



Estadística Inferencial

Cuaderno de trabajo

Ing. Joel Herrera Ordoñez

Carrera:

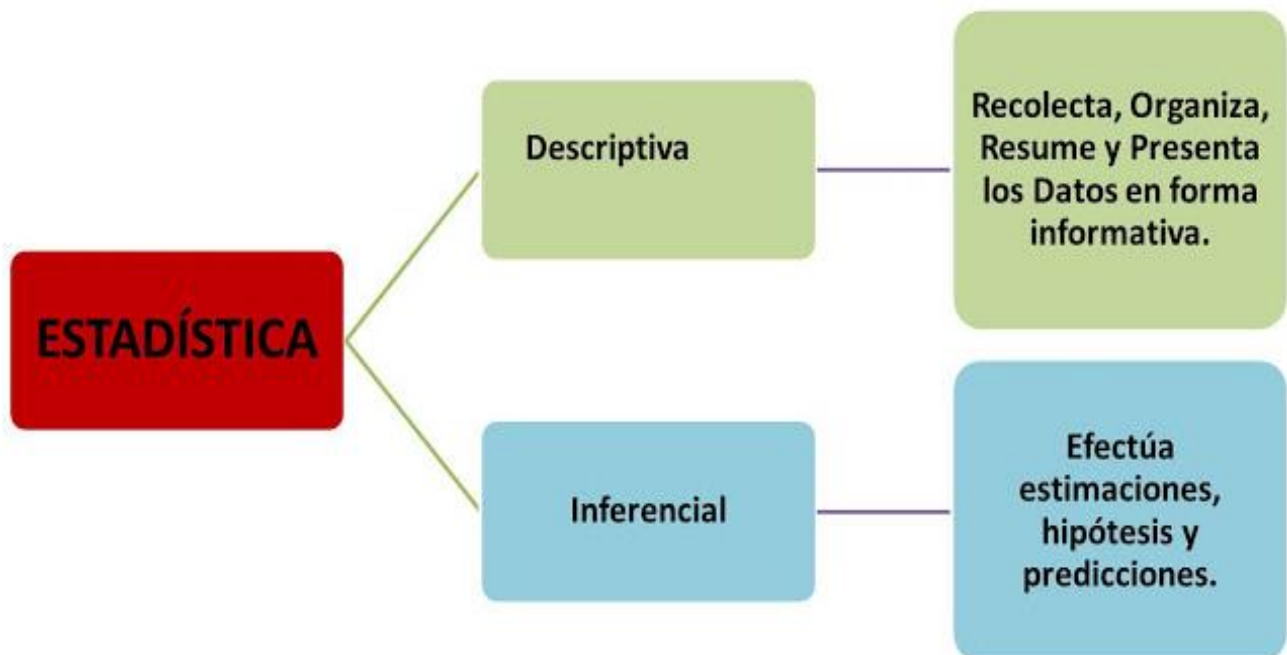
PASIÓN POR EDUCAR

L.A.E, L.C.P, L.PS.

Grupo, Semestre Y Modalidad:

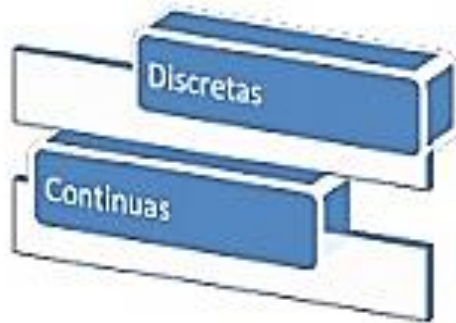
4to. Cuatrimestre "Semiescolarizado"

Concepto de Estadística



VARIABLE ESTADISTICA

Es la característica de los elementos de la población que se investiga.



Aquella variable que no es medible.

