como estudiante universitario, tuve la oportunidad de estudiar los conceptos fundamentales de la estadística, y uno de los más importantes y útiles, en mi opinión, es la **proporción poblacional**. En términos simples, una proporción poblacional es la fracción de la población que tiene una característica específica, y entender cómo calcularla y usarla me ha permitido hacer análisis más precisos sobre grupos y tomar decisiones informadas.

### ¿Qué es la proporción poblacional?

En primer lugar, considere que es crucial comprender qué es exactamente la proporción poblacional. Esta representa la parte de un conjunto que posee una característica específica, como el porcentaje de personas que prefieren un determinado producto o el porcentaje de estudiantes que aprueban un curso. En mi experiencia,



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Trabajar con proporciones me ha dado una forma clara de expresar y entender la relación entre una parte de la población y el total.

Las pruebas de hipótesis para proporciones poblacionales son una herramienta poderosa para verificar suposiciones sobre una población a partir de una muestra. Este proceso me ha permitido hacer inferencias con un nivel de certeza sobre características de la población, y es una parte fundamental del análisis estadístico. Al aprender y practicar este procedimiento, puedo tomar decisiones más informadas y fundamentadas, lo que es crucial tanto en el ámbito académico como en la vida profesional.

Una de las pruebas más útiles que he aprendido es la **prueba de hipótesis para la diferencia de dos medios**, que me permite comparar dos grupos diferentes, por ejemplo, dos tratamientos o dos métodos de enseñanza, para ver si sus medios son significativamente diferentes.

### ¿Cuándo usar la prueba?

Cuando tengo dos muestras independientes y quiero saber si las medias de los dos grupos son iguales o diferentes, utilizo esta prueba. Un ejemplo que me ayuda a entender cómo funciona esta prueba es cuando quiero comparar el rendimiento académico de dos grupos de estudiantes que han recibido diferentes métodos de enseñanza.

# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA

trabajar con proporciones me ha dado una forma clara de expresar y entender la  
**Hipótesis para la diferencia de dos medias**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



En este tipo de pruebas, planteo dos hipótesis:

- **Hipótesis nula ( $H_0$  o  $H_0$ )** : Se establece que no hay diferencia significativa entre los medios de los dos grupos. Por ejemplo, si quiero comparar dos métodos de enseñanza, la hipótesis nula sería que la media de los resultados del primer grupo es igual a la media del segundo grupo.
- **Hipótesis alternativa ( $H_1$  o  $H_1$ )** : Aquí se plantea que hay una diferencia significativa entre los medios de los dos grupos.

Por otro lado, otra prueba importante que he aprendido es la prueba de hipótesis para la diferencia de dos proporciones poblacionales . Esta prueba me permite comparar dos proporciones de diferentes grupos o poblaciones para ver si existe una diferencia significativa entre ellas.

¿Cuándo usar la prueba?

Esta prueba es especialmente útil cuando quiero comparar la proporción de éxito (o cualquier otra característica binaria) entre dos grupos. Por ejemplo, podría usarla para comparar la proporción de estudiantes aprobados en dos diferentes cursos.

Hipótesis para la diferencia de dos proporciones

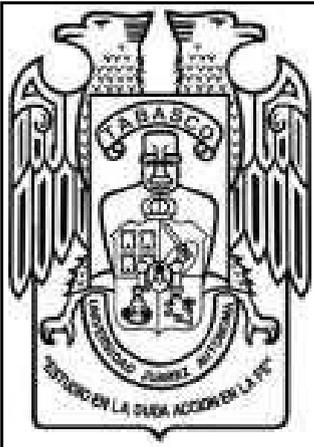
Al igual que con las medias, planteo las hipótesis para esta prueba:

Hipótesis nula ( $H_0$  o  $H_0$ ) : No hay diferencia significativa entre las proporciones

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

En este tipo de pruebas, planteo dos hipótesis:  
de los dos grupos. Es decir, las dos proporciones son iguales.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Hipótesis alternativa ( $H_1$ ) : Hay una diferencia significativa entre las proporciones de los dos grupos

tanto las pruebas de hipótesis para la diferencia de dos medios como para la diferencia de dos proporciones poblacionales son herramientas muy poderosas que utilizan para comparar dos grupos y determinar si existen diferencias significativas.

Como estudiante, me ha resultado muy útil entender cómo llevar a cabo estos procedimientos, ya que me permiten tomar decisiones informadas basadas en datos y evidencia empírica. Además, tanto en el ámbito académico como profesional, este tipo de análisis me ayuda a evaluar si las diferencias observadas son reales o simplemente el resultado de la variabilidad natural de las muestras.

### **La Importancia de las Pruebas de Hipótesis y el Control de Calidad: Mi**

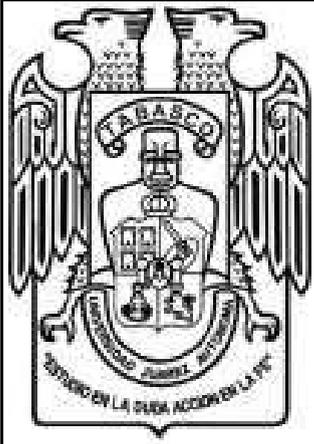
#### **Perspectiva**

Como estudiante universitario, he aprendido que las pruebas de hipótesis y el control de calidad son herramientas esenciales tanto en la investigación como en la toma de decisiones dentro de cualquier organización. Estas herramientas no solo permiten hacer inferencias estadísticas, sino que también son cruciales para garantizar la eficacia y la mejora continua de los procesos, productos y servicios. A lo largo de mis estudios, he comprendido que su importancia va más allá de las aulas y tiene aplicaciones prácticas en una amplia gama de industrias, desde la manufactura hasta la

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

Hipótesis alternativa (H1H\_1H1) : Hay una diferencia significativa entre las investigación científica.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### Las Pruebas de Hipótesis: Una Herramienta Clave para la Toma de Decisiones

Las pruebas de hipótesis, en mi opinión, son fundamentales porque proporcionan un marco estructurado para tomar decisiones basadas en evidencia. Al realizar una prueba de hipótesis, estamos validando o refutando suposiciones sobre una población basándonos en los datos de una muestra. Esto es esencial porque, en la mayoría de los casos, no tenemos acceso a toda la población, y las pruebas de hipótesis nos permiten hacer generalizaciones sobre ella a partir de una muestra representativa.

Por ejemplo, imagina que una empresa de tecnología quiere saber si un nuevo producto tiene una tasa de aceptación superior al 50% entre los consumidores. Para ello, realiza una encuesta a una muestra de clientes y, con base en esos datos, puede usar una prueba de hipótesis para determinar si la proporción de aceptación es estadísticamente significativa, es decir, si no es simplemente un resultado de la variabilidad aleatoria. Si se rechaza la hipótesis nula, la empresa podría decidir lanzar el producto al mercado con la seguridad de que tiene una alta probabilidad de ser bien recibido.

#### ¿Por Qué Son Importantes en la Investigación?

En el ámbito académico, las pruebas de hipótesis son esenciales para validar teorías o modelos propuestos. Como estudiante, he comprendido que no se trata solo



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



de observar patrones en los datos, sino de probar si esos patrones son estadísticamente significativos. Esto ayuda a evitar conclusiones erróneas y a asegurar que los resultados no se deben a una simple coincidencia o al azar.

Por ejemplo, en una investigación sobre la efectividad de un tratamiento médico, realizar una prueba de hipótesis es crucial para determinar si los efectos observados en el grupo experimental son significativamente diferentes de los observados en el grupo de control, y no simplemente producto de la variabilidad natural de la población. Sin estas pruebas, estaríamos tomando decisiones sin fundamentos sólidos, lo que podría tener graves consecuencias, especialmente en áreas como la medicina o la educación.

### **El Control de Calidad: Garantía de Productos y Servicios Excelentes**

El **control de calidad**, por otro lado, es otro aspecto vital que he aprendido a valorar en mis estudios. Esta práctica se refiere a las actividades y procedimientos que una empresa lleva a cabo para garantizar que sus productos o servicios cumplen con los estándares establecidos y las expectativas de los clientes. En mi opinión, el control de calidad es esencial no solo para asegurar que los productos sean seguros y eficaces, sino también para mantener la reputación de la empresa y satisfacer las demandas del mercado.

En industrias como la manufactura, por ejemplo, el control de calidad se utiliza

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

de observar patrones en los datos, sino de probar si esos patrones son para identificar defectos en los productos y corregirlos antes de que lleguen al consumidor. A través de métodos estadísticos, las empresas pueden monitorear los



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



procesos de producción y asegurarse de que los productos finales cumplan con las especificaciones. Esto no solo mejora la calidad, sino que también reduce los costos al minimizar el desperdicio y la necesidad de correcciones costosas a posteriori.

En un ejemplo práctico, si una fábrica de automóviles detecta un problema con una línea de ensamblaje que produce piezas defectuosas, el control de calidad puede alertar a la empresa antes de que los vehículos defectuosos lleguen al mercado. Esto es fundamental para garantizar la seguridad del consumidor y evitar daños a la marca.

### **El Control de Calidad y las Pruebas de Hipótesis Trabajando Juntas**

Las pruebas de hipótesis y el control de calidad a menudo trabajan de la mano. El control de calidad se basa en gran medida en los **intervalos de confianza** y las **pruebas de hipótesis** para asegurarse de que los procesos y productos estén dentro de los límites aceptables. Por ejemplo, una empresa puede usar pruebas de hipótesis para evaluar si la media de una característica de un producto, como su peso o resistencia, es igual al valor esperado, según las especificaciones del diseño.

El uso de pruebas de hipótesis en el control de calidad es especialmente útil para evaluar cambios en los procesos de producción. Si se implementa una nueva técnica o una nueva máquina, las pruebas de hipótesis pueden ayudar a verificar si esa modificación ha mejorado realmente la calidad del producto o si los cambios han

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

afectado alguna variable de manera significativa. Esto permite tomar decisiones

informadas sobre la continuidad o ajuste de los cambios.



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### Conclusión

En resumen, tanto las pruebas de hipótesis como el control de calidad son fundamentales para el análisis de datos y la toma de decisiones en cualquier contexto, ya sea académico, industrial o empresarial. Como estudiante, he aprendido que, más allá de las matemáticas y la teoría, estas herramientas me proporcionan una base sólida para enfrentar problemas complejos y tomar decisiones fundamentadas en datos confiables.

Al emplear correctamente estas herramientas, no solo puedo obtener conclusiones válidas y útiles, sino que también puedo garantizar la mejora continua de procesos, productos y servicios.

En el futuro, como profesional, sin duda seguiré aplicando estos principios en mi carrera. Ya sea evaluando la efectividad de una campaña de marketing, optimizando un proceso de fabricación o tomando decisiones políticas basadas en datos, las pruebas de hipótesis y el control de calidad serán herramientas clave para asegurar que las decisiones que tome estén basadas en hechos y no en suposiciones. Estas herramientas no solo mejoran la calidad de los productos y servicios, sino que también generan confianza en los resultados y en las decisiones que se tomen a partir de ello



# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



## **“MAPAS MENTALES”**

### **ALUMNO:**

Luis Enrique Muñoz Paulino

## **LICENCIATURA EN CONTADURIA PÚBLICA**

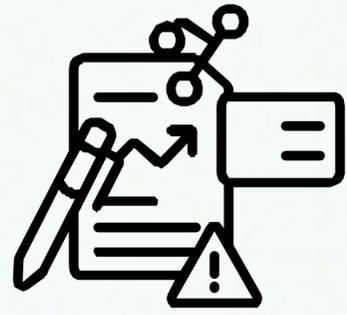
### **ASIGNATURA:**

Estadística Inferencial

### **DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 2025**



**Concepto**

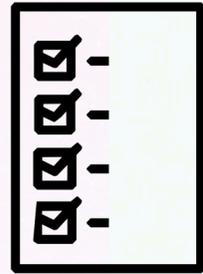
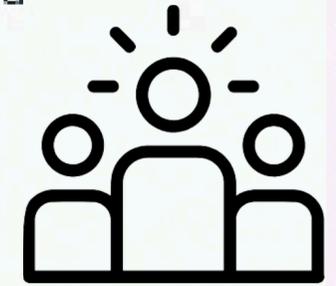
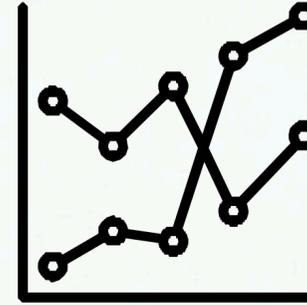
Procedimiento estadístico  
Permite tomar decisiones en  
datos muestrales

Evalua afirmaciones sobre  
una población



**Errores  
posibles**

- Error Tipo I ( $\alpha$ ): Rechazar  $H_0$  cuando es verdadera
- Error Tipo II ( $\beta$ ): No rechazar  $H_1$  cuando  $H_1$  es verdadera



**Elementos**

Hipótesis nula ( $H_0$ ):  
Afirmación inicial que se  
pone a prueba

Hipótesis alternativa ( $H_1$ ):  
Afirmación contraria a  $H_0$

Nivel de significancia ( $\alpha$ ):  
Probabilidad de cometer error ti-  
p I (generalmente 0.05)

Estadístico de prueba:  
Valor calculado para comparar  
con el valor crítico

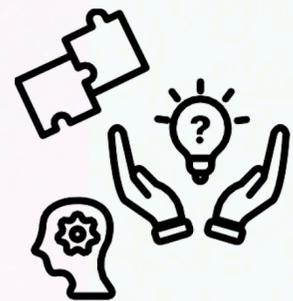
# PRUEBAS DE HIPOTESIS

1. Formular  $H_0$  y  $H_1$
2. Elegir el nivel de significancia ( $\alpha$ )
3. Determinar la distribución a usar
4. Calcular el estadístico de prueba
5. Determinar la región crítica o p-valor
6. Tomar la decisión: ¿Rechazar  $H_0$ ?



**Resultados**

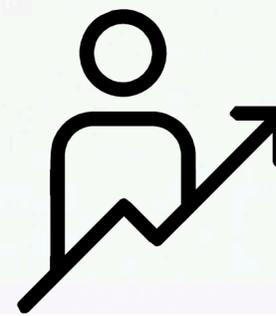
- Rechazo de  $H_0$ :  
Hay evidencia significativa
- No se rechaza  $H_0$ :  
No hay evidencia suficiente para rechazarla

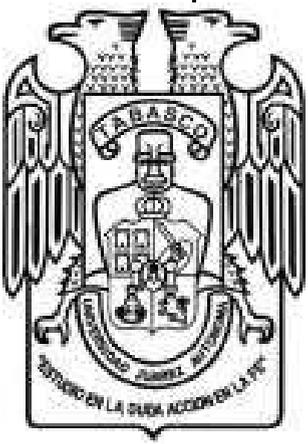


**Tipos de pruebas**

Región crítica:  
Zona donde se rechaza  $H_0$

Error Tipito I ( $\beta$ ):  
No rechazar  $H_0$  cuando  $H_1$   
es verdadera





**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



***“DOCUMENTO EXTENSO”***

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 2025**



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA



## PRUEBA DE HIPÓTESIS Y CONTROL DE CALIDAD.

En estadística, una prueba de hipótesis es un método que se usa para rechazar o aceptar una hipótesis. Es decir, una prueba de hipótesis sirve para determinar si se rechaza o se acepta una hipótesis que se tiene acerca del valor de un parámetro estadístico de una población.

En una prueba de hipótesis se analiza una muestra de datos y, a partir de los resultados obtenidos, se decide rechazar o aceptar una hipótesis de un parámetro poblacional que se había establecido previamente.

**Hipótesis y Niveles de Significancia** En la prueba de hipótesis se pone a prueba un reclamo hecho sobre la naturaleza de una población a base de la información de una muestra. El reclamo se llama hipótesis estadística.

**Definición de hipótesis:** Una hipótesis se define como una afirmación transitoria que debe ser sometida a prueba. La inferencia estadística propone un procedimiento para llevar a cabo la prueba de las hipótesis. Propone, primero, enunciarlas formalmente y luego contrastarlas con la evidencia de los datos. Son los datos, entonces, con sus características, los que dirán si una hipótesis es falsa o verdadera.

**Hipótesis Estadística:** Una hipótesis estadística es un reclamo hecho

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

sobre la naturaleza de una población. Por ejemplo, la premisa formulada por un



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA

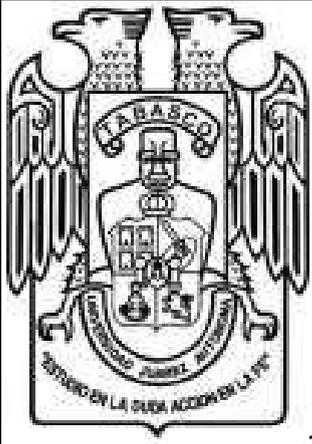


productor de baterías para autos de que su batería dura en promedio 48 meses,

Es una hipótesis estadística porque el manufacturero no inspecciona la vida de cada batería que él produce. Si surgieran quejas de parte de los clientes, entonces se pone a prueba el reclamo del manufacturero. La hipótesis estadística **sometida a prueba se llama la hipótesis nula, y se denota como  $H_0$** .

**Hipótesis Nula ( $H_0$ ):** Es una Premisa, reclamo, o conjetura que se pronuncia sobre la naturaleza de una o varias poblaciones. Por ejemplo, para probar o desaprobado el reclamo pronunciado por el productor de baterías debemos probar la hipótesis estadística de que  $\mu \geq 48$ . Por lo tanto, la hipótesis nula es:  $H_0 : \mu \geq 48$ . Luego se procede a tomar una muestra aleatoria de baterías y medir su vida media. Si la información obtenida de la muestra no apoya el reclamo en la hipótesis nula ( $H_0$ ), entonces otra cosa es cierta. La premisa alterna a la hipótesis nula se llama hipótesis alterna y se representa por  $H_1$ .

**Hipótesis Alterna:** Una premisa que es cierta cuando la hipótesis nula es falsa. Por ejemplo, para el productor de baterías  $H_0 : \mu \geq 48$  y  $H_1 : \mu < 48$  Para probar si la hipótesis nula es cierta, se toma una muestra aleatoria y se calcula la información, como el promedio, la proporción, etc. Esta información



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



muestral se llama estadística de prueba.

Estadística de Prueba: Una estadística de prueba se basa en la información de la muestra como la media o la proporción.