Cualitativas

Genero

Color de piel

Nacionalidad

Tipos de Variable

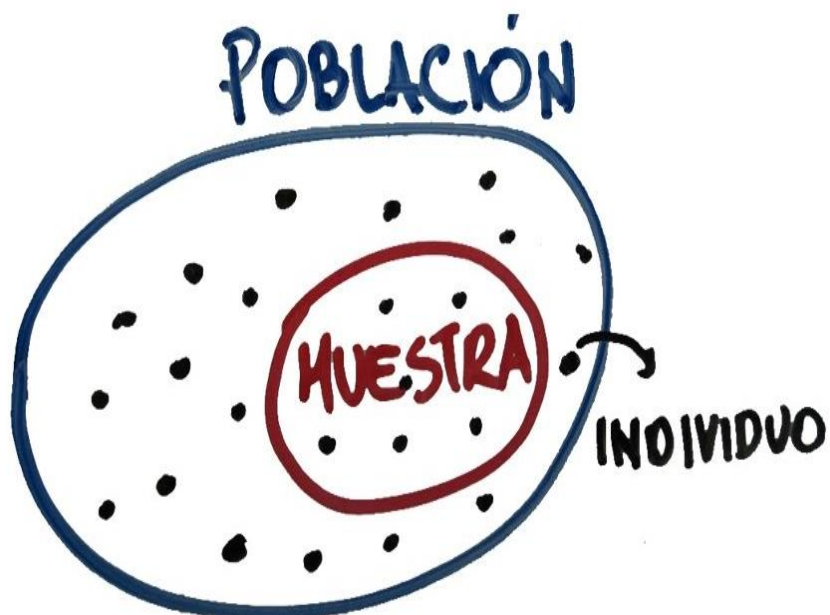
Cuantitativas

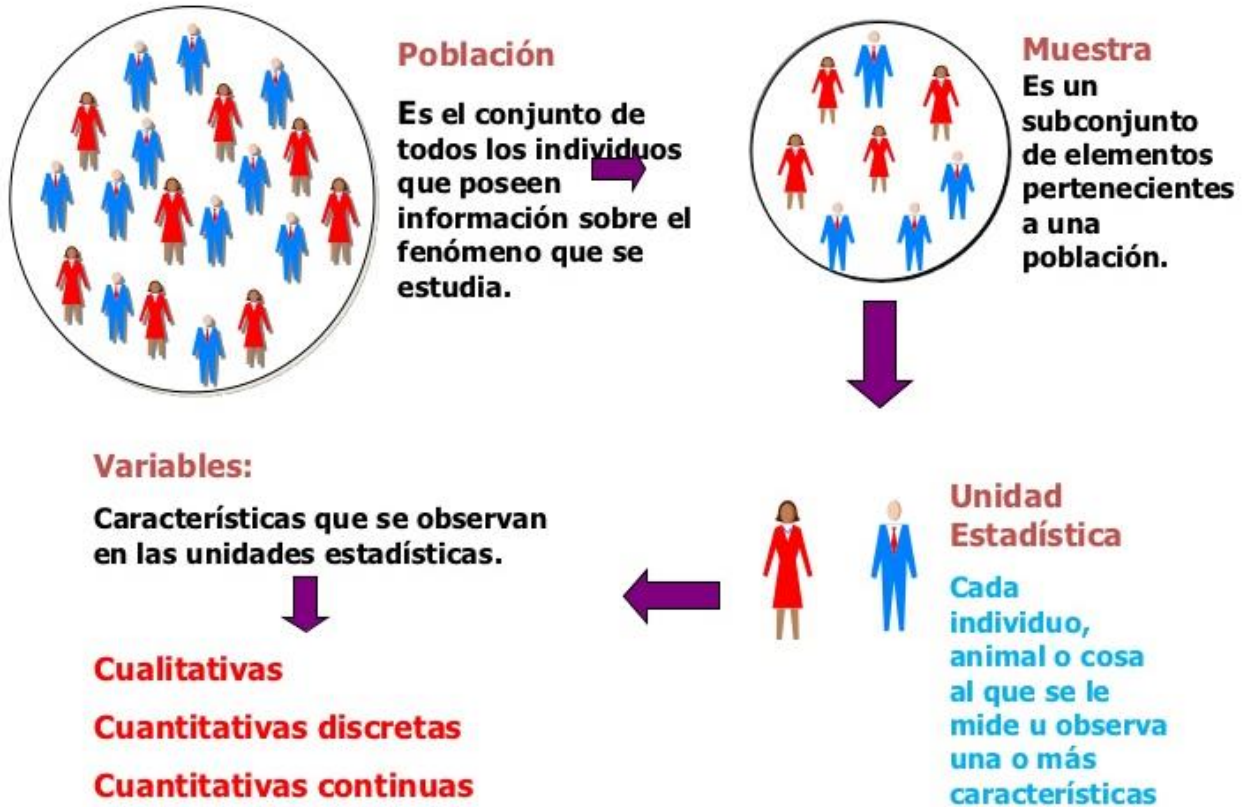
Edad

Estatura

Precios

Aquella variable que se puede contar o medir.





COMPONENTES DE UNA INVESTIGACION ESTADISTICA

1. Formulación del problema.
2. Diseño del experimento
3. Recolección de datos.
4. Tabulación y descripción de los resultados.
5. Conclusiones.

ESTADISTICA INFERENCIAL

EJERCICIOS

En Estadística inferencial se llama estimación al conjunto de técnicas que permiten dar un valor aproximado de un parámetro de una población a partir de los datos proporcionados por una muestra.

1. Estimación puntual

Consiste en la estimación del valor del parámetro mediante un sólo valor.

Ejercicio: Las puntuaciones en una muestra aleatoria de 10 estudiantes a un test fueron las siguientes: 25, 24, 22, 20, 25, 18, 17, 24, 16, 21

Determina la puntuación promedio de los estudiantes (μ)

2. Estimación por intervalos

Consiste en la obtención de un intervalo dentro del cual estará el valor del parámetro estimado con una cierta probabilidad.

2.1 Intervalo de confianza para la media de una población

Ejercicio 1: De 50,000 peones de la construcción que laboran en el Distrito Federal, se tomó una muestra aleatoria de 400 y se investigó su ingreso diario. En la siguiente tabla se muestran los resultados. Construya un intervalo de confianza de 90% para el ingreso diario de la población total de peones.

Ingreso diario (x)	Número de peones (f)
50 – 60	50
60 – 70	80
70 – 80	150
80 – 90	100
90 – 100	20
Total	400

Ejercicio 2: De 2,000 alumnos que asisten a una escuela secundaria, se tomó una muestra de 200 a los que se les pregunto el número de horas que dedican a la semana para practicar algún deporte. Haga una estimación de intervalo con un nivel de confianza de 97% para el número promedio de horas que dedica al deporte la población completa de alumnos.

Número de horas (x)	Número de alumnos (f)
2 – 5	23
5 – 8	43
8 – 11	64
11 – 14	50
14 – 17	20
Total	200

Ejercicio 3: Se realiza un muestreo sobre la calificación que los habitantes de una ciudad en el norte del país le darían a su gobierno; se determinó con una muestra de 400 personas que del 0 al 10 la calificación que le otorgarían es de 5 con una desviación estándar de 0.5. Determine un intervalo de confianza para la calificación que se le da a dicho gobierno con un nivel de confianza del 95% y 99%.

Ejercicio 4: Se tomó una muestra de 100 estudiantes los cuales tienen un gasto promedio en fotocopias cada módulo de \$ 30 pesos, con una desviación estándar de \$ 12 pesos. Determine el intervalo de confianza para la media con un nivel de confianza del 90% y 99%.