### 1.1 Conceptos básicos

**¿Qué es una variable?:** Una variable es una característica de interés, que tienen los individuos de una población.

- Ejemplo 1: el peso de un paquete de galletitas de cierta marca
- Ejemplo 2: la cantidad de alumnos en un mes determinado de una escuela que da clases de baile online
- Ejemplo 3: la localidad en la que está ubicada un comercio de una cierta cadena

**¿Qué es un parámetro?** En estadística, un parámetro es una constante asociada a la distribución de probabilidades de una variable aleatoria.

- Ejemplo 1: Si una variable tiene distribución binomial, sus parámetros son  $n$  y  $p$ .
- Ejemplo 2: Si una variable tiene distribución normal, sus parámetros son  $\mu$  y  $\sigma$ .

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

muestral se llama estadística de prueba.

- Ejemplo 3:



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Si una variable tiene distribución Bernoulli, su único parámetro es la probabilidad de éxito  $p$ .

**¿Qué es un estimador de un parámetro?** El estimador de un parámetro es un estadístico (estadístico: variable aleatoria función de las observaciones muestrales) que toma «valores cercanos» al verdadero valor del parámetro. Fundamentalmente nos interesan los siguientes estimadores:

- La media muestral  $\bar{X}$  es un estimador de la media poblacional  $\mu$ .
- El desvío estándar muestral  $S$  es un estimador del desvío estándar poblacional  $\sigma$ .
- La proporción muestral  $\hat{p}$ , es un estimador de la proporción poblacional  $p$ .

Parámetros poblacionales	Estimadores de los parámetros
$\mu$	$\bar{X}$
$\sigma^2$	$S^2$
$\sigma$	$S$
$p$	$\hat{p}$

(Parámetros poblacionales y estimadores de esos parámetros)

¿Qué es un estadístico de prueba? Esta es la definición más difícil de esta serie.

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

Si una variable tiene distribución Bernoulli, su único parámetro es la  
Por eso vamos a dar muchos ejemplos. Un estadístico de prueba es:



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



- Una variable aleatoria
- De distribución conocida
- Que vincula a un parámetro de interés, con un estimador de ese parámetro.

**¿Qué es una hipótesis estadística?** Una hipótesis estadística es una afirmación acerca de la distribución de una variable aleatoria.

- Si la afirmación es sobre el valor de un parámetro, es una hipótesis estadística paramétrica.
- Si la afirmación es sobre la forma de la distribución de probabilidades, es una hipótesis estadística no paramétrica.

**ejemplos de afirmaciones y establezcamos si son o no son hipótesis estadísticas.**

- Afirmación 1: «La molécula de agua está compuesta por dos átomos de hidrógeno y un átomo de oxígeno».
- Afirmación 2: «La variable X tiene distribución Binomial». 10
- Afirmación 3: «La media de una muestra de 100 observaciones es de  $X = 45,32$  gramos»



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



- Afirmación 4: «La media de la variable  $X \sim (\mu, \sigma)$  es  $\mu = 134$ »
- Afirmación 5: «La media de la variable  $X \sim (\mu, \sigma)$  es  $\mu < 134$ »

La afirmación 1 predica sobre la composición de una molécula. Será una afirmación... pero no es de tipo «estadístico». Y en particular no dice nada sobre ninguna variable aleatoria. No es una hipótesis estadística.

La afirmación 2 predica sobre la forma que tiene la distribución de una variable aleatoria. Está diciendo algo acerca de una variable aleatoria. Así que podemos decir que es una hipótesis estadística. Pero es una hipótesis estadística no paramétrica.

La afirmación 3 es sobre la media muestral de una variable aleatoria. La media muestral no es un parámetro, sino que es una variable aleatoria. No es una hipótesis estadística porque no afirma nada ni sobre el tipo de distribución (binomial, normal, etc...) ni sobre sus parámetros.

La afirmación 4 sí es una hipótesis estadística (paramétrica) porque asevera que el parámetro medio poblacional de cierta variable es igual a 134.

**Ejemplo de prueba de hipótesis detallado paso a paso (incluyendo los detalles**

**UNIVERSIDAD JUÁREZ AUTÓNOMA  
DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**

«técnicos»)



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Un fabricante de galletitas produce paquetes en los cuales el peso nominal impreso es de 500 gramos. Pero el contenido real es una variable aleatoria con distribución normal. No tienen exactamente 500 gramos todos los paquetes. El fabricante, basándose en información histórica, afirma que la media de esa variable  $X$  es  $\mu=500$  gramos con un desvío estándar de 5 gramos. Se desconfía de la afirmación del fabricante acerca de que  $\mu = 500$  gramos. Se quiere analizar si en realidad el peso promedio de los paquetes es inferior a 500 gramos. Para esto se toma una muestra de tamaño 100, y se obtiene una media muestral de 497,3 gramos. Realizar una prueba de hipótesis con un nivel de significación de 0,05.

**La variable** La variable sobre la que vamos a trabajar es  $X$ : peso real de un paquete de galletitas de 500 gramos de la fábrica. El enunciado afirma que la distribución de  $X$  es normal. No se conoce  $\mu$ , pero sí se conoce  $\sigma = 5$ .

**Hipótesis nula e hipótesis alternativa**  $H_0$  es la hipótesis nula. Hipótesis nula es la hipótesis de no cambio. Es la hipótesis de que todo queda igual. Hay otra hipótesis que es la hipótesis de cambio. Siempre en esta hipótesis va a estar el igual.  $H_0: \mu = 500$   $H_1$  es la hipótesis alternativa. Es complementaria a la nula. Niega a la hipótesis nula.

**Las hipótesis son complementarias.** La nula tiene el símbolo de igualdad siempre. Entonces en la otra no puede aparecer el igual. En la hipótesis alternativa se utiliza o bien el símbolo de distinto  $\neq$ , o bien mayor  $>$  o bien menor  $<$ .



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



### 1.2 TIPOS DE ERRORES

#### ERROR TIPO 1 Y ERROR TIPO 2

A base de la información de una muestra nosotros podemos cometer dos tipos de errores en nuestra decisión.

1. Podemos rechazar un  $H_0$  que es cierto.
2. Podemos aceptar un  $H_0$  que es falso.

El primero se llama error Tipo 1, **Error Tipo 1**: Cuando rechazamos una Hipótesis Nula que es cierta cometemos error tipo 1.

Y el segundo error se llama error Tipo 2.

**Error Tipo 2**: Cuando aceptamos una Hipótesis Nula que es falsa cometemos error tipo 2.

**NIVEL DE SIGNIFICANCIA ( $\alpha$ )**; Para ser muy cuidadosos en no cometer el error tipo 1, debemos especificar la probabilidad de rechazar  $H_0$ , denotada por  $\alpha$ . A ésta se le llama nivel de significancia.

**Nivel de Significancia**: La probabilidad ( $\alpha$ ) más alta de rechazar  $H_0$  cuando  $H_0$  es

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

cierto se llama nivel de significancia.



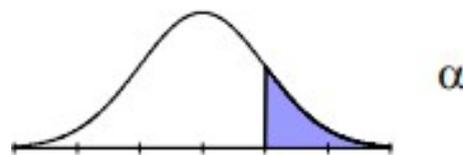
# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Comentario: Para mantener la probabilidad de cometer el error tipo 1 baja, debemos escoger un valor pequeño de  $\alpha$ . Usando un valor preasignado de  $\alpha$  se construye una región de rechazo o región crítica en la curva normal estándar o en la curva  $t$  que indica si debemos rechazar  $H_0$ .

Región Crítica o de Rechazo: Una región crítica o de rechazo es una parte de la curva de  $z$  o de la curva  $t$  donde se rechaza  $H_0$ . La región puede ser de una cola o de dos dependiendo de la hipótesis alterna.



### 2.3 PRUEBA DE HIPOTESIS PARA MEDIA Y PROPORCIÓN POBLACIONAL (POBLACIÓN FINITA - MUESTRAS GRANDE)

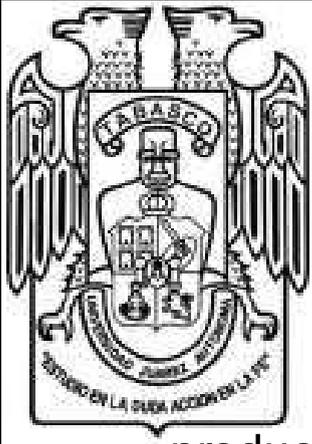
Debido a la dificultad de explicar este tema se enfocará un problema basado en un estudio en una fábrica de llantas.

En este problema la fábrica de llantas tiene dos turnos de operarios,

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

turno de día y turno mixto. Se selecciona una muestra aleatoria de 100 llantas



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



producidas por cada turno para ayudar al gerente a sacar conclusiones de cada una de las siguientes preguntas:

1.- ¿Es la duración promedio de las llantas producidas en el turno de día igual a 25 000 millas?

2.- ¿Es la duración promedio de las llantas producidas en el turno mixto menor de 25 000 millas? 3.- ¿Se revienta más de un 8% de las llantas producidas por el turno de día antes de las 10 000 millas?

Prueba De Hipótesis Para La Media En la fábrica de llantas la hipótesis nula y alternativa para el problema se plantearon como sigue:

$$H_0: \mu = 25\ 000$$

$$H_1: \mu \neq 25\ 000$$

Si se considera la desviación estándar  $\sigma$  las llantas producidas en el turno de día, entonces, con base en el teorema de límite central, la distribución en el muestreo de la media seguiría la distribución normal, y la prueba estadística que esta basada en la diferencia entre la media de la muestra y la media  $\mu$  hipotética se encontrara como sigue: Si el tamaño de la región  $\alpha$  de rechazo se estableciera en 5% entonces se podrían determinar los valores críticos de la distribución.

Dado que la región de rechazo esta dividida en las dos colas de la distribución, el

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

5% se divide en dos partes iguales de 2.5%.

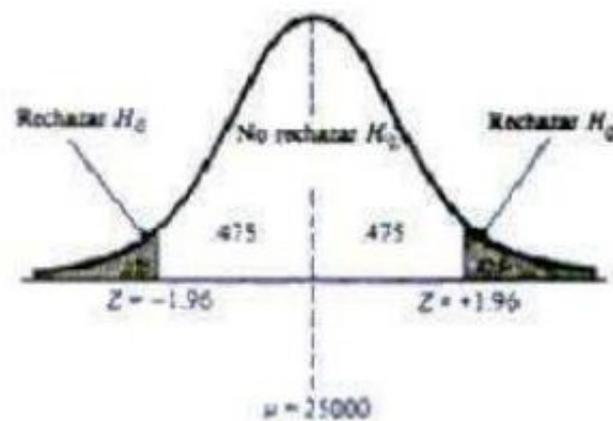


# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



Dado que ya se tiene la distribución normal, los valores críticos se pueden expresar en unidades de desviación. Una región de rechazo de 0.25 en cada cola de la distribución normal, da por resultado un área de .475 entre la media hipotética y el valor crítico. Si se busca esta área en la distribución normal, se encuentra que los valores críticos que dividen las regiones de rechazo y no rechazo son  $+ 1.96$  y  $- 1.96$



Por tanto, la regla para decisión sería: Rechazar  $H_0$  si  $Z > + 1.96$  o si  $Z < - 1.96$

De lo contrario, no rechazar  $H_0$ . No obstante, en la mayor parte de los casos se desconoce la desviación estándar de la población. La desviación estándar se estima al calcular  $S$ , la desviación estándar de la muestra. Si se supone que la



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



población es normal la distribución en el muestreo de la media seguiría una distribución t con n-1 grados de libertad.

En la práctica, se ha encontrado que siempre y cuando el tamaño de la muestra no sea muy pequeño y la población no este muy sesgada, la distribución t da una buena aproximación a la distribución de muestra de la media. La prueba estadística para determinar la diferencia entre la media de la muestra y la media de la población cuando se utiliza la desviación estándar S de la muestra, se expresa con:

$$\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \quad t_{n-1}$$

+

Para una muestra de 100, si se selecciona un nivel de significancia de .05, los valores críticos de la distribución t con  $100-1=99$  grados de libertad se puede obtener como se indica en la siguiente tabla: Como esta prueba de dos colas, la región de rechazo de .05 se vuelve a dividir en dos partes iguales de .025



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

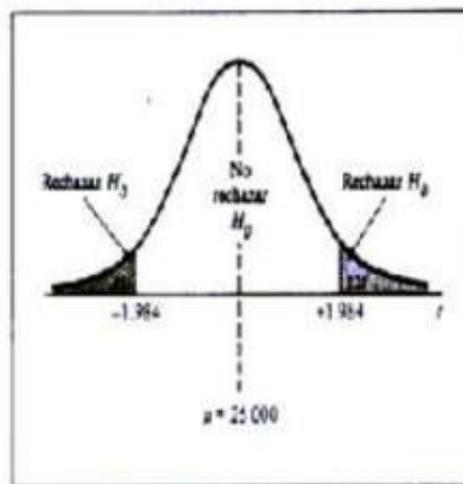
## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



para  $t$ , los valores críticos son  $-1.984$  y  $+1.984$ . la regla para la decisión es:

Rechazar  $H_0$  si  $t > +1.984$  o si  $t < -1.984$ . De lo contrario, no rechazar  $H_0$

al 10



JRA 03

Por ello, la decisión de no rechazar la hipótesis nula  $H_0$ . En conclusión es que la duración promedio de las llantas es 25 000 millas. A fin de tener en cuenta la posibilidad de un error de tipo II, este enunciado se puede redactar como "no hay pruebas de que la duración promedio de las llantas sea diferente a 25 000 millas en las llantas producidas en el turno de día".

### PRUEBA DE HIPÓTESIS PARA PROPORCIONES E

concepto de prueba de hipótesis se puede utilizar para probar hipótesis en relación con datos cualitativos. Por ejemplo, en el problema anterior el gerente de la fábrica de llantas quería determinar la proporción de llantas que se reventaban antes de 10,000 millas. Este es un ejemplo de una variable

# UNIVERSIDAD JUÁREZ AUTÓNOMA

cuantitativa, dado que se desea llegar a conclusiones en cuanto a la proporción



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



de los valores que tienen una característica particular. El gerente de la fábrica de llantas quiere que la calidad de llantas producidas, sea lo bastante alta para que muy pocas se revienten antes de las 10,000 millas. Si más de un 8% de las llantas se revientan antes de las 10,000 millas, se llegaría a concluir que el proceso no funciona correctamente. La hipótesis nula y alternativa se pueden expresar como sigue:  $H_0: p \leq .08$  (funciona correctamente)  $H_1: p > .08$  (no funciona correctamente)

$$Z = \frac{p - P}{\sqrt{\frac{P(1-P)}{n}}}$$

La prueba estadística se puede expresar en términos de la proporción de éxitos como sigue: En donde  $p$  = proporción de éxitos de la hipótesis nula

Ahora se determinará si el proceso funciona correctamente para las llantas producidas para el turno de día. Los resultados del turno de día indican que cinco llantas en una muestra de 100 se reventaron antes de 10,000 millas para este problema, si se selecciona un nivel de significancia de .05, las regiones de rechazo y no rechazo se establecerían como a continuación se muestra: Para ver

# UNIVERSIDAD JUÁREZ AUTÓNOMA

el gráfico seleccione la opción "Descargar" del menú superior Y la regla de



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



decisión sería: Rechazar  $H_0$  si  $>+ 1.645$ ; de lo contrario no rechazar  $H_0$ .

Con los datos que se tienen, Y entonces,

$$Z \cong \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}}$$

La hipótesis nula no se rechazaría por que la prueba estadística no ha caído en la región de rechazo. Se llegaría a la conclusión de que no hay pruebas de que más del 8% de las llantas producidas en el turno de día se revienten antes de 10,000 millas. El gerente no ha encontrado ninguna prueba de que ocurra un número excesivo de reventones en las llantas producidas en el turno de día.

## 2.4 PRUEBA DE HIPOTESIS PARA DIFERENCIA DE 2 MEDIAS Y PROPORCIONES POBLACIONALES.

La prueba de hipótesis para la diferencia de medias es una prueba estadística que sirve para rechazar o aceptar la hipótesis de que las medias de dos poblaciones son diferentes. Es decir, la prueba de hipótesis para la diferencia de medias se usa para determinar si dos medias poblacionales son iguales o distintas.

Ten en cuenta que las decisiones tomadas en las pruebas de hipótesis se basan

# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

## **DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA**

en un nivel de confianza previamente establecido, por lo que no se puede



# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



garantizar que el resultado de una prueba de hipótesis sea siempre el acertado, sino que es el resultado más probable de que sea verdad.

FORMULA: Para realizar la prueba de hipótesis para la diferencia entre dos medias, la aplicación realiza los siguientes pasos:

1. Plantear la hipótesis nula,  $H_0$ , y una hipótesis alternativa,  $H_1$ .
2. Obtener dos muestras de tamaño  $n_1$  y  $n_2$ , respectivamente.
3. Calcular las proporciones de la media,  $p_1$  y  $p_2$ , de cada muestra.
4. Obtener el valor  $z_1$  que corresponde al nivel de confianza dado. Si el tamaño de las dos muestras es mayor o igual a 30,  $z_1$  se obtiene a partir de la distribución normal. En otro caso,  $z_1$  se obtiene a partir de la  $t$  de Student.
5. Calcular el valor  $z_2$  usando las siguientes ecuaciones:

$$p = \frac{n_1 \times p_1 + n_2 \times p_2}{n_1 + n_2}$$
$$z_2 = \frac{p_1 - p_2}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}}$$

Act

dónde:

$n_1$  y  $p_1$  son el tamaño y la proporción de la media, respectivamente, de la primera muestra.  $n_2$  y  $p_2$  son el tamaño y la proporción de la media, respectivamente, de la segunda muestra.