2.2 Intervalo de confianza para la diferencia de medias

Ejercicio 1: La altura media de los alumnos de un centro se distribuye según una normal con desviación estándar de 15 cm y la de las alumnas sigue una normal con desviación estándar de 18 cm. Para estimar la diferencia de altura media de los chicos y las chicas se elige una muestra al azar de 40 alumnos y de 35 alumnas. Las alturas medias muestrales son: $X_h = 170$ cm y $X_m = 160$ cm. Hallar el intervalo de confianza para la diferencia de alturas medias al nivel del 90%.

Ejercicio 2: Se sabe que los pesos medios de los caballos de carreras se distribuyen normalmente. Los de la cuadra A con una desviación estándar de 45 kg y los de la cuadra B con una desviación estándar de 51 kg. Se desea estimar la diferencia de pesos medios de los caballos de ambas cuadras; para ello, se elige una muestra de 50 caballos de la cuadra A y 38 caballos de la cuadra B. Se calculan los pesos medios muestrales y se obtiene $X_A = 490$ kg. y $X_B = 475$ kg. Halla el intervalo de confianza para la diferencia de medias de pesos al nivel del 95%.

Ejercicio 3: Haya el intervalo de confianza al nivel del 90% para la diferencia de salarios medios de los trabajadores y trabajadoras de una gran empresa, cuando se ha elegido una muestra de 40 hombres y 35 mujeres, siendo el salario medio de los hombres de \$ 1051 y el de las mujeres \$ 1009.

- Suponiendo que las desviaciones estándar son 90 y 78 respectivamente.
- Suponiendo que las desviaciones estándar son 87 y 76 respectivamente.

2.3 Intervalo de confianza para proporciones

Ejercicio 1: Se hizo una encuesta a 325 personas mayores de 16 años y se encontró que 120 iban al teatro regularmente. Haya con un nivel de confianza del 94% un intervalo para estudiar la proporción de los ciudadanos que van al teatro regularmente.

Ejercicio 2: El ministerio de política social, desea conocer el interés de los padres por la introducción de la lengua extranjera en el primer curso de primaria. Encuestados 1024 padres elegidos al azar de los cuales el 80% está a favor. ¿Cuál es el intervalo de confianza para el porcentaje de padres que están a favor de esta medida con un nivel de confianza del 99%?

Ejercicio 3: Tomada una muestra de 500 personas de una determinada comunidad, se encontró que 300 leían la prensa regularmente. Haya con un nivel de confianza del 90% un intervalo para estimar la proporción de lectores entre las personas de la comunidad.

2.4 Intervalo de confianza para diferencia entre proporciones

Ejercicio 1: Se desea determinar si un cambio en el proceso de fabricación de cierto tipo de piezas ha sido efectivo o no. Para esto se toman dos muestras aleatorias una antes y otra después del cambio. Los resultados obtenidos son:

	Antes	Después
No de piezas defectuosas	75	80
No de piezas analizadas	1500	2000

Construir un intervalo de confianza del 90% para determinar la diferencia entre proporciones y concluya.

Ejercicio 2: Un fabricante afirma que su nuevo producto de consumo popular lo prefieren más los hombres que las mujeres. Para comprobar tal afirmación, se toma una muestra aleatoria de 250 hombres y otra de 200 mujeres, y se encuentra que 175 hombres y 120 mujeres prefieren el nuevo producto. Utilizando un intervalo de confianza del 95% para la diferencia entre proporciones de preferencia entre los hombres y mujeres ¿se puede concluir que el fabricante del nuevo producto tiene razón?

Ejercicio 3: Dos muestras aleatorias de 250 mujeres y 200 hombres indican que 75 mujeres y 80 hombres consumen un nuevo producto unisex que acaba de salir al mercado. Utilizando un intervalo de confianza del 95% ¿se puede aceptar que es igual la proporción de preferencias de mujeres y hombres en toda la población?

2.5 Intervalo de confianza para varianzas

Ejercicio 1: Una muestra de 25 hombres participo en una prueba de resistencia. La desviación estándar calculada a partir de los datos de la muestra es de 1.5. Construya un intervalo de confianza de 95% para la varianza poblacional.

Ejercicio 2: En un estudio de circulación sanguínea se obtuvieron los tiempos de circulación aparente en una muestra de 30 personas con enfermedad arterial. Se encontró que la varianza de la muestra es de 1.03. Construya un intervalo de confianza del 99% para la varianza poblacional.

Ejercicio 3: A un grupo de 12 individuos se le sometió a una dieta especial y al final se les midió el nivel de colesterol. La varianza calculada fue de 0.1527. Suponiendo que la población tiene una distribución normal, construya un Intervalo de confianza del 95% para la varianza poblacional.

Ejercicio 4: A un grupo de individuos se le sometió a una dieta especial y al final se les midió el nivel de colesterol. Los resultados fueron los siguientes: 6.0, 6.4, 7.0, 5.8, 6.0, 5.8, 5.9, 6.7, 6.1, 6.5, 6.3, 5.8 mmol/litro. Suponiendo que la población tiene una distribución normal, construya un Intervalo de confianza del 95% para la varianza poblacional del nivel de colesterol.

Ejercicio 5: Se midieron las concentraciones de hemoglobina en 5 animales expuestos a un compuesto químico nocivo. Se registraron los siguientes valores: 15.6, 14.8, 14.4, 16.6, 13.8. Construya un intervalo de confianza de 95% para la varianza poblacional.

2.6 Intervalo de confianza para razones de dos varianzas

Ejercicio 1: En una delegación política se realizaron encuestas en dos colonias para saber su opinión acerca de la construcción de una obra pública. En la colonia uno, se tomó una muestra de $n_1 = 11$ personas con una desviación estándar de 5.8 y en la colonia dos se tomó una muestra de $n_2 = 4$ personas con una desviación estándar de 3.4. Construya un intervalo de confianza de 95% para la razón de las varianzas de las dos poblaciones.

Ejercicio 2: 32 adultos afásicos sometidos a terapia del habla fueron divididos en dos grupos. El grupo uno consta de 17 personas y recibió el tratamiento 1 y el grupo dos consta de 15 personas y recibió el tratamiento dos. El análisis estadístico de los resultados de la eficacia de los tratamientos dio lo siguiente: Varianza 1 = 15, Varianza 2 = 8. Construya un intervalo de confianza de 95% para la razón de las dos varianzas poblacionales.

Ejercicio 3: Una empresa desea analizar las horas promedio de trabajo a la semana de las áreas de finanzas y de recursos humanos. Para lo cual se tomaron dos muestras independientes de 21 personas de cada uno de estos departamentos. Las varianzas de las muestras fueron 12 y 10 respectivamente. Construya un intervalo de confianza de 95% para la razón de las varianzas de las dos poblaciones.

Ejercicio 4: Se analizó estadísticamente la cantidad de artículos que se venden en dos áreas de una empresa de artículos deportivos. Del área uno se muestrearon 16 artículos con una varianza de 9.90 y del área dos se muestrearon 13 artículos con una varianza de 6.30. Construya un intervalo de confianza de 95% para la razón de las varianzas de las dos poblaciones.

SELECCIÓN DEL TAMAÑO DE MUESTRA PARA ESTIMAR LA MEDIA POBLACIONAL

Nivel de confianza

90% = 1.645
91% = 1.69
92% = 1.75
93% = 1.81
94% = 1.88
95% = 1.96
96% = 2.05
97% = 2.17
98% = 2.33
99% = 2.575

Población Finita	Población infinita
$n = \frac{NZ^2\sigma^2}{(N-1)e^2 + Z^2\sigma^2}$	$n = \frac{Z^2\sigma^2}{e^2}$

EJERICICOS CON POBLACION INFINITA (SIGNIFICA QUE NO SE CONOCE EL TAMAÑO DE LA POBLACIÓN)

Ejercicio 1: Se desea estimar el contenido medio de un refresco con un nivel de confianza del 93% y un error máximo de estimación de 5 ml. Una muestra previa de 5 refrescos indica que la desviación del contenido es de 12 ml. Calcular el tamaño de la muestra.