Dependiendo de la hipótesis alternativa, la aplicación toma una decisión:

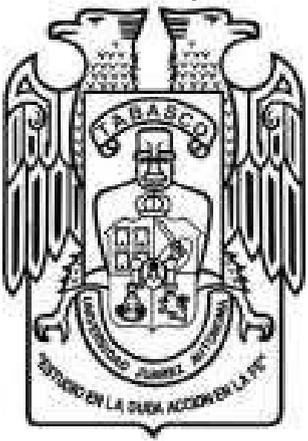


# UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

## DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO ADMINISTRATIVA



- a) Para el caso  $H_1: p_1 \neq p_2$ , si  $|z_2| > z_1$ , entonces se rechaza  $H_0$  y se acepta  $H_1$ . En otro caso, se acepta  $H_0$ .
- b) Para el caso  $H_1: p_1 > p_2$ , si  $z_2 < z_1$ , entonces se rechaza  $H_0$  y se acepta  $H_1$ . En otro caso, se acepta  $H_0$ .
- c) Para el caso  $H_1: p_1 < p_2$ , si  $z_2 > z_1$ , entonces se rechaza  $H_0$  y se acepta  $H_1$ . En otro caso, se acepta  $H_0$ .



# **UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO**

**DIVICIÓN ACADÉMICA DE CIENCIAS ECONÓMICO  
ADMINISTRATIVA**



**“PRACTICAS DE LABORATORIO DE COMPUTO”**

**ALUMNO:**

Luis Enrique Muñoz Paulino

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:**

Estadística Inferencial

**DOCENTE:**

Cesar Enrique Barrera Ovando

**VILLAHERMOSA, TABASCO, 07 DE MARZO  
DEL 202**

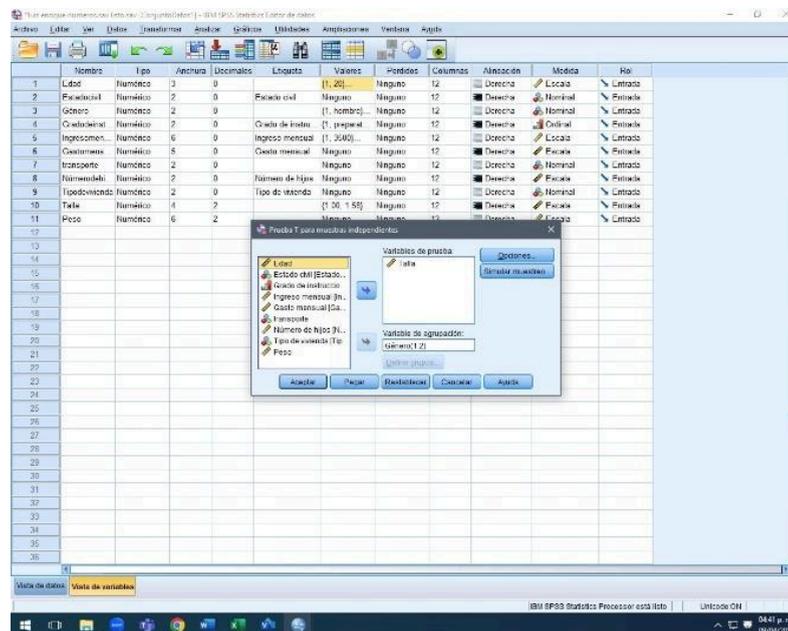
## INTRODUCCIÓN

El presente documento tiene como objetivo servir como evidencia de los temas abordados en la primera unidad, titulada "Teorema del Límite Central y Estimación", de la materia de Estadística Inferencial. Para ello, se llevó a cabo una práctica en el centro de cómputo de la universidad, la cual debía resolverse utilizando el software IBM SPSS Statistics.

El documento incluye una recopilación de los ejercicios realizados con variables de un conjunto de datos, así como capturas de pantalla que muestran los resultados obtenidos, incluyendo tablas y gráficos generados por el software. Además, se presenta un informe que contiene la interpretación de los resultados obtenidos durante el análisis.

SPSS Statistics es un software de análisis de datos que podemos utilizar para llevar a cabo tareas estadísticas y de investigación. Con él, podemos gestionar y analizar grandes volúmenes de datos, realizar análisis descriptivos, crear gráficos y aplicar modelos estadísticos complejos. Nos permite trabajar con diferentes tipos de datos, como encuestas, experimentos o bases de datos numéricas, y tiene herramientas que facilitan el análisis de regresión, ANOVA, correlaciones, y más. También tiene una interfaz gráfica que me permite realizar estos análisis sin necesidad de ser experto.

# CONJUNTO DE DATOS



## (VISTA DE DATOS)

En esta actividad utilicé el programa SPSS para aplicar una prueba t de muestras independientes, que sirve para comparar las medias de dos grupos y ver si hay diferencias significativas entre ellas.

En la ventana que aparece en la imagen, seleccioné como variable dependiente la "Talla", lo que significa que estoy analizando si el promedio de la talla es diferente entre dos grupos.

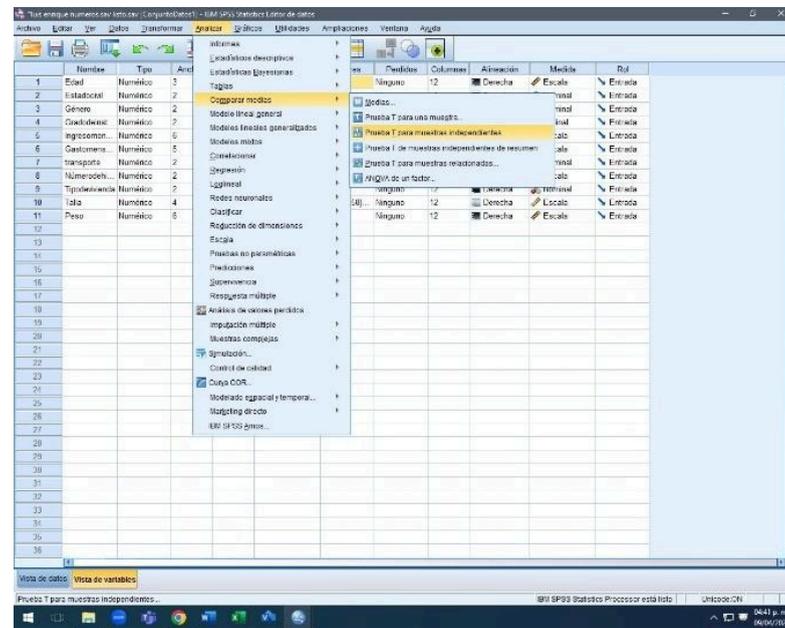
Como variable de agrupación, elegí "Grado del Estrés", lo cual indica que quiero comparar la talla de los participantes en función del nivel de estrés que tienen.

Primero, definí los grupos dentro de la variable "Grado del Estrés", y después pasé la variable "Talla" al cuadro correspondiente. Esto me permitirá saber si existe una diferencia estadísticamente significativa entre las tallas de las personas con diferentes niveles de estrés.

Este procedimiento es útil porque me ayuda a ver si una característica como el estrés tiene algún efecto en otra como la talla, al menos dentro de esta muestra.

Además, con SPSS es mucho más fácil hacer estos análisis, ya que el software realiza los cálculos automáticamente y

proporciona los resultados necesarios para interpretar si la diferencia entre grupos es significativa o no, basándose en el valor de p (significancia).



En esta imagen muestro cómo accedí al procedimiento para realizar una prueba t para muestras independientes en SPSS. Para hacerlo, primero me fui al menú "Analizar", que se encuentra en la parte superior de la ventana del programa. Después seleccioné la opción "Comparar medias", y dentro del submenú que se despliega, elegí "Prueba T para muestras independientes... ". Esta opción me permite comparar las medias de dos grupos diferentes para ver si hay una diferencia estadísticamente significativa entre ellas.

## 1. comparación de la Talla con los grupos hombre y mujer.

The image shows the SPSS output window for a T-test. The main table displays the results for the 'Talla' variable, comparing 'Hombre' and 'Mujer' groups. The output includes descriptive statistics, Levene's test for homogeneity of variances, and the T-test results.

Prueba de hipótesis de igualdad de medias									
Prueba de Levene de igualdad de varianzas									
		F	Sig.		t	df	Sig. (bilateral)	Diferencia de medias	95% de intervalo de confianza de la diferencia
Talla	Se asumen varianzas iguales	30.417	.000		8.814	878	.000	743.344	611.173 - 875.515
	NO SE ASUMEN VARIANZAS IGUALES				9.092	877.425	.000	743.249	61.843 - 873.219

Prueba T									
Estadísticas de grupo									
		N	Media	Desviación estándar	Desv. est. corregida				
Talla	Hombre	370	1,7301	10,760	,08559				
	Mujer	511	1,7223	11,210	,08592				

Prueba de hipótesis de igualdad de medias									
Prueba de Levene de igualdad de varianzas									
		F	Sig.		t	df	Sig. (bilateral)	Diferencia de medias	95% de intervalo de confianza de la diferencia
Talla	Se asumen varianzas iguales	.000	.957		-.257	878.394	.797	-.0102	-.30752 - .28710
	NO SE ASUMEN VARIANZAS IGUALES								

2.

### **Introducción:**

Realicé una **prueba t de muestras independientes** para comparar la variable **Talla** entre dos grupos: **hombres y mujeres**. El objetivo era determinar si existe una diferencia significativa en la media de la talla entre ambos géneros.

### **Interpretación:**

- La media de talla en hombres fue de **1.7901** con una desviación estándar de **0.10760**.
- La media de talla en mujeres fue de **1.7920** con una desviación estándar de **0.11370**.

Esto ya me da una idea de que las medias son bastante similares, En la **Prueba de Levene para igualdad de varianzas**, el valor de **significación (Sig.)** fue **0.437**, lo cual es mayor que 0.05. Esto indica que **se asumen varianzas iguales**, así que interpreto la primera fila de la prueba t. En la fila de "**Se asumen varianzas iguales**", el valor de significancia (Sig. bilateral) fue de **0.799**, que también es mucho mayor que 0.05. Esto quiere decir que **no hay una diferencia estadísticamente significativa** entre la talla de hombres y mujeres en esta muestra.

### **Conclusión:**

La diferencia es tan pequeña que, estadísticamente, no se considera significativa. Por lo tanto, puedo concluir que **no existe evidencia suficiente para afirmar que hay una diferencia real de talla entre hombres y mujeres** en este conjunto de datos..

### **3. Gastos mensuales**

The screenshot shows the SPSS output window with the following data:

**Prueba de Levene para igualdad de varianzas**

	F	Sig.	1	2	Sig. (bilateral)	Diferencia de medias	Diferencia de intervalos	% de intervalos de confianza de la diferencia
Prueba de Levene para igualdad de varianzas	12,97	,000	3,07	879	,000	2,642	891	4,285
Prueba de Levene para igualdad de varianzas			3,184	874,87	,000	-2,642	891	4,275

**Estadísticas de grupo**

Grupo	N	Media	Desv. Est.
Prueba de Levene para igualdad de varianzas	370	3343,74	1045,09
Prueba de Levene para igualdad de varianzas	511	4087,09	1383,59

**Prueba de muestras independientes**

	F	Sig.	1	2	Sig. (bilateral)	Diferencia de medias	Diferencia de intervalos	% de intervalos de confianza de la diferencia
Prueba de Levene para igualdad de varianzas	37,412	,000	-3,653	879	,000	-743,349	85,510	861,877
Prueba de Levene para igualdad de varianzas			-4,803	877,428	,000	-743,349	85,510	861,878

**Introducción:** En esta parte del análisis trabajé con la variable Gasto mensual para comparar si existe una diferencia significativa entre hombres y mujeres. Utilicé una prueba t para muestras independientes, y como se puede ver en la imagen, los resultados ya están generados.

**Interpretación:** en la tabla de **Estadísticas de grupo**, observamos que:

- Los hombres (n = 370) tienen un gasto mensual promedio de **3343.74** con una desviación estándar de **1045.09**.
- Las mujeres (n = 511) tienen un gasto mensual promedio de **4087.09** con una desviación estándar de **1383.59**.

A simple vista, ya se nota que las mujeres tienen un gasto promedio más alto que los hombres. La **Prueba de Levene para igualdad de varianzas** tuvo un valor de **significancia (Sig.) de 0.000**, que es menor a 0.05, es decir **no se asumen varianzas iguales**

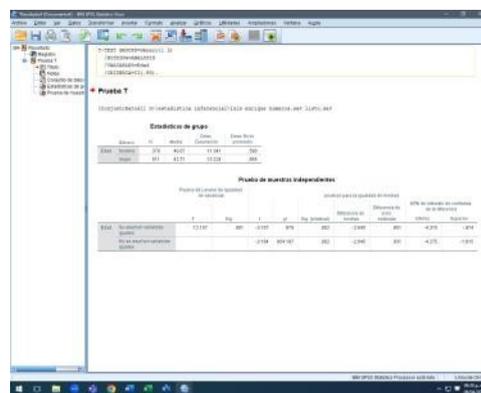
En la segunda fila, el valor de **Sig. (bilateral)** es también **0.000**, lo cual es menor que 0.05.

Esto me indica que **sí existe una diferencia significativa** en el gasto mensual entre

hombres y mujeres.

**Conclusión:** la diferencia de medias es de -743.349, lo que confirma que, en promedio, las mujeres gastan 743.35 unidades más que los hombres. El intervalo de confianza del 95% va desde -903.979 hasta -582.719, y como no contiene el cero, se refuerza la conclusión de que la diferencia es significativa.

#### 4. EDAD PROMEDIO ENTRE HOMBRES Y MUJERES



The screenshot shows the SPSS output for an independent samples t-test. The 'Estadísticas de grupo' table shows the mean age for men (40.07) and women (42.71). The 'Prueba de muestras independientes' table shows the t-statistic (-2.291) and the p-value (.025).

Estadísticas de grupo	
Edad	Sexo
Media	Hombres
40,07	370
Desviación estándar	
11,341	
Media	Mujeres
42,71	511
Desviación estándar	
13,224	

Prueba de muestras independientes				
Prueba de Levene para la igualdad de varianzas	Prueba t para la igualdad de medias			
F	t	df	Sig.	Intervalo de confianza del 95%
1,048	-2,291	879	,025	-743,349 (-903,979, -582,719)
Edad				
Prueba de Levene para la igualdad de varianzas				
F	t	df	Sig.	Intervalo de confianza del 95%
1,048	-2,291	879	,025	-743,349 (-903,979, -582,719)

**Introducción:** En este análisis realicé una **prueba t para muestras independientes** para comparar la **edad promedio** entre hombres y mujeres. Utilicé el programa SPSS con el archivo de datos correspondiente y seleccioné las variables "Edad" y "Género".

**Interpretación:** En la tabla de Estadísticas de grupo, podemos ver que:

Los hombres (n = 370) tienen una edad promedio de 40.07 años con una desviación estándar de 11.341.

Las mujeres (n = 511) tienen una edad promedio de 42.71 años con una desviación estándar de 13.224.

Ya desde aquí se observa una ligera diferencia en las edades, pero para saber si esa diferencia es estadísticamente significativa, revisé los resultados de la prueba t.

observamos la Prueba de Levene, que evalúa si las varianzas son iguales. El valor de

significancia es 0.001, que es menor a 0.05, no se asumen varianzas iguales, y debemos interpretar la segunda fila del análisis t.

En esa segunda fila, el valor de Sig. (bilateral) es 0.002, que también es menor a 0.05, lo cual indica que sí hay una diferencia significativa en la edad promedio entre hombres y mujeres.

Conclusión: puedo concluir que existe una diferencia estadísticamente significativa en la edad promedio entre hombres y mujeres, siendo la edad promedio de las mujeres mayor que la de los hombres.

## CONCLUSIÓN

A lo largo de estas prácticas con SPSS, realicé varias pruebas t para muestras independientes con el objetivo de comparar distintos grupos en base a variables como edad, talla y gasto mensual. Cada uno de estos análisis me permitió aplicar los conocimientos adquiridos en la materia de Estadística Inferencial, especialmente en lo relacionado con el Teorema del Límite Central, la estimación de parámetros y la prueba de hipótesis.

Con base en los resultados obtenidos: Se encontró que existen diferencias estadísticamente significativas entre hombres y mujeres en variables como edad y gasto mensual, ya que los valores de significancia fueron menores a 0.05. Esto indica que las diferencias observadas no son producto del azar. En el caso de la talla, la diferencia no fue significativa, lo que implica que, en esta muestra, hombres y mujeres no difieren de manera importante en esa variable.

También aprendí a interpretar adecuadamente la Prueba de Levene para verificar la igualdad de varianzas, lo cual es esencial para seleccionar correctamente qué fila de resultados interpretar dentro de la prueba t. Además, comprendí cómo utilizar e interpretar los intervalos de confianza, las medias, las desviaciones estándar y los valores p, reforzando mi capacidad de análisis estadístico.

**UNIVERSIDAD JUAREZ AUTONOMA DE TABASCO**

**DIVISION ACADEMICA DE CIENCIAS ECONOMICO ADMINISTRATIVAS**

**EXAMEN 2do. Parcial ESTADISTICA INFERENCIAL**

NOMBRE: LUIS ENRIQUE MUÑOZ PAULINO FECHA: 21/05/2025

Grupo: ELS P

UTILICE LA BASE DE DATOS "Para ponencia.sav", y realice los siguientes ejercicios tal y como lo realizo en la practica

- 1.- DE ACUERDO A LA VARIABLE "APOYO EN LA TECNOLOGIA" (APTE), ELIJA UNA PREGUNTA Y REALICE EL CONTRASTE DE PRUEBA DE HIPOTESIS (con cualquier variable cualitativa), CON LA ESTRUCTURA DE HIPOTESIS INVESTIGADA.
- 2.- DE ACUERDO A LA VARIABLE "APOYO Y CONFIANZA" (APCO), ELIJA UNA PREGUNTA Y REALICE EL CONTRASTE DE PRUEBA DE HIPOTESIS (con cualquier variable cualitativa), CON LA ESTRUCTURA DE HIPOTESIS INVESTIGADA.

## ¿Cuántos hombres y cuantas mujeres utilizan el apoyo tecnológico para sus actividades académicas?

### Prueba T

**Estadísticas de grupo**

	Género del docente	N	Media	Desviación estándar	Media de error estándar
Utilizo la tecnología como apoyo para el desarrollo de mis actividades académicas	HOMBRE	47	4.57	.500	.073
	MUJER	56	4.48	.632	.084

**Prueba de muestras independientes**

		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
									Inferior	Superior
Utilizo la tecnología como apoyo para el desarrollo de mis actividades académicas	Se asumen varianzas iguales	4.806	.031	.811	101	.419	.092	.114	-.134	.318
	No se asumen varianzas iguales			.827	100.667	.410	.092	.112	-.129	.314

En la imagen se muestra una prueba estadística t para muestras independientes, cuyo propósito es determinar si realmente existen diferencias estadísticamente significativas entre hombres y mujeres en cuanto al uso de la tecnología como herramienta de apoyo en sus actividades académicas. En la muestra que se realizó estaban compuesta por 47 hombres y 56 mujeres, en el caso de los hombres obtuvieron una media de 4.57, mientras que en las mujeres su media fue de 4.48, lo que nos indica que hay una ligera diferencia a favor de los hombres,

Verifique la prueba para poder igualar varianzas lo cual el resultado me dio  $p=0.031$ , lo que nos indica que las varianzas no son iguales, por esta razón pude tomar los valores de la fila donde no se asumen varianzas iguales.

En la fila el valor t fue de 0.827 con el gl de 1000'.67 y un valor de p bilateral de 0.410, así obstante no se rechaza la hipótesis nula.

Con base en los resultados, concluí que no existen diferencias estadísticamente significativas entre hombres y mujeres respecto al uso de la tecnología como apoyo en sus actividades académicas. Aunque los hombres muestran una media ligeramente superior a la de las mujeres, esta diferencia no es lo suficientemente grande como para considerarse significativa.

## 2.- ¿Cómo se sienten las personas debido a los tratos de las autoridades?

### → Prueba T

[Conjunto\_de\_datos1] C:\Users\Administrador\Downloads\para ponencia.sav

Estadísticas de grupo

	Género del docente	N	Media	Desviación estándar	Media de error estándar
Me siento a gusto con la imagen y prestigio que tengo ante mis pares	HOMBRE	47	4.57	.651	.095
	MUJER	56	4.70	.464	.062

Prueba de muestras independientes

		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
									Inferior	Superior
Me siento a gusto con la imagen y prestigio que tengo ante mis pares	Se asumen varianzas iguales	7.609	.007	-1.107	101	.271	-.122	.110	-.341	.097
	No se asumen varianzas iguales			-1.076	81.240	.285	-.122	.113	-.348	.104

Pude realizar una prueba t para comparar si existen diferencias significativas entre los hombres y las mujeres, respecto a la siguiente imagen pude comprender que en la muestra se incluyeron 47 docentes hombres y 56 docentes mujeres.

Primero, verifiqué la prueba el archivo para comprobar si las varianzas eran iguales. En el valor de significancia fue 0.007, menor a 0.05, por lo que no se asumen varianzas iguales y tomé la segunda fila de la prueba t.

En cuanto a la comparación de medias, los hombres obtuvieron una media muy acertada de 4.57 y las mujeres 4.70, lo que indica que las mujeres reportaron sentirse ligeramente mejor respecto a su imagen y prestigio. Sin embargo, el valor p lo que significa bilateral, la cual fue de 0.285, y lo que nos dice que es mayor a 0.05.

Por tanto, puedo concluir que no existen diferencias estadísticamente significativas entre los hombres y mujeres en cuanto a cómo se sienten con su imagen y prestigio ante sus pares. Aunque las mujeres tienen una media más alta, esta diferencia no es lo suficientemente grande como para ser considerada significativa en términos estadísticos.

## CONCLUSIÓN SOBRE EL PARCIAL

El Teorema del Límite Central me ayudó a entender cómo, al aumentar el tamaño de una muestra, la distribución de la media muestral tiende a ser normal, sin importar la forma de la población original. Esto es clave para justificar muchas de las técnicas inferenciales que usamos en estadística.

Además, al trabajar con estimaciones puntuales e intervalos de confianza, pude ver cómo es posible hacer inferencias sobre parámetros poblacionales a partir de datos muestrales, lo que considero una habilidad muy valiosa aplicación práctica de estos conceptos utilizando el software SPSS. Al desarrollar ejercicios y prácticas con esta herramienta, pude llevar la teoría a un contexto más real y aplicable, reforzando mis habilidades para el análisis de datos con tecnología especializada.

En general, esta unidad me permitió integrar conocimientos teóricos y prácticos, desarrollar mi pensamiento crítico y mejorar mis competencias en el uso de herramientas estadísticas. Me siento más preparado para enfrentar los siguientes temas del curso y aplicar estos aprendizajes en contextos académicos y profesionales.

# **UNIDAD 3 Y 4**

## **ANOVA, CORRELACIÓN, REGRESIÓN LINEAL Y SERIES DE TIEMPO**

### **CONTENIDO:**

**ENSAYO  
MAPA MENTAL  
DOCUMENTO EXTENSO  
PRACTICAS DEL LABORATORIO DE COMPUTO  
EXAMEN DE LA UNIDAD  
CONCLUSIÓN DE LA UNIDAD**



**UNIVERSIDAD JUÁREZ  
AUTÓNOMA DE TABASCO**  
**DIVICIÓN ACADÉMICA DE CIENCIAS  
ECONÓMICO ADMINISTRATIVA**



**"ENSAYO"**

**ALUMNO:** Luis Enrique Muñoz Paulino

**MATRICULA:** 241B38018

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:** Estadística Inferencial

**DOCENTE:** Cesar Enrique Barrera Ovando

VILLAHERMOSA, TABASCO, 30 DE MAYO DEL 2025



## INTRODUCCIÓN

El siguiente ensayo representa mi análisis, opinión y el aprendizaje sobre los temas de las unidades tres y cuatro de la materia de Estadística Inferencial. En esta ocasión, vimos sobre los temas de ANOVA, correlación, regresión lineal y series de tiempo, los cuales me permitieron ampliar mis conocimientos sobre el análisis de datos y su aplicación en distintas situaciones reales. A lo largo del estudio de estos contenidos, se vio cómo podemos identificar relaciones entre variables, comparar grupos mediante análisis de varianza, y construir modelos que permiten hacer predicciones con base en datos previos.

Además, el tema de series de tiempo aun no lo manejo al cien por ciento, pero si pude entender cómo se pueden analizar fenómenos que cambian a lo largo del tiempo, lo cual es muy útil en áreas como la economía, la climatología o la administración. En este informe también compartiré algunas de las dificultades que enfrenté al momento de aprender a usar el software estadístico, ya que al principio fue un reto interpretar los resultados correctamente.



## MI REFLEXIÓN SOBRE EL PARCIAL

Durante el tercer parcial de la materia de Estadística Inferencial viví una de las etapas más complicadas del curso. Desde el inicio de la unidad, me sentí bastante confundido con los nuevos temas que empezamos a ver. A diferencia de los contenidos anteriores, en esta ocasión me costó mucho más trabajo comprender conceptos como el análisis de varianza (ANOVA), la regresión lineal, la correlación y las series de tiempo. Aunque el maestro lo explicó, reconozco que no le puse el interés ni la atención que realmente necesitaban. Dejé pasar oportunidades valiosas para preguntar, participar o practicar, y eso me pasó factura al momento de realizar las prácticas.

Uno de los errores más importantes que reconozco fue no haber participado en clase. Me mantenía en silencio, incluso cuando tenía dudas, por inseguridad o simplemente por no sentirme preparado. Esa falta de participación me desconectó del ritmo de aprendizaje del grupo, y mientras los demás avanzaban, yo sentía que me iba quedando atrás.

También me costó mucho trabajo exponer o pasar al frente a explicar algún tema. Me sentía nervioso, con poca confianza, y prefería evitarlo en lugar de afrontarlo. Pero ahora sé que esa actitud me impidió reforzar mi comprensión y ganar seguridad en los temas.

Otro punto que afectó bastante mi desempeño fue no entregar las prácticas. Eso me costó a la hora de que el profesor me dio mi calificación pues

Sali con un 7. Desde el inicio de la unidad, fui postergando el investigar por mi propia cuenta los temas que



íbamos viendo, para que, al momento de verlo en clases, yo ya tuviera un conocimiento previo sobre los temas.

Al igual que tuve muchos problemas para trabajar con el software SPSS. No sabía bien cómo usarlo, cómo ingresar los datos ni cómo interpretar los resultados. Me frustraba fácilmente cuando algo no salía como esperaba, y eso me desmotivaba aún más y lo que hacía era no hacer nada y mejor pedírselo a unos de mis amigos, pero si estoy consciente que a la hora de hacer el examen no sabía qué hacer. Al final, no realizar y entregar las prácticas fue un reflejo de esa falta de constancia y organización, algo que reconozco.

También me faltó acercarme al maestro a pedir ayuda. Tuve varias oportunidades para hacerlo, pero por pena o por no saber bien cómo formular mis dudas, simplemente lo dejaba pasar.

Ahora que lo reflexiono, me doy cuenta de que preguntar no es una muestra de debilidad, sino de interés. Si me hubiera tomado el tiempo de aclarar mis dudas a tiempo, probablemente el parcial habría sido menos estresante y habría podido responder con mayor seguridad.

Durante el examen me sentí muy inseguro. Hubo ejercicios en los que no sabía por dónde empezar, especialmente aquellos que requerían el uso del software o la interpretación de resultados estadísticos. Sentí que no estaba preparado y que todo me rebasaba. Fue una sensación frustrante, pero al mismo

}  
tiempo, me hizo darme cuenta de que no se trata solo de pasar un examen, sino  
de realmente entender lo que estoy aprendiendo.



A pesar de todo esto, creo que este parcial me dejó una lección muy valiosa. Me hizo ver que necesito cambiar mis hábitos, participar más en clase, practicar con constancia y no dejar pasar las dudas sin resolver.

También entendí que el interés y la actitud que uno le ponga al aprendizaje marcan una gran diferencia. Si quiero mejorar y terminar bien el curso, tengo que tomar decisiones diferentes, ser más disciplinado y aprovechar todas las herramientas y apoyos que tengo a mi alcance.

Este momento, aunque difícil, me ayudó a reflexionar sobre mis errores y sobre lo que puedo hacer para no repetirlos. Todavía tengo la oportunidad de corregir el rumbo y cerrar el semestre de una mejor manera, y estoy dispuesto a hacerlo.

En este periodo, me adentré en temas como ANOVA, correlación, regresión lineal y series de tiempo, los cuales me brindaron la oportunidad de ampliar mi perspectiva sobre el análisis de datos y su relevancia en diferentes contextos prácticos.

A medida que fuimos viendo los temas si pude entender y comprender cómo identificar relaciones entre variables, comparar grupos a través del análisis de varianza y como los modelos que facilitan la predicción basada en información previa.

}  
El tema de series de tiempo me permitió entender cómo los fenómenos cambian y evolucionan a lo largo del tiempo, y cómo esta información puede ser fundamental en campos como la economía, la climatología y la administración.



Ahora iré explicando con mis propias palabras lo que he llegado a comprender, entender, aprender y de igual forma explicare las partes que aún me quedan dudas, de los temas de la unidad 3 **Anova y Correlación, Regresión lineal.**

### **I. Análisis de Varianza (ANOVA)**

Básicamente, ANOVA es una herramienta que se usa cuando queremos saber si hay alguna diferencia real entre las medias de tres o más grupos. Por ejemplo, si queremos comparar el rendimiento académico de estudiantes de tres diferentes escuelas, ANOVA nos ayuda a saber si esas diferencias en las calificaciones son significativas o si simplemente ocurrieron por casualidad.

Antes de conocer ANOVA, yo pensaba que la única forma de comparar grupos era usando la famosa prueba t. Pero esa solo sirve para **comparar dos grupos**. Si tienes más de dos, hacer muchas pruebas t puede dar resultados confusos. ANOVA, en cambio, **compara todos los grupos al mismo tiempo**, lo que lo hace más práctico y confiable.

Lo que hace esta técnica es comparar la **variación entre los grupos** con la **variación dentro de cada grupo**. Si los grupos son muy diferentes entre sí (es decir, si hay mucha variabilidad entre ellos) y esa diferencia es mayor que la variabilidad dentro de los mismos grupos, entonces se puede decir que hay una

**diferencia significativa.**

}

Eso sí, ANOVA no se puede aplicar a lo loco. Hay algunas condiciones que se deben cumplir para que funcione bien:



- **Independencia:** los datos de los grupos no deben influirse entre sí. Lo ideal es que los datos se obtengan de manera aleatoria.
- **Homocedasticidad:** suena complicado, pero solo significa que todos los grupos deben tener una **varianza similar**, o sea, que los datos se dispersen más o menos igual.
- **Normalidad:** los residuos (o errores) deben tener una **distribución normal**.
- **Continuidad:** la variable que estás analizando tiene que ser **continua**, no categórica.

Una vez que entendí todo esto, vi que ANOVA es una herramienta muy útil, sobre todo cuando necesitas comparar más de dos grupos sin meterte en líos con muchas pruebas separadas. Lo mejor es que se puede aplicar en muchos contextos: educación, salud, psicología, negocios... en fin, donde sea que haya datos y quieras entender si hay diferencias importantes.

En general, pienso que ANOVA es una herramienta muy útil para comparar varios grupos y entender mejor los datos, y me siento más seguro ahora para usarla y analizar los resultados que arroja.

Cuando empecé a ver el tema de **ANOVA** en clase, al principio me costó un poco entender que no se trata solo de una técnica, sino que en realidad **hay varios tipos de ANOVA**, dependiendo de cuántas variables estés analizando. Por eso, aquí te explico de manera sencilla cuáles son esos tipos y para qué sirven, sin

tanto enredo.

}

## **1. ANOVA de un solo factor (One-way ANOVA)**



Este es el más sencillo. Aquí solo se analiza una sola **variable independiente**.

Por ejemplo, si quiero comparar cómo afecta el tipo de desayuno (ligero, regular o completo) en el rendimiento académico, ANOVA me permite saber si hay una diferencia real entre esos grupos. Lo que hace es comparar las medias y decirnos si esa variable (el desayuno) influye o no. Es ideal cuando solo tienes un factor que cambia entre grupos.

## 2. ANOVA de dos factores (Two-way ANOVA)

Este tipo me pareció un poco más completo. En lugar de analizar una sola variable independiente, aquí se trabajan **dos factores al mismo tiempo**.

Siguiendo con el ejemplo anterior, además del desayuno, también podríamos ver si el **turno de clases** (mañana o tarde) tiene algún efecto.

Lo interesante de este ANOVA es que también analiza si hay una **interacción entre los dos factores**, es decir, si juntos tienen un impacto diferente al que tendrían por separado.

## 3. MANOVA (Análisis multivariante de la varianza)

Este ya es más avanzado. Lo que entendí es que aquí no solo se analiza una variable dependiente, sino **varias** al mismo tiempo. Por ejemplo, si se quiere ver cómo el tipo de entrenamiento y la alimentación afectan tanto el peso **como** la energía de una persona, MANOVA te ayuda a estudiar todo eso junto. Es útil cuando no te interesa solo una medida de resultado, sino varias que están

relacionadas entre sí.

}



Después de estudiar estos tres tipos, me quedó más claro cuándo usar cada uno. Si solo tienes un factor, vas con el ANOVA de un solo factor. Si quieres ver dos factores y su posible interacción, usas el de dos vías. Y si tienes varias cosas que estás midiendo como resultados, entonces lo mejor es aplicar MANOVA.

Ahora entiendo mejor cómo se pueden aplicar estos análisis en la vida real, y aunque al principio me sentí confundido, con ejemplos todo empezó a tener más sentido.

### ¿Qué son los datos bivariados?

Me di cuenta de que sirven para analizar relaciones entre dos cosas al mismo tiempo. No es como cuando solo ves una variable (por ejemplo, la estatura de varias personas), sino que aquí se trata de **dos variables que están relacionadas entre sí**.

Lo que entendí es que los datos bivariados se usan para estudiar cómo **una variable se comporta con respecto a otra**. Por ejemplo, si analizamos el tiempo que una persona estudia y la calificación que obtiene en un examen, ahí estamos trabajando con datos bivariados. La idea es ver si existe alguna relación entre esas dos cosas.

Una de las formas más comunes de visualizar este tipo de datos es con **gráficas de dispersión**, que son esos diagramas donde se ponen puntos en un

}  
plano, y cada punto representa un par de valores (uno para cada variable). Al ver la forma de esos puntos, puedes darte cuenta si hay una tendencia o no.

Con los datos bivariados también se puede calcular la **correlación**, que es una medida que indica qué tan fuerte es la relación entre las dos variables. Si la correlación



es positiva, significa que cuando una sube, la otra también tiende a subir. Si es negativa, pasa lo contrario: cuando una sube, la otra baja.

Otra herramienta importante que se relaciona con los datos bivariados es la **regresión lineal**, que básicamente sirve para hacer una línea que se ajuste a los datos y nos permita **predecir** el valor de una variable con base en la otra.

## II. Correlación

En palabras simples, es una forma de saber si dos cosas están relacionadas. Por ejemplo, si cuando estudio más saco mejor calificación, entonces hay una relación entre las horas de estudio y mis resultados en los exámenes.

La correlación se usa para medir qué tan fuerte es esa relación entre dos variables. Si los cambios en una variable van acompañados de cambios en la otra, entonces hay correlación. Y si no, pues no. nos ayuda a **ver patrones** entre dos variables. No significa que una causa a la otra (eso es otra historia), pero sí nos da pistas sobre cómo se comportan juntas.

Tipos de correlación que aprendí:

- Correlación positiva o directa: Esto pasa cuando una variable sube y la otra también. Por ejemplo, entre más tiempo pasas haciendo ejercicio, más calorías quemas. Ambas van en el mismo sentido.
- Correlación negativa o inversa: Aquí es al revés. Si una variable

}  
sube, la otra baja. Por ejemplo, mientras más horas duermes,  
menos cansado te sientes. Una sube, la otra baja.



- Sin correlación (correlación nula): A veces, no hay relación para nada entre las dos cosas que estás midiendo. Como si intentaras relacionar la cantidad de zapatos que tienes con tu promedio escolar... ahí no hay conexión.

Cuando empecé a estudiar correlación hubieron dos conceptos que al principio me confundieron un poco fueron el **coeficiente de Pearson (r)** y el **coeficiente de determinación ( $r^2$ )**..

¿Qué es el coeficiente de Pearson (r)? Este valor, conocido como "r", nos dice qué tan fuerte y en qué dirección está la relación entre dos variables. Va de -1 a 1, y dependiendo del número, se interpreta así:

Si r es cercano a 1, hay una correlación positiva fuerte, es decir, cuando una variable sube, la otra también.

Si r es cercano a -1, hay una correlación negativa fuerte, o sea, cuando una sube, la otra baja.