Ejercicio 2: Un gerente de personal de una empresa desea estimar con una confianza del 95%, el tiempo promedio de capacitación que recibieron los empleados de la compañía y desea que el error de la estimación no rebase los 30 minutos (0.5 horas). Con base en los resultados de estudios anteriores estima que la desviación estándar es de 3 horas. Calcular el tamaño de la muestra.

Ejercicio 3: Un restaurante desea conocer con un nivel de confianza de 90%, la cantidad de clientes promedio que acuden entre 11 a.m. y 6 p.m. los lunes, y considera que el error máximo es de 2 clientes; basándose en las muestras del resto de los días de la semana, estima que la desviación estándar es de 5 clientes. Calcule el tamaño de la muestra para llevar a cabo dicha estimación.

EJERCICIOS CON POBLACIÓN FINITA (SIGNIFICA QUE SI SE CONOCE EL TAMAÑO DE LA POBLACIÓN)

Ejercicio 1: Un productor quiere estimar con un nivel de confianza de 99% y un margen de error de 1, el promedio de manzanas por caja que llegan en mal estado al almacén de donde se surten todos los pedidos, se sabe que la desviación estándar es de 2.5. Calcule el tamaño de la muestra, sabiendo que el número de cajas en cada trayecto es de 5,000.

Ejercicio 2: Una empresa desea estimar el promedio de trabajadores que requieren utilizar lentes con un nivel de confianza de 90% y un margen de error de 2. Se sabe que la desviación estándar es de 8. Calcule el tamaño de muestra sabiendo que el número de trabajadores de la empresa es de 3225.

Ejercicio 3: Un gerente de personal desea estimar el número promedio de horas de capacitación que toman los vendedores de la empresa con un margen de error de 2 horas y un nivel de confianza de 95%. Con base en estudios anteriores se sabe que la desviación estándar es de 10 horas. Calcule el tamaño de la muestra sabiendo que el número de vendedores de la empresa es de 2000.

SELECCIÓN DEL TAMAÑO DE MUESTRA PARA ESTIMAR PROPORCIONES

Población Finita	Población infinita
$n = \frac{N Z^2 p (1 - p)}{(N - 1)e^2 + Z^2 p (1 - p)}$	$n = \frac{Z^2 p (1 - p)}{e^2}$

POBLACION INFINITA

Ejercicio 1: Se desea estimar con un nivel de confianza del 97% el porcentaje de clientes potenciales que comprarían n nuevo producto. Para esto se toma una muestra previa de 80 clientes de los cuales 65 manifiestan que comprarían el nuevo producto. Si se desea un error máximo de estimación de 6%, calcule el tamaño de la muestra.

Ejercicio 2: Se desea tomar una muestra para estimar con una confianza de 95% la proporción de artículos defectuosos en un embarque y se desea que el error sea de 5%. Si se sabe que la proporción de artículos defectuosos en esta clase de embarques fue de 2% en el pasado, determine el tamaño mínimo necesario para la muestra.

Ejercicio 3: La dirección de una escuela primaria planea incorporar clases de natación al programa escolar, por lo cual quiere estimar la proporción de alumnos que saben nadar con un nivel de confianza de 99% y un error no mayor a 11%. Determine el tamaño de la muestra.

POBLACION FINITA

Ejercicio 1: El director de una escuela secundaria desea estimar la proporción de alumnos que cuentan con el esquema completo de vacunas con un margen de error de 3% y un nivel de confianza de 95%. Se sabe que, en el pasado, 70% tenía cubierto el esquema de vacunación. Calcule el tamaño de la muestra sabiendo que asisten 1350 alumnos a la escuela.

Ejercicio 2: El departamento de administración escolar desea estimar la proporción de alumnos en el último semestre que pretende estudiar alguna maestría, con un nivel de confianza de 97% y un error de 8.5%; anteriormente 31% de los estudiantes expresaron interés por seguir estudiando. Calcule el tamaño de muestra si el total de alumnos en el 9º semestre es de 1340.