Si r es cercano a 0, significa que no hay relación clara entre las variables.

Por ejemplo, si tienes un  $r = 0.85$  entre horas de estudio y calificación, eso indica una relación positiva bastante fuerte: a más estudio, mejor calificación.

¿Qué es el coeficiente de determinación ( $r^2$ )? Este es simplemente r al cuadrado. Se llama " $r^2$ " y se interpreta como el porcentaje de variación de una variable que se puede explicar por la otra.

}  
Siguiendo el mismo ejemplo de estudio y calificación:

Si  $r = 0.85$ , entonces  $r^2 = 0.7225$ , o sea, el 72.25% del cambio en la calificación puede



explicarse por las horas de estudio. Lo que queda (el otro 27.75%) puede depender de otras cosas: el estado de ánimo, el tipo de examen, si comiste antes, etc.

Tanto  $r$  como  $r^2$  te ayudan a entender relaciones entre datos. Pearson te dice si la relación es positiva o negativa, y qué tan fuerte es. Y  $r^2$  te dice qué tanto influye una variable sobre la otra, lo cual es clave cuando quieres hacer predicciones.

### III. Regresión Lineal

Básicamente, la **regresión lineal** es una herramienta que se usa para ver cómo se relacionan dos o más variables. Por ejemplo, si quieres saber si el tiempo que estudias afecta tu calificación, puedes usar regresión lineal para **encontrar una fórmula que describa esa relación.**

#### Regresión lineal simple

Aquí solo se usan **dos variables**: una que creemos que influye (la variable independiente, como las horas de estudio) y otra que es el resultado que queremos analizar (la variable dependiente, como la calificación).

La idea es **trazar una línea recta que se acerque lo más posible a los puntos** que representan los datos. Esa línea tiene una fórmula como esta:

$$y = \beta_0 + \beta_1 x$$

Donde:

}

**y** es lo que queremos predecir (como la calificación),

**x** es la variable que creemos que influye (como las horas de estudio),



$\beta_0$  es el punto donde la línea cruza el eje Y (como un valor base),  
y  $\beta_1$  es cuánto cambia y cuando x cambia (o sea, la pendiente de la recta).

Ya para realizar la línea empleamos el método de mínimos cuadrados.

Suena complicado, pero lo que hace es buscar la línea que cometa el menor error posible al predecir los datos. Ese "error" es simplemente la diferencia entre los valores reales y los que predice la línea.

El método lo que hace es sumar esos errores al cuadrado y buscar que esa suma sea lo más pequeña posible. Eso nos ayuda a la hora de obtener la línea esta representa de forma bastante precisa la tendencia de los datos, y nos ayuda a hacer predicciones con cierta confianza.

Aunque el método de mínimos cuadrados es muy útil para encontrar la mejor línea o curva que ajuste los datos, también tiene sus limitaciones. Es importante entender estas limitaciones para no sacar conclusiones equivocadas de los resultados que obtenemos.

Por ejemplo, en la regresión lineal el ajuste se representa con una línea recta porque se asume que la relación entre las variables es lineal. Pero si la relación es más complicada y no sigue una línea recta, se usan otros tipos de regresión, como la regresión no lineal, donde la mejor forma de ajuste puede ser una curva en lugar de una línea recta.

Comprendí que no siempre una línea recta explica bien los datos, y que hay

}  
que elegir el modelo correcto para interpretar bien lo que está pasando.



Al principio me parecían solo letras en una fórmula, aun no me quedo al 100 por ciento el tema, pero puedo comentar que el coeficiente  $b_0$  (que también se conoce como la intersección o término independiente) representa el valor que tendría la variable dependiente (Y) cuando la variable independiente (X) vale cero.

Es como el punto de partida de la recta en el eje Y. Por ejemplo, si estamos analizando cómo las horas de estudio influyen en una calificación,  $b_0$  sería la calificación esperada sin estudiar nada. Puede sonar raro, pero matemáticamente es útil para construir el modelo.

Luego está  $b_1$ , que es el coeficiente de pendiente, y para mí este es el que más sentido práctico tiene.  $b_1$  nos dice cuánto cambia Y por cada unidad que cambia X. Siguiendo el ejemplo de estudio y calificación: si  $b_1$  vale 2, significa que, por cada hora extra de estudio, la calificación subiría 2 puntos (en promedio). Este valor indica si la relación es positiva o negativa, y qué tan fuerte es ese cambio.

Así que al final, lo que aprendí es que  $b_0$  y  $b_1$  no solo son números en una fórmula, sino que nos ayudan a entender cómo se comportan los datos, y eso hace que todo el análisis tenga sentido en la vida real.

Cuando vimos el tema de la prueba de hipótesis de significatividad en la regresión, entendí que es una parte clave para saber si la ecuación que obtuvimos

}  
realmente sirve para explicar algo o si solo fue casualidad. O sea, no basta con tener una recta con ciertos valores de  $b_0$  y  $b_1$ , también hay que comprobar si esos valores tienen sentido estadísticamente.



Esta prueba básicamente nos ayuda a responder una pregunta: ¿la variable independiente realmente influye en la variable dependiente o no? Para eso se plantea una hipótesis nula ( $H_0$ ) que dice que el coeficiente  $b_1$  es igual a cero, es decir, que no hay relación entre las variables. Y la hipótesis alternativa ( $H_1$ ) dice que sí hay una relación, o sea, que  $b_1$  es diferente de cero.

Después se hace el análisis usando una prueba t. Si el valor p que obtenemos es menor que el nivel de significancia (por ejemplo, 0.05), entonces rechazamos la hipótesis nula, lo que significa que sí hay evidencia de que la variable X tiene un efecto sobre Y, y por lo tanto, la ecuación de regresión es significativa.

Puedo notar que no solo se trata de trazar una línea y ya, sino de verificar si esa línea tiene sentido y es confiable para hacer predicciones o tomar decisiones con base en los datos. al principio me costó un poco entender para qué servía. Pero ya que lo fui viendo con ejemplos, me quedó más claro que es una forma de tener una idea más completa y realista de los valores de esos coeficientes.

Lo que pasa es que cuando hacemos una regresión, el software o los cálculos nos dan un solo valor para  $b_0$  y  $b_1$ , pero ese valor es solo una estimación puntual, basada en los datos de la muestra. Entonces, en vez de quedarnos solo con ese número, lo que se hace es calcular un intervalo de confianza, que nos dice un rango donde probablemente se encuentra el valor

real del coeficiente en la población. }

Por ejemplo, si para  $b_1$  (la pendiente) me sale un intervalo de [1.5, 2.3], eso significa que estoy, por ejemplo, un 95% seguro de que el valor real de la pendiente está



entre 1.5 y 2.3. Es como decir: "no sé el valor exacto, pero con estos datos, estoy bastante seguro de que está dentro de este rango" .

comprendí que en estadística nunca hay certezas absolutas, pero sí herramientas que nos permiten tomar decisiones con cierto nivel de confianza. Y los intervalos son justo eso: una forma de expresar cuánta seguridad tenemos sobre lo que estamos estimando.

#### IV. SERIES DE TIEMPO

Nunca me había puesto a pensar que muchos datos que usamos en la vida diaria tienen una relación directa con el paso del tiempo. Por ejemplo, las ventas que hace una tienda cada mes, el clima diario, o incluso la cantidad de veces que se reproduce una canción en una plataforma cada semana... todo eso puede analizarse con series de tiempo.

**No solo se trata de mirar el pasado**, sino también de usar esa información para intentar **predecir qué podría pasar más adelante**. Y para eso se analizan ciertos patrones que aparecen en los datos. Y para eso hay cuatro cosas clave que debemos observar cuando trabajamos con este tipo de datos:

Primero está la **tendencia**, que es básicamente si los valores van subiendo, bajando o se mantienen más o menos estables con el tiempo.



Luego está lo **estacional**, que son cambios que se repiten en ciertos periodos, como cuando suben las ventas en diciembre o bajan las visitas a una página los fines de semana.

También está la **parte cíclica**, que son como patrones más largos, que pueden tardar años en repetirse. Como en economía, donde hay épocas buenas y otras no tanto. Por último, están las **variaciones aleatorias**, que son cosas inesperadas: una tormenta, una falla técnica, una noticia que cambia todo de un día para otro.

Para que todos esos datos no se vean tan dispersos y podamos entender mejor lo que está pasando, usamos técnicas que nos ayudan a "suavizar" la información. Una de ellas es el **promedio móvil**, que lo que hace es tomar ciertos puntos de la serie, sacar el promedio y moverlo para que se vea una línea más clara. Otra opción es el **alisamiento exponencial**, que le da más importancia a los datos más recientes para resaltar mejor los cambios actuales.

Otra técnica que aprendí para trabajar con series de tiempo es el **alisamiento exponencial**, y la verdad es que me pareció muy interesante por cómo funciona. Su objetivo es el mismo que el del promedio móvil: **suavizar los datos** para que se pueda ver con más claridad alguna tendencia o patrón. Pero lo que lo hace diferente es la manera en que **le da importancia a los datos**.

A diferencia del promedio móvil, que trata a todos los datos por igual (es

}  
decir, les da el mismo peso al hacer el promedio), el alisamiento exponencial **le da más peso a los datos más recientes**. Esto tiene mucho sentido si lo pensamos bien, porque en



}  
muchos casos los datos más actuales son los que tienen más influencia en lo que puede pasar después.

Por ejemplo, si estamos analizando las ventas de los últimos meses, el alisamiento exponencial va a tomar en cuenta todos los datos anteriores, pero le va a dar **más relevancia a lo que ocurrió en los últimos días o semanas**. Y eso ayuda a hacer pronósticos más ajustados a la realidad actual.



## CONCLUSIÓN

Después de haber estudiado todos estos temas en la materia de Estadística Inferencial, puedo decir que mi perspectiva sobre el análisis de datos ha cambiado bastante. Al principio me costó entender muchas cosas, pero poco a poco fui comprendiendo cómo herramientas como ANOVA, la correlación, la regresión lineal y las series de tiempo nos ayudan a interpretar mejor la información y tomar decisiones basadas en datos reales.

Cada tema tuvo su grado de dificultad, sobre todo al momento de usar el software SPSS, que me causó varios problemas al principio. También reconozco que me faltó acercarme al maestro para aclarar mis dudas y participar más en clase. Eso influyó en que me sintiera inseguro al exponer o al realizar ciertas prácticas. Sin embargo, con el tiempo entendí que equivocarse también es parte del aprendizaje.

Ahora sé que el análisis de varianza nos sirve para comparar grupos, que la correlación nos muestra la relación entre dos variables, que con la regresión lineal podemos hacer predicciones, y que las series de tiempo nos permiten entender cómo cambian las cosas a lo largo del tiempo. Todo esto es muy valioso, no solo para esta materia, sino para aplicarlo en distintas áreas de la vida real.

En conclusión, aunque este parcial fue un reto para mí, también fue una oportunidad para reflexionar sobre lo que necesito mejorar como estudiante. Me queda claro que debo involucrarme más, preguntar cuando no entienda algo y

}  
practicar más con las herramientas que usamos. Sé que todavía tengo mucho  
por aprender, pero también sé que voy por el camino correcto.



**UNIVERSIDAD JUÁREZ  
AUTÓNOMA DE TABASCO**  
**DIVICIÓN ACADÉMICA DE CIENCIAS  
ECONÓMICO ADMINISTRATIVA**



**"MAPAS MENTALES"**

**ALUMNO:** Luis Enrique Muñoz Paulino

**MATRICULA:** 241B38018

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:** Estadística Inferencial

**DOCENTE:** Cesar Enrique Barrera Ovando

VILLAHERMOSA, TABASCO, 30 DE MAYO DEL 2025

# ANOVA

## CONCEPTO

Es una técnica estadística utilizada para comparar las medias de tres o más grupos.

## TIPOS DE ANOVA

- **ANOVA de un factor;** Compara las medias de varios grupos basándose en un solo factor.
- **ANOVA de dos factores:** Evalúa el efecto de dos factores simultáneamente y sus interacciones.

## HIPOTESIS DE ANOVA

- **Hipótesis nula:** Establece que no hay diferencias significativas entre las medias de los grupos.
- **Hipótesis alternativa:** Establece que hay al menos una diferencia significativa entre las medias de los grupos.

- **Nivel de significancia:** Es el valor utilizado para determinar si se rechaza o no la hipótesis nula.

## SUPUESTOS DEL ANOVA

## ANÁLISIS DE RESULTADOS DE ANOVA

# $f(x)$

VALOR F

el valor f en anova es el resultado de dividir la varianza entre grupos por la varianza dentro de los grupos

**Puntuaciones de significancia**

indica la probabilidad de obtener un valor F igual o mayor observando si la hipótesis nula fuera verdadera.

**INTERPRETACIÓN DE LOS RESULTADOS**

Si el valor F es mayor que el nivel de

significancia se rechaza la hipótesis nula y se concluye que hay diferencias significativas entre las medias de los grupos

- **Normalidad:** Las distribuciones de los grupos deben ser aproximadamente normales.
- **Homogeneidad de varianzas:** Las varianzas de los grupos deben ser similares.
- **Independencia:** Las observaciones deben ser independientes entre sí.

## ANALISIS DE VARIANZA



# CORRELACIÓN Y REGRESIÓN LINEAL

## DEFINICIÓN DE CORRELACIÓN

LA CORRELACIÓN ES UNA MEDIDA ESTADÍSTICA QUE INDICA LA RELACIÓN ENTRE DOS VARIABLES

## DEFINICIÓN DE REGRESIÓN LINEAL

LA REGRESIÓN LINEAL ES UN MÉTODO ESTADÍSTICO QUE PERMITE PREDECIR EL VALOR DE UNA VARIABLE A PARTIR DE OTRA

## TIPOS DE CORRELACIÓN

1. **CORRELACIÓN POSITIVA:** INDICA QUE A MEDIDA QUE UNA VARIABLE AUMENTA LA OTRA TAMBIÉN LO HACE.
2. **CORRELACIÓN NEGATIVA:** INDICA QUE A MEDIDA QUE UNA VARIABLE AUMENTA LA OTRA DISMINUYE
3. **CORRELACIÓN NULA:** INDICA QUE NO HAY RELACIÓN ENTRE LAS VARIABLES

## EL MODELO DE REGRESIÓN LINEAL

EL MODELO DE REGRESIÓN LINEAL SE BASA EN UNA LÍNEA RECTA QUE REPRESENTA LA RELACIÓN ENTRE LAS VARIABLES

## COEFICIENTE DE CORRELACIÓN

EL COEFICIENTE DE CORRELACIÓN ES UN VALOR NUMÉRICO QUE MIDE LA FUERZA Y DIRECCIÓN DE LA RELACIÓN ENTRE DOS VARIABLES

## COEFICIENTE DE DETERMINACIÓN

EL COEFICIENTE DE DETERMINACIÓN ES UNA MEDIDA QUE INDICA QUE TAN BIEN SE AJUSTA EL MODELO DE REGRESIÓN LINEAL A LOS DATOS OBSERVADOS

# Métodos de Alisamiento Promedios Móviles

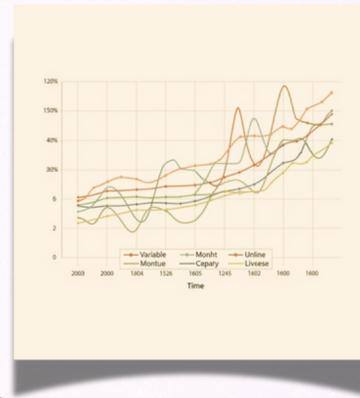
## y Alisamiento Exponencial

son técnicas utilizadas para eliminar el ruido o las fluctuaciones aleatorias en una serie de tiempo, revelando la tendencia subyacente y los patrones a largo plazo.

## Promedios Móviles

calcula la media de un número fijo de observaciones consecutivas para "alisar" la serie original..

- Promedio Móvil Simple
- Promedio Móvil Ponderado
- Promedio Móvil Exponencial



# SERIES DE TIEMPO

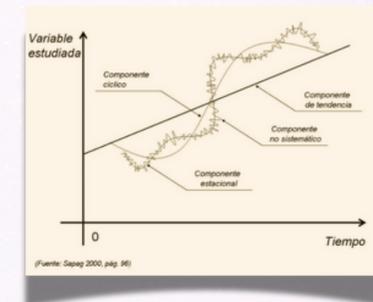
## ¿QUE SON LAS SERIES DE TIEMPO?

Las series de tiempo (o series temporales) son secuencias de datos observados en intervalos de tiempo y ordenados cronológicamente.

## COMPONENTES de UNA series

### de tiempo

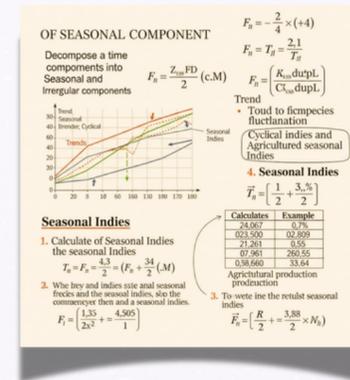
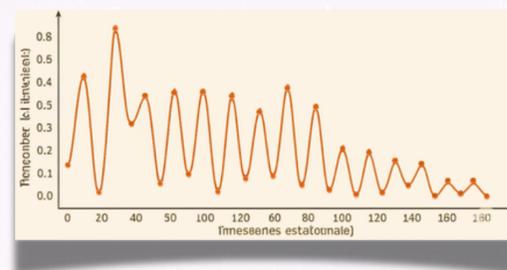
- Tendencia Variación
- estacional Variación
- cíclica Variación
- aleatoria



## Alisamiento Exponencial

utilizan un promedio ponderado de las observaciones pasadas, donde los pesos disminuyen exponencialmente con la antigüedad de la observación.

- Alisamiento Exponencial Simple
- Alisamiento Exponencial Doble
- Alisamiento Exponencial Triple



## Análisis de L componente estacional

Se refiere a la identificación y cuantificación de los patrones de comportamiento que repiten en intervalos fijos y predecibles,



**UNIVERSIDAD JUÁREZ  
AUTÓNOMA DE TABASCO**  
**DIVICIÓN ACADÉMICA DE CIENCIAS  
ECONÓMICO ADMINISTRATIVA**



**“DOCUMENTO EXTENSO”**

**ALUMNO:** Luis Enrique Muñoz Paulino

**MATRICULA:** 241B38018

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:** Estadística Inferencial

**DOCENTE:** Cesar Enrique Barrera Ovando

VILLAHERMOSA, TABASCO, 30 DE MAYO DEL 2025



## ANOVA y CORRELACIÓN, REGRESIÓN LINEAL

El análisis de la varianza, también llamado ANOVA (del inglés Analysis of Variance), es una técnica que permite comparar las varianzas entre las medias de diferentes muestras.