Ejercicio 3: Se desea estimar la proporción de alumnos en el programa de becas institucionales de una escuela preparatoria que mantiene un promedio de 9 o más y tiene derecho a renovarla, con un margen de error de 4% y un nivel de confianza de 96%. En años anteriores 58% de los becarios renovaron dicha beca. Calcule el tamaño de muestra, si el patrón total de becarios es de 2720 alumnos.

ANALISIS DE VARIANZA DE UN FACTOR

Ejercicio 1: Se hace un estudio sobre la efectividad de 3 marcas de spray para matar moscas. Para ello cada producto se aplica a un grupo de 100 moscas, y se cuenta el número de moscas muertas expresado en porcentajes. Se hacen 6 réplicas y los resultados son los siguientes:

MARCA 1	MARCA 2	MARCA 3
72	55	64
65	59	74
67	68	61
75	70	58
62	53	51
73	50	69

- Formule la hipótesis
- Realice el Anova al 5%
- Calcule el C.V.

Ejercicio 2: Una empresa ensambla cuadros para un solo modelo de bicicleta. La planta trabaja tres turnos: El matutino, el vespertino y el nocturno. Los trabajadores se rotan cada uno de los turnos, por lo que, al cabo de un trimestre, todos ellos laboraron en cada uno de los tres turnos. El director de producción quisiera saber si existe diferencia entre la producción promedio de los 3 turnos, para ello, obtiene una muestra aleatoria de la producción de 6 días de cada turno. Los datos se muestran a continuación:

MATUTINO	VESPERTINO	NOCTURNO
129	138	118
141	142	120
128	140	132
145	149	118
135	129	136
144	148	138

- Formule la hipótesis
- Realice el Anova al 5%
- Calcule el C.V.

Ejercicio 3: Los fabricantes de automóviles compactos de 4 marcas distintas promocionan sus autos como los más económicos en cuanto a consumo de combustible. Una asociación civil de defensa del consumidor supone que los rendimientos de los autos son en promedio iguales. Para comprobar este supuesto, sujeta a una muestra de carros nuevos y en buenas condiciones mecánicas a un mismo recorrido mixto de ciudad y carretera, a las mismas horas del día, y en el mismo día de la semana. Los resultados de rendimiento en km/l se muestran en la siguiente tabla:

M A R C A			
A	B	C	D
16	18	22	20
19	19	21	17
18	17	20	16
16	18	19	20
17	18	21	16
20	21	20	20

- Formule la hipótesis
- Realice el Anova al 5%
- Calcule el C.V.

Ejercicio 4: Una planta de productos químicos tiene 4 plataformas de embarque distintas para sus productos, que deben trabajar de manera uniforme. Existen también 4 supervisores que se rotan entre las plataformas todas las semanas. Los supervisores tienen la facultad de autorizar horas extra a su personal para realizar labores extraordinarias. Existe la preocupación en la gerencia de embarques de que alguno o algunos de los supervisores tengan más propensión a autorizar horas extra que otros. Para investigarlo se obtuvieron muestras aleatorias de las horas extra por semana autorizadas por cada uno de los supervisores. Los datos obtenidos son los siguientes:

LOPEZ	MANCILLA	ORTIZ	OROPEZA
56	77	61	71
52	70	54	79
57	75	72	64
68	86	75	85

- Formule la hipótesis
- Realice el Anova al 5%
- Calcule el C.V.

PRUEBAS DE HIPOTESIS

PRUEBA DE HIPÓTESIS PARA LA MEDIA CON DISTRIBUCIÓN T DE STUDENT

Ejercicio 1: Un fabricante de papas fritas empaca su producto en bolsas de 28 gramos, los contenidos tienen una distribución normal. Se toma una muestra de 16 bolsas y se obtuvo un peso promedio de 29.22 gramos, con una desviación estándar de 2.11. Desde el punto de vista del fabricante ¿se está cumpliendo con lo prometido en el empaque? Use un nivel de significancia de 10%.