El análisis de la varianza (ANOVA) sirve para analizar si existe alguna diferencia entre las medias de más de dos poblaciones. De manera que el análisis de la varianza permite determinar si las medias poblacionales de dos o más grupos son diferentes analizando la variabilidad entre las medias de las muestras.

Para poder hacer un análisis de la varianza (ANOVA) se deben cumplir los siguientes requisitos:

**Independencia:** los valores observados son independientes entre sí. Una manera de garantizar la independencia de las observaciones es añadiendo aleatoriedad en el proceso de muestreo.

**Homocedasticidad:** debe haber homogeneidad en las varianzas, es decir, la variabilidad de los residuos es constante.

**Normalidad:** los residuos deben estar distribuidos normalmente, o dicho de otra forma, deben seguir una distribución normal.

**Continuidad:** la variable dependiente debe ser continua. por H1.

Existen tres tipos de análisis de la varianza (ANOVA):

1. Análisis de la varianza de un factor (One-way ANOVA): en el análisis de la varianza solo hay un factor, es decir, solo hay una variable independiente.



2. Análisis de la varianza de dos factores (Two-way ANOVA): el análisis de la varianza tiene dos factores, por lo tanto, se analizan dos variables independientes y la interacción entre ellas.
3. Análisis multivariante de la varianza (MANOVA): en el análisis de la varianza hay más de una variable dependiente. Se pretende determinar si las variables independientes modifican su valor al variar las variables dependientes.

### **Ventajas y desventajas del análisis de la varianza (ANOVA)**

La principal ventaja del análisis de la varianza (ANOVA) es que permite comparar más de dos grupos a la vez. A diferencia de la prueba t, en la que solo se puede analizar la media de una o dos muestras, el análisis de la varianza sirve para determinar si múltiples poblaciones tienen la misma media o no.

No obstante, el análisis de la varianza no nos dice cuál es el grupo del estudio que tiene una media distinta, solamente nos permite averiguar si hay alguna media significativamente diferente o si todas las medias son similares.

Asimismo, otra desventaja del análisis de la varianza es que se deben cumplir cuatro supuestos previos (vistos arriba) para poder realizar el análisis ANOVA, de lo contrario,

### **Aplicaciones del ANOVA de un solo factor.**

El análisis de varianza (ANOVA) de un factor nos sirve para comparar

varios grupos en una variable cuantitativa. Esta prueba es una generalización del contraste de igualdad de medias para dos muestras independientes.



Se aplica para contrastar la igualdad de medias de tres o más poblaciones independientes y con distribución normal. Supuestas  $k$  poblaciones independientes, las hipótesis del contraste son siguientes:

1.  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  Las medias poblacionales son iguales
2.  $H_1$ : Al menos dos medias poblacionales son distintas

Para realizar el contraste ANOVA, se requieren  $k$  muestras independientes de la variable de interés. Una variable de agrupación denominada Factor  $y$  clasifica las observaciones de la variable en las distintas muestras.

Suponiendo que la hipótesis nula es cierta, el estadístico utilizado en el análisis de varianza sigue una distribución  $F$  de Fisher-Snedecor con  $k-1$  y  $n-k$  grados de libertad, siendo  $k$  el número de muestras y  $n$  el número total de observaciones que participan en el estudio.

## CORRELACIÓN

¿Qué es la correlación? La correlación es una medida estadística que indica el grado de relación entre dos variables. En concreto, la correlación lineal sirve para determinar cuánto de correlacionadas linealmente están dos variables distintas. Dos variables están relacionadas cuando al variar los valores de una variable también cambian los valores de la otra variable. Por ejemplo, si al aumentar la variable  $A$  también aumenta la variable  $B$ , existe una correlación entre las variables  $A$  y  $B$ .

Según cómo sea la relación que hay entre dos variables aleatorias, se distinguen los siguientes tipos de correlación lineal:



- Correlación directa (o correlación positiva): una variable aumenta cuando la otra también aumenta.
- Correlación inversa (o correlación negativa): cuando una variable aumenta la otra disminuye, y al revés, si una variable disminuye la otra aumenta.
- Correlación nula (sin correlación): no existe ninguna relación entre las dos variables.

### **Coefficiente de correlación de Pearson**

El coeficiente de correlación de Pearson, también llamado coeficiente de correlación lineal o simplemente coeficiente de correlación, es una medida estadística que indica la relación entre dos variables. Para calcular el coeficiente de correlación de Pearson entre dos variables se debe dividir la covarianza de dichas variables por la raíz cuadrada del producto de sus varianzas.

De manera que el coeficiente de correlación de Pearson trata de cuantificar la dependencia lineal entre dos variables aleatorias cuantitativas. A priori, valorar numéricamente la correlación entre dos variables es complicado porque resulta difícil determinar qué pareja de variables está más correlacionada entre sí, así pues, el objetivo de coeficiente de correlación de Pearson es dar un valor a la relación entre variables para luego poder comparar entre ellas. El valor del índice de correlación de Pearson está entre -1 y +1, ambos incluidos.

La regresión lineal es un modelo estadístico que relaciona una o varias variables independientes con una variable dependiente. Es decir, la regresión lineal es una técnica



que sirve para encontrar una ecuación que aproxime la relación entre una o varias variables explicativas y una variable respuesta.

Por ejemplo, la ecuación  $y=2+5x_1-3x_2+8x_3$  es un modelo de regresión lineal, ya que relaciona matemáticamente tres variables independientes ( $x_1$ ,  $x_2$ ,  $x_3$ ) con una variable dependiente ( $y$ ) y, además, la relación entre las variables es lineal.

### **Regresión lineal simple**

La regresión lineal simple se usa para relacionar una variable independiente  $X$  con una variable dependiente  $Y$ . Es decir, en una regresión lineal simple solo hay dos variables (la variable explicativa  $X$  y la variable respuesta  $Y$ ) y se intenta aproximar la relación que hay entre ambas variables.

La ecuación de un modelo de regresión lineal simple es una recta, por lo que está formada por dos coeficientes: la constante de la ecuación ( $\beta_0$ ) y el coeficiente de la correlación entre las dos variables ( $\beta_1$ ). Por lo tanto, la ecuación de un modelo de regresión lineal simple es  $y=\beta_0+\beta_1x$ .

### **método de mínimos cuadrados**

Es una técnica estadística fundamental en el análisis de datos que busca encontrar la mejor línea o curva de ajuste para un conjunto de datos dispersos, minimizando la suma de los errores al cuadrado entre los valores observados y estimados. Esta metodología se utiliza ampliamente en diversas disciplinas,

incluyendo la economía, la física y las ciencias sociales, debido a su capacidad para ofrecer proyecciones precisas y análisis significativos.



La fórmula de mínimos cuadrados es uno de los elementos más importantes para entender cómo este método opera. Su simplicidad permite que sea accesible a investigadores y expertos de varias disciplinas. A nivel práctico, el método puede ser implementado de manera efectiva a través de una serie de pasos claros, que incluyen la recolección de datos, la determinación de la pendiente y la intersección de la línea o curva de ajuste, y la validación del modelo resultante.

Sin embargo, a pesar de su utilidad, este método también presenta limitaciones que deben ser comprendidas para garantizar la correcta interpretación de los resultados obtenidos. Este ajuste se puede representar como una línea recta en el caso de la regresión lineal o como una curva para la regresión no lineal. Los mínimos cuadrados fórmula para una regresión lineal simple es:

$$y = a + bx$$

donde  $y$  es la variable dependiente,  $x$  es la variable independiente,  $a$  es la intersección y  $b$  es la pendiente de la línea de ajuste. El valor de  $b$  se calcula para minimizar la suma de las diferencias al cuadrado entre los valores observados y los valores estimados.

### **Cómo interpretar la pendiente y el intercepto**

En particular, debemos tener cuidado al hablar de los dos componentes

clave: la pendiente y el intercepto.

El intercepto te indica dónde cruza la recta de regresión el eje y. En términos prácticos, representa el valor de la variable dependiente cuando la variable independiente es cero. Es importante saber que el intercepto no siempre es interpretable.



La pendiente indica cuánto se espera que cambie la variable dependiente con un aumento de una unidad en la variable independiente. Una pendiente positiva sugiere una relación positiva, en la que la variable dependiente aumenta a medida que aumenta la variable independiente. Una pendiente negativa indica lo contrario.

### Prueba de la significancia de la regresión

La prueba de significancia de la regresión es una de las pruebas de hipótesis utilizadas para medir la bondad de ajuste del modelo. Esta prueba determina si existe

una relación lineal entre la variable respuesta  $Y$  y alguna de las variables regresoras  $X_1, X_2, \dots, X_k$ . La hipótesis estadística adecuada es:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \beta_j \neq 0 \text{ para al menos un } j; j = 1, 2, \dots, k$$

Al rechazar la hipótesis nula se concluye que al menos una de las variables regresoras contribuye significativamente al modelo. La prueba estadística utilizada es

$$\begin{aligned} F &= \frac{CM_{\text{Regresión}}}{CM_{\text{Error}}} \\ &= \frac{SC(\text{Regresión}/b_0)/(p-1)}{CM_{\text{Error}}} \\ &= \frac{(\mathbf{b}'\mathbf{X}'\mathbf{Y} - n\bar{Y}^2)/(p-1)}{CM_{\text{Error}}} \end{aligned}$$

La cual asumiendo que la hipótesis nula es cierta se distribuye  $F$  con  $k$  grados de libertad en el numerador y  $n - k - 1$  grados de libertad en el denominador.

Se rechaza la hipótesis nula si el valor calculado de la estadística de prueba es mayor que el valor teórico de la distribución  $F(\alpha, k, n - k - 1)$



## La estimación por intervalo para los coeficientes

en una regresión lineal simple nos permite construir rangos de valores dentro de los cuales es muy probable que se encuentren los verdaderos valores de estos parámetros, con un cierto nivel de confianza (por ejemplo, 95%).

Cuando se realiza una regresión lineal simple, el modelo que se ajusta es de la forma:

$$y = \beta_0 + \beta_1 x + \varepsilon, y = \hat{\beta}_0 + \hat{\beta}_1 x + \varepsilon,$$

donde  $\varepsilon$  es un término de error aleatorio. A partir de los datos, obtenemos estimaciones de los coeficientes verdaderos:  $\hat{\beta}_0$  y  $\hat{\beta}_1$ . Sin embargo, como estos valores se obtienen a partir de una muestra y no de toda la población, existe cierta incertidumbre. Por eso se construyen intervalos de confianza, que nos indican el rango dentro del cual es probable que se encuentre el valor real del coeficiente.

Para construir estos intervalos, primero se calculan los errores estándar de los coeficientes, que miden la variabilidad de las estimaciones. Estos errores estándar dependen del tamaño de la muestra, de la dispersión de los valores de  $x$ , y del error residual del modelo.

Luego, se elige un nivel de confianza, comúnmente del 95%, lo cual implica usar un valor crítico de la distribución  $t$  de Student con  $n - 2$  grados de libertad (donde  $n$  es el número de datos). Este valor se multiplica por el error

estándar del coeficiente para obtener el margen de error.



Finalmente, el intervalo de confianza se construye con la fórmula. Esto da como resultado dos rangos de valores: uno para el intercepto y otro para la pendiente.

La interpretación es la siguiente: si, por ejemplo, el intervalo de confianza para  $\beta_1$  al 95% es de [1.2, 2.8], eso significa que, con un 95% de confianza, el valor real de la pendiente se encuentra dentro de ese intervalo. Además, si este intervalo no incluye el cero, eso indica que hay una relación significativa entre las variables  $x$  e  $y$ .

## SERIES DE TIEMPO

En estadística, las series de tiempo (o series temporales) son secuencias de datos observados en intervalos de tiempo y ordenados cronológicamente. Normalmente las series de tiempo se representan en gráficos para analizar la evolución temporal de los datos.

Las series de tiempo sirven para analizar un conjunto de datos y, además, ayudan a hacer predicciones pues se pueden utilizar técnicas estadísticas para extrapolar la información e intentar predecir un valor futuro. Para analizar una serie de datos se deben tener en cuenta los siguientes componentes:

- o **Tendencia:** se trata de la evolución a largo plazo de la serie de tiempo. Por ejemplo, la serie de tiempo del apartado de arriba tiene una tendencia creciente, ya que el número de ventas va subiendo a

medida que avanza el tiempo.

- o Variación estacional: son movimientos a corto plazo que se repiten de manera periódica. Por ejemplo, el número de visitas de una página web informativa como esta disminuye durante los fines de semana.



- o Variación cíclica: componente de la serie que recoge las oscilaciones periódicas que tienen una amplitud superior a un año. Por ejemplo, las empresas que venden helados aumentan considerablemente su número de ventas cada verano.
- o Variación aleatoria: variaciones que se deben debido a fenómenos puntuales como tormentas, inundaciones, huelgas, guerras, avances tecnológicos, etc.

### **Métodos de alisamiento promedios móviles y alisamiento exponencial**

Los **métodos de alisamiento** son técnicas estadísticas utilizadas para reducir la variabilidad aleatoria en los datos de una serie temporal, con el fin de identificar con mayor claridad comportamientos como tendencias o ciclos. Dos de los métodos más conocidos son el **promedio móvil** y el **alisamiento exponencial**.

El método de **promedios móviles** consiste en calcular el promedio de un número fijo de observaciones consecutivas. Este promedio se va desplazando (o “moviendo” ) a lo largo de la serie temporal, de ahí su nombre.

Por ejemplo, un **promedio móvil de 3 periodos** calcula el promedio de los primeros 3 datos, luego se mueve un paso adelante y calcula el promedio del segundo al cuarto dato, y así sucesivamente. La fórmula general es:

$$PM_t = \frac{y_t + y_{t-1} + \dots + y_{t-k+1}}{k},$$

donde:

- $PM_t$  es el promedio móvil en el periodo  $t$ ,
- $k$  es el número de periodos que se están promediando.



Este método es útil cuando se desea suavizar datos con fluctuaciones pequeñas y detectar la tendencia general, pero tiene una desventaja: puede perder información en los extremos de la serie y no responde bien a cambios repentinos.

### **Alisamiento Exponencial**

El alisamiento exponencial también tiene como objetivo suavizar los datos, pero lo hace asignando mayor peso a las observaciones más recientes. A diferencia del promedio móvil, que da igual peso a todos los datos considerados, el alisamiento exponencial pondera cada dato con un factor que disminuye exponencialmente a medida que los datos son más antiguos. La fórmula básica del alisamiento exponencial simple es:

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1},$$

donde:

- $S_t$  es el valor suavizado en el tiempo  $t$ ,
- $y_t$  es el valor real observado en el tiempo  $t$ ,
- $\alpha$  es el coeficiente de alisamiento (entre 0 y 1),
- $S_{t-1}$  es el valor suavizado anterior.

Este método es más flexible que los promedios móviles y puede ajustarse para pronósticos, especialmente cuando no hay una tendencia o estacionalidad fuerte.

### **El análisis de la tendencia por regresión lineal simple**

Es una técnica estadística que permite modelar y entender la dirección general o el comportamiento a largo plazo de una serie temporal. Esta técnica es especialmente útil cuando los datos presentan una tendencia creciente o decreciente a lo largo del tiempo, ya que busca ajustar una recta que represente esa tendencia.



Se utiliza la regresión lineal simple, cuyo modelo general es:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

donde:

- $y_t$  es el valor de la variable en el tiempo  $t$ ,
- $t$  es el tiempo (por ejemplo, años, trimestres, meses),
- $\beta_0$  es el intercepto (valor estimado cuando  $t = 0$ ),
- $\beta_1$  es la pendiente o coeficiente de tendencia (indica el cambio promedio de  $y$  por cada unidad de tiempo),
- $\varepsilon_t$  es el término de error (lo que no explica la recta).

Ventajas del análisis de tendencia por regresión:

- Permite hacer proyecciones a futuro si la tendencia es estable.
- Facilita la comprensión del comportamiento global de los datos.
- Es fácil de aplicar y de interpretar.

### **Análisis de la componente e índices estacionales.**

El análisis de la componente estacional busca identificar y medir los patrones que se repiten de manera regular en determinados periodos de tiempo, como meses, trimestres o estaciones del año. Este comportamiento se conoce como estacionalidad y es común en datos de ventas, producción, turismo, consumo de energía, entre otros.

Por ejemplo, las ventas de juguetes tienden a aumentar en diciembre, mientras que las de uniformes escolares crecen antes del inicio del ciclo escolar. Estos cambios no se deben al azar, sino a factores estacionales que influyen cada año de manera similar.

¿Qué es la componente estacional? La componente estacional es una de las partes en que se puede descomponer una serie temporal, junto con la tendencia, el ciclo y la



variación aleatoria. Representa el efecto sistemático y recurrente que se repite dentro de cada ciclo temporal (como los 12 meses del año).

¿Qué son los índices estacionales? Los índices estacionales son valores que cuantifican el efecto de la estacionalidad en cada periodo del ciclo (meses, trimestres, etc.). Estos índices permiten comparar cada valor observado con el comportamiento promedio del conjunto.

- Un índice estacional mayor que 1 (o mayor que 100%) indica que ese periodo tiene un efecto positivo sobre la variable (está por encima del promedio).
- Un índice menor que 1 (o menor que 100%) indica que el periodo tiene un efecto negativo (está por debajo del promedio).
- Un índice igual a 1 (o 100%) indica que ese periodo está en el promedio general.

### **Cálculo de índices estacionales (método clásico):**

1. Eliminar la tendencia: Si los datos tienen una tendencia clara, esta debe eliminarse para aislar la estacionalidad. Esto se puede hacer dividiendo (modelo multiplicativo) o restando (modelo aditivo) el valor observado por el valor de tendencia estimado.
2. Agrupar los datos por periodo estacional: Por ejemplo, agrupar todos los eneros, febreros, etc., si la estacionalidad es mensual.
3. Calcular el promedio de cada periodo: Se obtiene el valor

promedio de cada mes o trimestre.



#### 4. Obtener los índices estacionales:

- En el modelo multiplicativo, se calcula:

$$\text{Índice estacional} = \frac{\text{valor de periodo}}{\text{promedio general}}$$

- En el modelo aditivo :

$$\text{Índice estacional} = \text{valor de periodo} - \text{Promedio general}$$

5. Ajustar los índices: En el modelo multiplicativo, la suma de los índices debe igualar el número de periodos (por ejemplo, 12 si son meses); en el aditivo, deben sumar cero.

#### ¿Para qué sirven los índices estacionales?

- o Para ajustar series temporales eliminando el efecto estacional.
- o Para hacer pronósticos más precisos, especialmente cuando se combinan con modelos de tendencia.
- o Para comparar datos de distintos periodos sin que la estacionalidad afecte la interpretación.



## Referencias

*ANOVA de un factor.* (s. f.).

<https://www.jmp.com/es/statistics-knowledge-portal/one-way-anova>

Estadística, P. Y. (2023a, marzo 1). *Análisis de la varianza (ANOVA)*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/analisis-de-la-varianza-anova/>

Estadística, P. Y. (2023b, marzo 8). *Regresión lineal simple*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal-simple/>

Estadística, P. Y. (2023c, marzo 10). *Regresión lineal*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal/>

Estadística, P. Y. (2023d, marzo 10). *Regresión lineal*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal/>

Estadística, P. Y. (2023e, noviembre 11). *Correlación*. Probabilidad y Estadística.

<https://www.probabilidadyestadistica.net/correlacion/>

Estadística, P. Y. (2024a, enero 2). *Promedio móvil*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/promedio-movil/>

Estadística, P. Y. (2024b, enero 2). *Serie de tiempo*. Probabilidad y

Estadística. <https://www.probabilidadyestadistica.net/series-de-tiempo/>

Estadística, P. Y. (2024c, enero 5). *Coefficiente de correlación de Pearson*.

Probabilidad y Estadística.

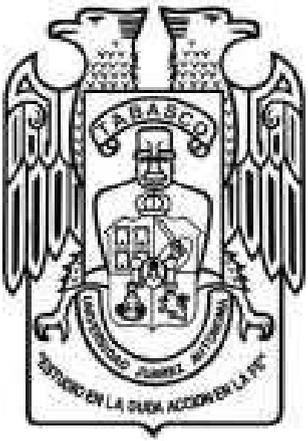
<https://www.probabilidadyestadistica.net/coeficiente-de-correlacion-de-pearson/>

Mate. (2025, 12 abril). Método de mínimos cuadrados: definición y aplicaciones clave.

Matematix. <https://matematix.org/formula-de-minimos-cuadrados/>

*Promedios móviles y métodos de suavizamiento.* (s. f.). calameo.com.

<https://www.calameo.com/books/000764714c5ae035f87f1>



**UNIVERSIDAD JUÁREZ  
AUTÓNOMA DE TABASCO**  
**DIVICIÓN ACADÉMICA DE CIENCIAS  
ECONÓMICO ADMINISTRATIVA**



**"PRACTICA NO.3"**

**ALUMNO:** Luis Enrique Muñoz Paulino

**MATRICULA:** 241B38018

**LICENCIATURA EN CONTADURIA PÚBLICA**

**ASIGNATURA:** Estadística Inferencial

**DOCENTE:** Cesar Enrique Barrera Ovando

VILLAHERMOSA, TABASCO, 30 DE MAYO DEL 2025



## INTRODUCCIÓN

El presente documento tiene como objetivo servir como evidencia de los temas abordados en las unidades 3 y 4 de la materia de Estadística Inferencial, centrada en "ANOVA, Correlación, Regresión Lineal y Series de Tiempo". Para ello, se llevó a cabo una práctica en el centro de cómputo de la universidad, la cual debía resolverse utilizando el software IBM SPSS Statistics.

El documento incluye una recopilación de los ejercicios realizados con variables de un conjunto de datos, así como capturas de pantalla que muestran los resultados obtenidos, incluyendo tablas y gráficos generados por el software. Además, se presenta un informe que contiene la interpretación de los resultados obtenidos durante el análisis estadístico.

SPSS Statistics es un software de análisis de datos que podemos utilizar para llevar a cabo tareas estadísticas y de investigación. Con él, podemos gestionar y analizar grandes volúmenes de datos, realizar análisis descriptivos, crear gráficos y aplicar modelos estadísticos complejos.

En el caso de esta unidad, el software fue fundamental para realizar análisis de varianza (ANOVA), identificar relaciones entre variables a través de la correlación, construir modelos predictivos mediante regresión lineal, y examinar patrones temporales con análisis de series de tiempo.

Además, su interfaz gráfica facilita enormemente la realización e

}  
interpretación de estos procedimientos sin necesidad de conocimientos  
avanzados en programación.



## 1. INGRESO MENSUAL SEGÚN EL ESTADO CIVIL

ANOVA						
Ingreso mensual	Suma de cuadrados	gl	Media cuadrática	F	Sig.	
Entre grupos	507413636,2	3	169137878,7	30,779	,000	
Dentro de grupos	4819398509	877	5495323,271			
Total	5326812145	880				

Pruebas post hoc						
Comparaciones múltiples						
Variable dependiente: Ingreso mensual						
Bonferroni						
(i) Estado civil	(j) Estado civil	Diferencia de medias (i-j)	Desv. Error	Sig.	Limite inferior	Limite superior
casado	divorciado	-1046,228 <sup>a</sup>	265,535	,001	-1748,37	-344,09
	soltero	-2207,869 <sup>a</sup>	257,509	,000	-2888,79	-1526,95
	viudo	-1963,018 <sup>a</sup>	255,310	,000	-2638,12	-1287,91
divorciado	casado	1046,228 <sup>a</sup>	265,535	,001	344,09	1748,37
	soltero	-1161,640 <sup>a</sup>	214,447	,000	-1728,69	-594,59
	viudo	-916,790 <sup>a</sup>	211,802	,000	-1476,85	-356,73
soltero	casado	2207,869 <sup>a</sup>	257,509	,000	1526,95	2888,79
	divorciado	1161,640 <sup>a</sup>	214,447	,000	594,59	1728,69
	viudo	244,851	201,648	1,000	-288,36	778,06
viudo	casado	1963,018 <sup>a</sup>	255,310	,000	1287,91	2638,12
	divorciado	916,790 <sup>a</sup>	211,802	,000	356,73	1476,85
	soltero	-244,851	201,648	1,000	-778,06	288,36

\*. La diferencia de medias es significativa en el nivel .05.

### Introducción de la variable:

La variable que estoy analizando se llama "Ingreso mensual" y se refiere a la cantidad de dinero que una persona recibe cada mes, ya sea por su trabajo u otras fuentes de ingreso. Esta variable es cuantitativa y continua, y me permite observar el nivel económico de las personas dentro del conjunto de datos. En este caso, utilicé el ingreso mensual como variable dependiente para analizar si existen diferencias significativas según el estado civil de las personas,

### Hipótesis:

**Hipótesis nula ( $H_0$ ):** No existen diferencias significativas en el ingreso mensual promedio entre los diferentes grupos de estado civil.

**Hipótesis alternativa ( $H_1$ ):** Al menos uno de los grupos de estado civil tiene un

}  
ingreso mensual promedio diferente.



### **Interpretación:**

Realicé un análisis ANOVA para comparar si el ingreso mensual varía según el estado civil de las personas (casado, soltero, divorciado y viudo). El resultado mostró un valor de significancia de .000, lo cual es menor que .05, por lo tanto, hay evidencia suficiente para afirmar que sí existen diferencias significativas en los ingresos entre al menos dos de estos grupos.

- casados y solteros: tienen una diferencia promedio de -2207.869, con una significancia de .000,
- También hay diferencias significativas entre divorciados y casados, así como entre solteros y viudos, entre otros.T

Estas comparaciones que tienen un valor de significancia (Sig.) menor a .05 muestran que realmente existen diferencias entre esos grupos.

### **Conclusión de la variable de estudio**

al analizar la variable "Ingreso mensual" , pude comprobar que el estado civil tiene un impacto significativo en los ingresos que perciben las personas. A través del análisis ANOVA y la prueba post hoc, se identificaron diferencias claras entre varios grupos, especialmente entre personas casadas, solteras y divorciadas. Esto me permite concluir que el ingreso mensual no es igual para todos, y que el estado civil es un factor que puede influir en la situación económica de los individuos.



## 2. INGRESO MENSUAL SEGÚN EL GÉNERO

ANOVA						
Ingreso mensual						
	Suma de cuadrados	gl	Media cuadrática	F	Sig.	
Entre grupos	507413636,2	3	169137878,7	30,779	.000	
Dentro de grupos	4819398509	877	5495323,271			
Total	5326812145	880				

Pruebas post hoc						
Comparaciones múltiples						
Variable dependiente: Ingreso mensual						
Bonferroni						
(I) Estado civil	(J) Estado civil	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
casado	divorciado	-1046,228 <sup>*</sup>	265,535	,001	-1748,37	-344,09
	soltero	-2207,869 <sup>*</sup>	257,509	,000	-2888,79	-1526,95
	viudo	-1963,018 <sup>*</sup>	255,310	,000	-2638,12	-1287,91
divorciado	casado	1046,228 <sup>*</sup>	265,535	,001	344,09	1748,37
	soltero	-1161,640 <sup>*</sup>	214,447	,000	-1728,69	-594,59
	viudo	-916,790 <sup>*</sup>	211,802	,000	-1476,85	-356,73
soltero	casado	2207,869 <sup>*</sup>	257,509	,000	1526,95	2888,79
	divorciado	1161,640 <sup>*</sup>	214,447	,000	594,59	1728,69
	viudo	244,851	201,648	1,000	-288,36	778,06
viudo	casado	1963,018 <sup>*</sup>	255,310	,000	1287,91	2638,12
	divorciado	916,790 <sup>*</sup>	211,802	,000	356,73	1476,85
	soltero	-244,851	201,648	1,000	-778,06	288,36

\*. La diferencia de medias es significativa en el nivel .05.

### Introducción de la variable

Al analizar estas variables veremos si existen diferencias significativas en el ingreso mensual entre los distintos géneros. Este análisis no solo busca identificar posibles brechas salariales, sino también relevar y comprender las desigualdades económicas que pueden estar asociadas al género.

Esto es fundamental para promover la igualdad de oportunidades en el ámbito laboral, ya que las diferencias en remuneración pueden reflejar prácticas discriminatorias o desigualdades estructurales que afectan principalmente a grupos vulnerables.

### Hipótesis:

**Hipótesis nula:** El ingreso mensual no difiere significativamente entre los hombres y las mujeres.

**Hipótesis Alternativa:** El ingreso mensual difiere significativamente entre los

hombres y las mujeres.

}



### **Interpretación:**

Se aplicó una prueba t para muestras independientes con el fin de evaluar si existe una diferencia significativa entre los grupos analizados. Cuando el valor p es menor que 0.05, se rechaza la hipótesis nula, lo que indica que hay una diferencia estadísticamente significativa entre las medias de las dos muestras.

Además, se presentan las estadísticas descriptivas correspondientes, tales como la media, la mediana y la desviación estándar, para ofrecer una mejor comprensión de la distribución y variabilidad de los datos en cada grupo.

### **Conclusión de la variable de estudio**

Existen diferencias significativas en el ingreso mensual entre los géneros, dado que el valor p fue menor a 0.05, lo que permite rechazar la hipótesis nula de igualdad de medias. Esto sugiere que el género sí influye en el nivel de ingresos mensuales, evidenciando una posible desigualdad salarial basada en esta variable.

Las estadísticas descriptivas, como la media y la desviación estándar, muestran además cómo varían los ingresos dentro de cada grupo, lo que aporta un panorama más completo sobre la dispersión y tendencia central de los datos.



### 3. TALLA SEGÚN EL NIVEL DE INSTRUCCIÓN

ANOVA						
Talla	Suma de cuadrados	gl	Media cuadrática	F	Sig.	
Entre grupos	.037	3	.012	.988	.398	
Dentro de grupos	10.829	877	.012			
Total	10.866	880				

Pruebas post hoc						
Comparaciones múltiples						
Variable dependiente: Talla						
Bonferroni						
(I) Grado de instrucción	(J) Grado de instrucción	Diferencia de medias (I-J)	Desv. Error	Sig.	Límite inferior	Límite superior
1	2	-.01544	.01981	1.000	-.0678	.0369
	3	-.02419	.01530	.686	-.0647	.0163
	4	-.02456	.01537	.662	-.0652	.0161
2	1	.01544	.01981	1.000	-.0369	.0678
	3	-.00875	.01489	1.000	-.0481	.0306
	4	-.00913	.01496	1.000	-.0487	.0304
3	1	.02419	.01530	.686	-.0163	.0647
	2	.00875	.01489	1.000	-.0306	.0481
	4	-.00038	.00809	1.000	-.0218	.0210
4	1	.02456	.01537	.662	-.0161	.0652
	2	.00913	.01496	1.000	-.0304	.0487
	3	.00038	.00809	1.000	-.0210	.0218

#### Introducción de la variable

El objetivo de este análisis es determinar si existen diferencias significativas en la variable "Talla" según el grado de instrucción. Para ello, se aplicó un análisis de varianza (ANOVA) de un factor, seguido de una prueba post hoc de comparaciones múltiples (Bonferroni), para explorar específicamente entre qué grupos podrían existir diferencias.

#### Hipótesis:

**Hipótesis nula ( $H_0$ ):** No existen diferencias significativas en la talla entre los diferentes grados de instrucción.

**Hipótesis alternativa ( $H_1$ ):** Existen diferencias significativas en la talla entre al menos dos grupos de grados de instrucción.



### **Interpretación:**

El análisis de varianza (ANOVA) se realizó con el fin de identificar si existían diferencias estadísticamente significativas en la variable "Talla" entre los distintos grados de instrucción. Los resultados mostraron un valor de  $F = 0.988$  con un nivel de significancia ( $p$ ) de 0.398. Este valor es mayor al umbral comúnmente aceptado de 0.05, lo que indica que no hay evidencia suficiente para rechazar la hipótesis nula. En otras palabras, no se observaron diferencias significativas en la talla entre los diferentes niveles educativos.

### **Conclusión de la variable de estudio**

Los resultados obtenidos en este análisis permiten concluir que no existen diferencias estadísticamente significativas en la talla entre los diferentes grados de instrucción. El valor de significancia del ANOVA ( $p = 0.398$ ) fue mayor al nivel crítico de 0.05, lo que indica que las variaciones en la altura entre los grupos no son lo suficientemente grandes como para atribuir las al grado de instrucción.

Ya que ninguna de las diferencias entre pares de grupos resultó significativa, y todos los intervalos de confianza incluyeron el valor cero. Esto refuerza la idea de que, en esta muestra, el nivel educativo no tiene un impacto importante sobre la estatura de las personas.



#### 4.

→ Unidireccional

ANOVA

Peso	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Entre grupos	1687.911	2	843.956	2.082	.125
Dentro de grupos	355926.270	878	405.383		
Total	357614.181	880			

Pruebas post hoc

Comparaciones múltiples

Variable dependiente: Peso  
Bonferroni

(I) Tipo de vivienda	(J) Tipo de vivienda	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
1	2	1.86309	1.78612	.892	-2.4211	6.1472
	3	3.91944	1.94942	.134	-.7564	8.5953
2	1	-1.86309	1.78612	.892	-6.1472	2.4211
	3	2.05635	1.56878	.571	-1.7065	5.8192
3	1	-3.91944	1.94942	.134	-8.5953	.7564
	2	-2.05635	1.56878	.571	-5.8192	1.7065

### PESO SEGÚN EL TIPO DE VIVIENDA

#### Introducción de la variable

En este análisis, lo que busco es identificar si hay diferencias significativas en el peso de las personas dependiendo del tipo de vivienda en la que viven. Para eso utilicé un ANOVA unidireccional, que es una prueba estadística que me permite comparar las medias de más de dos grupos en una sola variable. En este caso específico, estoy comparando los promedios de peso entre tres tipos diferentes de vivienda para ver si el lugar donde vive una persona podría estar relacionado con su peso.

**Hipótesis nula ( $H_0$ ):** No existen diferencias significativas en el peso promedio de las personas según el tipo de vivienda. Es decir, el tipo de vivienda no influye en el peso.

**Hipótesis alternativa ( $H_1$ ):** Sí existen diferencias significativas en el peso promedio de las personas dependiendo del tipo de vivienda. Esto implicaría que al menos un grupo difiere de los demás en cuanto al peso.



### **Interpretación:**

Al aplicar el ANOVA unidireccional, obtuve un valor de  $F = 2.082$  con una significancia (p) de 0.125. Este valor es mayor al nivel de significancia comúnmente utilizado de 0.05, por lo tanto, no puedo rechazar la hipótesis nula. Esto significa que no se encontraron diferencias estadísticamente significativas en el peso de las personas según el tipo de vivienda

Los resultados muestran que todas las comparaciones tienen valores de significancia mayores a 0.05, y los intervalos de confianza incluyen el cero, lo cual refuerza que no hay diferencias significativas entre los tipos de vivienda en cuanto al peso.

### **Conclusión de la variable de estudio**

Con base en los resultados obtenidos del ANOVA y la prueba post hoc de Bonferroni, puedo concluir que no existen diferencias estadísticamente significativas en el peso de las personas según el tipo de vivienda en la que residen. El valor de p fue de 0.125, lo cual es mayor al nivel de significancia de 0.05, por lo tanto, no se rechaza la hipótesis nula.

Esto quiere decir que, en esta muestra, el lugar donde vive una persona — ya sea el tipo de vivienda 1, 2 o 3— no tiene un efecto claro o significativo sobre su peso.