Ejercicio 2: Se aplica una prueba de autoestima a 25 personas quienes obtienen una calificación promedio de 62.1 con una desviación estándar de 5.83. se sabe que el valor correcto de la prueba debe ser mayor a 60. ¿Existe suficiente evidencia para comprobar que no hay problemas de autoestima en el grupo seleccionado? Considere un nivel de significancia de 0.05

Ejercicio 3: Una fábrica de electrodomésticos afirma que una aspiradora consume un promedio de 46 kw – hora al año. Si en un estudio se toma una muestra aleatoria de 12 aspiradoras y se encuentra que consumen un promedio de 42 kilovatios –hora al año con una desviación estándar de 11.9 kw-hora, ¿se podría afirmar con un nivel de significancia de 0.05 que las aspiradoras consumen, en promedio, menos de 46 kilovatios?

PRUEBA DE HIPÓTESIS PARA LA MEDIA CON DISTRIBUCIÓN NORMAL

Ejercicio 1: En un colegio se hizo un estudio para determinar qué tan lejos de la institución viven los estudiantes. Para ello se tomó una muestra de 50 estudiantes y se encontró que el promedio de la distancia del colegio a las casas de ellos es de 102.6 kilómetros, con un nivel de significancia de 0.10. Probar la hipótesis según la cual el promedio de las distancias es menos de 105 km si se sabe que la desviación estándar es igual a 5 km.

Ejercicio 2: El dueño de un café desea saber si la edad promedio de las personas que entran a su negocio es menor o igual a 20 años. Para determinar lo anterior se realiza una muestra aleatoria de 40 personas dando un promedio de 22 años y una desviación estándar de 3.74 años. Realiza la prueba de hipótesis con un nivel de significancia de 0.1

Ejercicio 3: Fumar cigarrillos de la marca X produce en promedio 0.6 mg de nicotina. El departamento de ingeniería del fabricante propone un filtro nuevo que supuestamente reduce la producción de nicotina. Se toma una muestra de 50 cigarrillos con el nuevo filtro y se encuentra que el promedio de nicotina es de 0.55 mg con desviación estándar de 0.56 ¿se debe aceptar la aseveración del departamento de ingeniería con un nivel de significación de 2.5%?

Ejercicio 4: Una institución de enseñanza de idiomas registra en promedio 4 personas semanalmente en los cursos que ofrece, por lo que hace una campaña publicitaria de gran magnitud. Después de un año toma una muestra aleatoria de 32 semanas y encuentra que en promedio se registran semanalmente 5 alumnos nuevos con una desviación estándar de 0.05. Determine si la campaña publicitaria cumplió su objetivo de que el número de estudiantes inscritos semanalmente sea mayor a 4.

Ejercicio 5: Una psicóloga desea determinar si el tiempo promedio que tarde un conductor adulto en reaccionar a cierta situación de urgencia es en realidad 0.56 segundos. De la información contenida en estudios similares ella puede suponer que la variabilidad de esas mediciones está dada por una desviación típica de 0.082 segundos. Así mismo decide basar la demostración en una muestra aleatoria de tamaño 35 y utilizar el nivel de significancia de 0.05 ¿Qué concluirá si sus datos producen una media de 0.59 segundos?

Ejercicio 6: Un fabricante de llantas para automóviles afirma que la duración promedio de determinado modelo de llantas en un coche de cierto peso es de 40,000 km bajo condiciones normales de manejo. Se analiza una muestra aleatoria de 100 llantas de este tipo bajo las condiciones especificadas y se encuentra que la duración promedio fue de 39000 km con una desv. estándar de 8500 km. Realice una prueba de hipótesis para decidir si la afirmación del fabricante es aceptable o no con un nivel de significancia del 0.05