## 5. GASTO MENSUAL SEGÚN EL ESTADO CIVIL

ANOVA						
Gasto mensual						
	Suma de cuadrados	gl	Media cuadrática	F	Sig.	
Entre grupos	121995429,0	3	40665143,00	25.919	.000	
Dentro de grupos	1375930672	877	1568906.126			
Total	1497926101	880				

**Pruebas post hoc**

**Comparaciones múltiples**

Variable dependiente: Gasto mensual  
Bonferroni

(I) Estado civil	(J) Estado civil	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
casado	divorciado	-437.787*	141.881	.013	-812.96	-62.62
	soltero	-1070.366*	137.592	.000	-1434.20	-706.54
	viudo	-905.563*	136.417	.000	-1266.29	-544.84
divorciado	casado	437.787*	141.881	.013	62.62	812.96
	soltero	-632.580*	114.583	.000	-935.57	-329.59
	viudo	-467.776*	113.170	.000	-767.03	-168.53
soltero	casado	1070.366*	137.592	.000	706.54	1434.20
	divorciado	632.580*	114.583	.000	329.59	935.57
	viudo	164.804	107.745	.759	-120.10	449.71
viudo	casado	905.563*	136.417	.000	544.84	1266.29
	divorciado	467.776*	113.170	.000	168.53	767.03
	soltero	-164.804	107.745	.759	-449.71	120.10

\*. La diferencia de medias es significativa en el nivel .05.

### Introducción de la variable

En este análisis quise investigar si el estado civil de una persona influye en su gasto mensual. Para ello, utilicé un ANOVA unidireccional, que me permite comparar las medias de gasto entre más de dos grupos. En este caso, analicé las diferencias entre personas casadas, divorciadas, solteras y viudas.

La idea es ver si existen diferencias significativas en el nivel de gasto mensual dependiendo del estado civil. Para complementar el análisis, también apliqué una prueba post hoc con el método Bonferroni, con el objetivo de identificar entre qué grupos específicos se encuentran esas posibles diferencias.



Hipótesis nula ( $H_0$ ): No existen diferencias significativas en el gasto mensual promedio entre los distintos estados civiles. Es decir, el estado civil no influye en cuánto gasta una persona mensualmente.

Hipótesis alternativa ( $H_1$ ): Sí existen diferencias significativas en el gasto mensual promedio según el estado civil. Esto significa que al menos uno de los grupos presenta un gasto mensual distinto a los demás.

### **Interpretación:**

Después de realizar el ANOVA unidireccional, obtuve un valor de  $F = 25.919$  con una significancia de  $p = 0.000$ . Este valor es menor a  $0.05$ , lo que indica que sí existen diferencias estadísticamente significativas en el gasto mensual según el estado civil. Es decir, el estado civil influye en cuánto gasta mensualmente una persona. Los resultados muestran que hay diferencias significativas en varias comparaciones, especialmente entre personas casadas y solteras, casadas y viudas, solteras y divorciadas, entre otras. Todas estas comparaciones tienen un valor de  $p$  menor a  $0.05$ , y sus intervalos de confianza no incluyen el cero, lo que confirma que esas diferencias son significativas.

### **Conclusión de la variable de estudio**

Con base en los resultados del ANOVA y la prueba post hoc de Bonferroni, puedo concluir que el estado civil sí influye de manera significativa en el gasto mensual de las personas. El valor de significancia obtenido ( $p = 0.000$ ) indica que existen diferencias reales entre los grupos analizados. Las comparaciones múltiples mostraron que ~~personas casadas, solteras, viudas y divorciadas tienen patrones de gasto diferentes.~~

}  
Por ejemplo, se observaron diferencias importantes entre casados y solteros, así como entre solteros y viudos. Estas diferencias son estadísticamente significativas, lo que significa que no se deben al azar.



## CONCLUSIÓN

En esta práctica, lo que hice fue aplicar el análisis de varianza, o ANOVA unidireccional, para analizar si ciertas variables categóricas, como el tipo de vivienda o el estado civil, influyen en variables numéricas como el peso, la talla o el gasto mensual. Este tipo de análisis me permitió comparar varios grupos a la vez y ver si había diferencias significativas entre ellos.

En algunos casos, como el peso según el tipo de vivienda o la talla según el grado de instrucción, no encontré diferencias significativas. Eso me mostró que no siempre una variable categórica tiene un efecto en los resultados. Pero en otros casos, como en el gasto mensual según el estado civil, sí hubo diferencias importantes, y eso lo confirmé también con la prueba post hoc de Bonferroni, que me ayudó a ver exactamente entre qué grupos estaban esas diferencias.

En general, esta práctica me sirvió mucho para entender cómo se aplica el ANOVA y cómo interpretar sus resultados, además de ver la importancia de usar pruebas adicionales cuando el ANOVA indica que sí hay diferencias. Fue una herramienta muy útil para analizar datos de forma más profunda y precisa.

## CONCLUSIÓN SOBRE EL PARCIAL

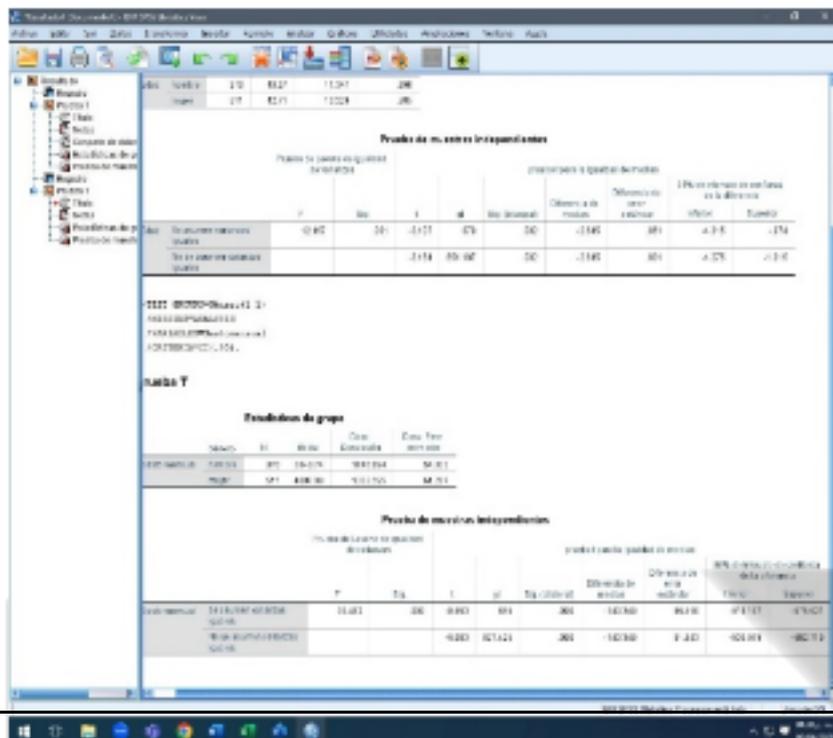
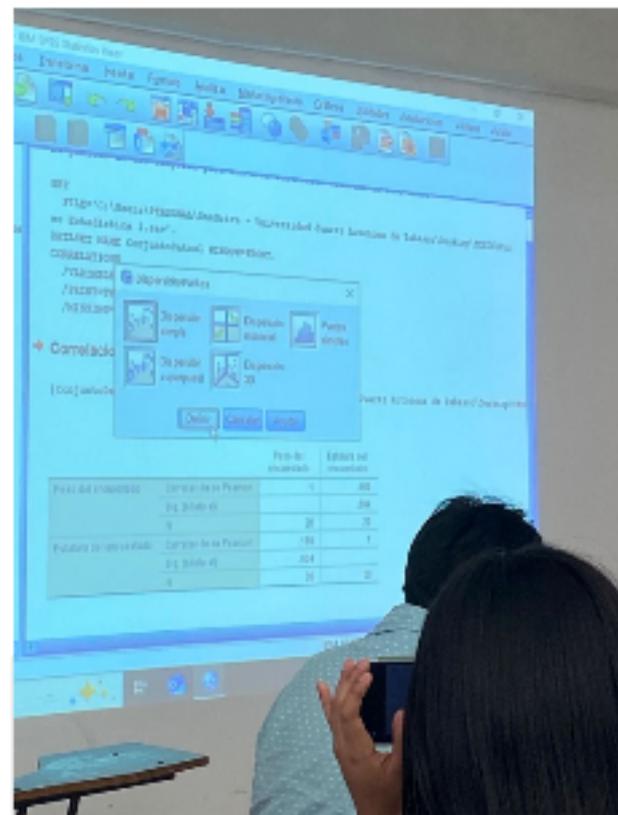
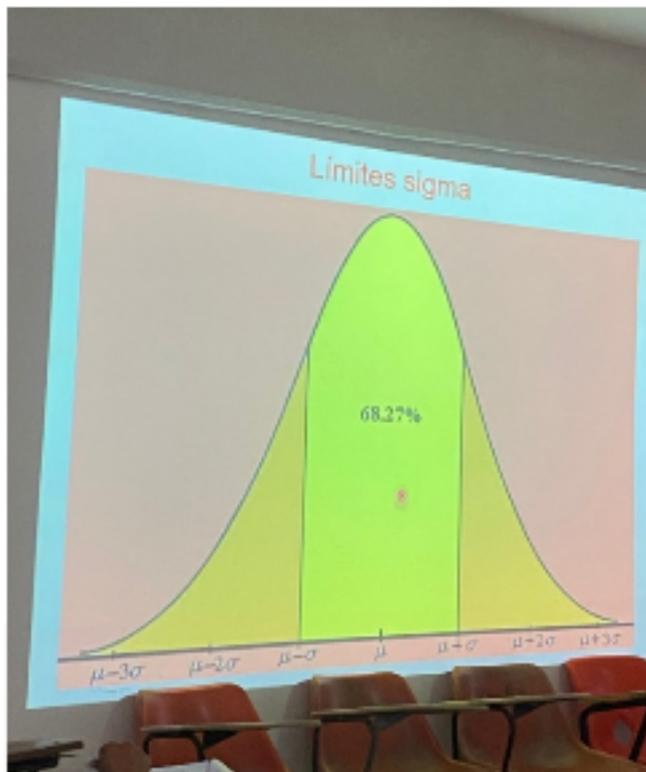
Después de estudiar estas dos unidades, me doy cuenta de lo útiles y potentes que son las herramientas estadísticas cuando se trata de analizar información y tomar decisiones fundamentadas. En la Unidad 3, aprendí a aplicar el análisis de varianza (ANOVA), que me permite comparar varias medias al mismo tiempo para saber si hay diferencias significativas entre grupos. Esto es especialmente útil cuando se quiere evaluar el efecto de diferentes tratamientos o condiciones.

También conocí la correlación, que me ayudó a entender el grado de relación entre dos variables, y la regresión lineal, que va un paso más allá al permitirme construir modelos para predecir una variable a partir de otra. Estas técnicas me resultaron bastante prácticas porque no solo explican relaciones, sino que también ayudan a proyectar comportamientos futuros con base en datos concretos.

Por otro lado, en la Unidad 4 aprendí a trabajar con series de tiempo, es decir, con datos recolectados a lo largo del tiempo. Entendí cómo identificar patrones como la tendencia, la estacionalidad y las variaciones irregulares, y cómo usar métodos de suavizamiento, como los promedios móviles, para hacer predicciones más claras y precisas. Esta parte me hizo ver que muchas

situaciones reales, desde las ventas de una empresa hasta el comportamiento del clima, pueden analizarse y entenderse mejor con estas herramientas

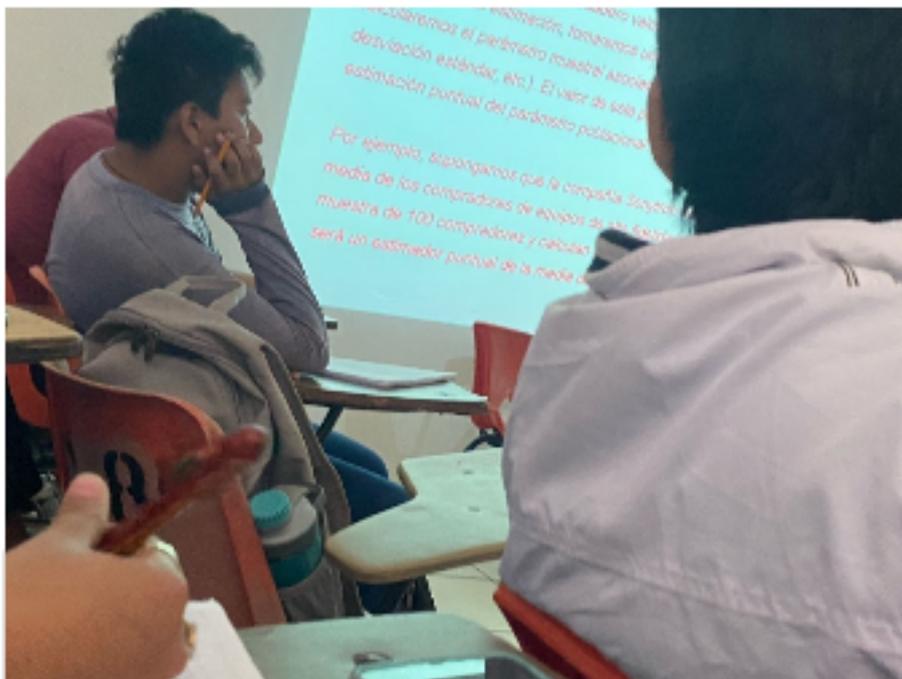
## EVIDENCIAS



**La familia de las distribuciones de probabilidad normal**

- ✓ No existe una sola distribución de probabilidad normal, sino más bien se trata de toda una "familia" de ellas.
- ✓ Cada una de las distribuciones puede tener una media distinta ( $\mu$ ) y desviación estándar distinta ( $\sigma$ ).





## CONCLUSIÓN FINAL

Al llegar al final de este portafolio de evidencias, me tomo

un momento para reflexionar sobre todo lo que he aprendido y cómo he evolucionado en el manejo de herramientas estadísticas. Cada unidad representó un reto distinto, pero también una oportunidad para fortalecer mis conocimientos y habilidades analíticas.

Desde la comprensión de las pruebas de hipótesis y el control de calidad, pasando por el análisis de varianza, la correlación y la regresión, hasta llegar al estudio de series de tiempo, me doy cuenta de que la estadística va mucho más allá de los números: es una forma de interpretar el mundo, de tomar decisiones con fundamento y de entender procesos con mayor profundidad.

Este recorrido no solo me dejó aprendizajes técnicos, sino también una mayor conciencia sobre la importancia de analizar con criterio, de cuestionar los

resultados y de confiar en el poder de los datos cuando se usan de manera correcta. Además, aprendí a valorar el proceso, incluso cuando hubo dificultades, porque cada error o duda se convirtió en una oportunidad de mejora.

En definitiva, este portafolio refleja no solo los temas estudiados, sino también mi compromiso personal con el aprendizaje. Me voy con herramientas útiles, una visión más clara y, sobre todo, con la satisfacción de haber dado un paso más en mi formación académica y profesional.

## Referencias

*ANOVA de un factor.* (s. f.).

<https://www.jmp.com/es/statistics-knowledge-portal/one-way-anova>

Benites, L. (2021, 7 mayo). *Correlación versus regresión: ¿Cuál es la diferencia?*

Statologos. <https://statologos.com/correlacion-vs-regresion>

*¿Cuáles son los componentes de una serie temporal?* (s. f.).

<https://todosloshechos.es/cuales-son-los-componentes-de-una-serie-temporal>

De los Datos, E. M. (2023, 17 noviembre). *Pruebas de Hipótesis: Teoría y ejemplos prácticos*. El Mundo de los Datos.

<https://elmundodelosdatos.com/pruebas-de-hipotesis-teoria-y-ejemplos-practicos/>

Estadística, P. Y. (2023a, marzo 1). *Análisis de la varianza (ANOVA)*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/analisis-de-la-varianza-anova/>

Estadística, P. Y. (2023b, marzo 8). *Regresión lineal simple*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal-simple/>

Estadística, P. Y. (2023c, marzo 10). *Regresión lineal*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal/>

Estadística, P. Y. (2023d, marzo 10). *Regresión lineal*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/regresion-lineal/>

Estadística, P. Y. (2023e, noviembre 11). *Correlación*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/correlacion/>

Estadística, P. Y. (2024a, enero 2). *Promedio móvil*. Probabilidad y Estadística. <https://www.probabilidadyestadistica.net/promedio-movil/>

Estadística, P. Y. (2024b, enero 2). *Series de tiempo*. Probabilidad y Estadística.

<https://www.probabilidadyestadistica.net/series-de-tiempo/>

Estadística, P. Y. (2024c, enero 5). *Coeficiente de correlación de Pearson*.

Probabilidad y Estadística.

<https://www.probabilidadyestadistica.net/coeficiente-de-correlacion-de-pearson/>

Eva. (2021, 12 noviembre). Tema 4b. Correlación y regresión. -

Estadisticaparatodos.com. *Estadisticaparatodos.com*.

<https://estadisticaparatodos.com/correlacion-y-regresion/>

Ferreira, M. (2025, 30 mayo). *¿Qué es el control de calidad estadístico?* Checklist

Fácil | Blog. <https://es.checklistfacil.com/blog/control-calidad-estadistico/>

Kafesu, A. (2024, 14 mayo). *Análisis de la varianza (ANOVA) explicado en menos de 5 minutos*. Geekflare Spain. <https://geekflare.com/es/analysis-of-variance-explained/>

Learn Statistics Easily. (2024, 21 agosto). *Qué es: Estacionalidad - APRENDE ESTADÍSTICAS FÁCILMENTE*. LEARN STATISTICS EASILY.

<https://es.statisticseasily.com/glosario/%C2%BFQu%C3%A9-es-la-estacionalidad-en-el-an%C3%A1lisis-de-datos%3F/>

Mate. (2025, 12 abril). Método de mínimos cuadrados: definición y aplicaciones clave. *Matematix*. <https://matematix.org/formula-de-minimos-cuadrados/>

Ortega, C. (2025, 7 febrero). *Prueba de Hipótesis: Qué es, pasos y ejemplos*. QuestionPro. <https://www.questionpro.com/blog/es/prueba-de-hipotesis/>

*Promedios móviles y métodos de suavizamiento*. (s. f.). calameo.com. <https://www.calameo.com/books/000764714c5ae035f87f1>

Teststrials. (2022, 20 enero). *El Control Estadístico de la Calidad en los procesos de producción*. Tests And Trials.

<https://www.testsandtrials.com/blog/control-estadistico-de-la-calidad/>

*Tutorial sobre la correlación y la regresión*. (s. f.).

<https://www.jmp.com/es/online-statistics-course/correlation-and-regression>

Virtualpro, R. (s. f.). *Virtualpro - Noticia: Historia del control estadístico de la calidad*.

<https://www.virtualpro.co/actualidad/historia-del-control-estadistico-de-la-calidad-56715>

