UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

FACULTAD DE INGENIERÍA

“POTENCIA DE UNA PRUEBA ESTADÍSTICA”

TESINA

Que para obtener el título de:

INGENIERO EN AUTOMATIZACIÓN

(SISTEMAS INDUSTRIALES)

Presentan:

Andrea Beltrán Pérez

Expediente 159846

Marcel De Jesús Moctezuma Reyes

Expediente 150176

Dr. Juan José Méndez Palacios

Asesor De Tesina

Santiago de Querétaro, Qro.

Noviembre del 2011

ÍNDICE

DEDICATORIAS

Andrea Beltrán Pérez

Le doy gracias adiós por tener esta familia tan maravillosa.

Con mucho amor y cariño les dedico principalmente a mis padres que me dieron la vida y han estado conmigo en todo momento. Gracias por todo papá y mamá por darme nuevamente otra carrera para mi futuro, pero sobre todo por creer en mí, aunque hemos pasado momentos difíciles siempre han estado apoyándome y brindándome todo su amor, por todo esto les agradezco de todo corazón. Darme la mejor herencia: mi Ingeniería GRACIAS PAPÁS.

A mis hermanos Blanca, Alberto y Cristina gracias por estar por apoyarme y estar conmigo en una de las etapas más felices de mi, los quiero mucho.

También a estos dos peques Faby y Gael que en momentos de estrés se acercaban a mí para hacerme reír, con sus locuras y travesuras los adoro.

Le doy gracias a mis abuelos por estar conmigo y compartir los momentos que viví durante la carrera, aunque uno partió a mitad de mi carrera su alma me acompañara por siempre.

También le doy gracias a mi novio por apoyarme, darme consejos, sobre todo por darme fuerzas y estar a mi lado durante este largo camino que hoy llega a su fin. GRACIAS amor te amo.

Marcel De Jesús Moctezuma Reyes

A Dios:

Porque a través de mis padres, hermanos, familiares y amigos me has obsequiado una vida maravillosa. Por todas las bendiciones con las que llenas mi vida, pero sobre todo, por las dificultades que has colocado en ella. Esos obstáculos han forjado mi carácter y me han permitido valorar y amar mucho más todo lo que tú me has dado. Gracias Dios Mío.

A Víctor Zacarías Sánchez:

Un amigo, un padre, un benefactor. A usted especialmente, que se durmió en la espera de éste momento, le dedico todo. Gracias por aparecer en mi vida, por apoyarme en momentos verdaderamente inciertos, por abrirme las puertas de su hogar y por ser parte fundamental de mi carrera. Los momentos que viví a su lado, los consejos que me dio, la amistad que me brindó y el cariño que siempre me demostró, vivirán en mí para toda la eternidad.

"Tienes derecho a todo, menos a fallarle a tu familia"

Sus palabras siempre harán eco en mi vida. Gracias Don Vic.

I. RESUMEN

La filosofía Seis Sigma busca ofrecer mejores productos o servicios, de una manera cada vez más rápida y a más bajo costo, mediante la reducción de la variación de cualquiera de los procesos. Aunque a muchas personas les ha costado entender, una de las grandes enseñanzas del Dr. Deming fue buscar el control de variación de los procesos lo cual es medido por medio de la desviación estándar. Decía el Dr. Deming: “el enemigo de todo proceso es la variación, por lo que es ahí en donde debemos concentrar el esfuerzo hacía la mejora continua”, pero sobre todo porque “La variación es el enemigo de la satisfacción de nuestros clientes”.

El Seis Sigma es un parámetro cuya base principal es la desviación estándar y su enfoque es reducir la variación y/o defectos en lo que hacemos.

El principal planteamiento lo podemos encontrar cuando se considera la variación de un proceso, con una fluctuación entre más 6 sigma y Menos 6 sigma del valor promedio, la probabilidad de que se salga del valor especificado es de 3.4 partes por millón.

El valor de Seis Sigma sirve como parámetro de comparación común entre compañías iguales o diferentes e inclusive entre los mismos departamentos de una empresa, tan diferentes como compras, cuentas por cobrar, mantenimiento, ingeniería, producción, recursos humanos, etc.

Es una filosofía que busca obtener mejores resultados (productos, servicios), por medio de procesos robustos que permitan reducir los defectos y los errores. Se podría considerar como una metodología (Lógica y/o disciplinada) de pasos, por medio de herramientas probadas para la solución de problemas.

Considerando que en el uso de Seis Sigma se necesita del método potencia de una prueba estadística dentro del estudio de la inferencia, se describe como se puede tomar una muestra aleatoria y a partir de esta muestra estimar el valor de un parámetro poblacional en la cual se puede emplear el método de muestreo y el teorema del valor central lo que permite explicar cómo a partir de una muestra se puede inferir algo acerca de una población, lo cual nos lleva a definir y elaborar una distribución de muestreo de

medias muestrales que permitan explicar el teorema del límite central y utilizar este teorema para encontrar las probabilidades de obtener las distintas medias muestrales de una población.

Pero es necesario tener conocimiento de ciertos datos de la población como la media, la desviación estándar o la forma de la población, pero a veces no se dispone de esta información.

En este caso es necesario hacer una estimación puntual que es un valor que se usa para estimar un valor poblacional. Pero una estimación puntual es un solo valor y se requiere un intervalo de valores a esto se denomina intervalo de confianza y se espera que dentro de este intervalo se encuentre el parámetro poblacional buscado. También se utiliza una estimación mediante un intervalo, el cual es un rango de valores en el que se espera se encuentre el parámetro poblacional.

En éste caso, se desarrolla un procedimiento para probar la validez de una aseveración acerca de un parámetro poblacional. Éste método es denominado prueba de hipótesis para una muestra.
El presente trabajo mostrará algunos de los conceptos más importantes en el diseño investigativo de Seis Sigma, centrándose en tres aspectos cada vez más utilizados en su metodología: la potencia estadística, la sensibilidad y el tamaño de efecto.

El objeto de éste trabajo será pues mostrar la importancia de tener un control óptimo en la variación de los tres factores antes mencionados, ejemplificando algunas de sus posibles aplicaciones prácticas.

“Medimos para entender,
Entendemos para decidir,
Decidimos para mejorar,
Mejoramos para evolucionar…”

II. JUSTIFICACIÓN

La inferencia estadística, como ya se mencionó, está relacionada con los métodos para obtener conclusiones o generalizaciones acerca de una población. Estas conclusiones sobre la población pueden estar relacionadas ó con la forma de la distribución de una variable aleatoria, ó con los valores de uno o varios parámetros de la misma.

El campo de la inferencia estadística se divide en dos: Por un lado se tiene el problema de la estimación de los parámetros de una distribución, y por el otro, las pruebas de hipótesis. En el problema de estimación se trata de elegir el valor de un parámetro de la población, mientras que en las pruebas de hipótesis se trata de decidir entre aceptar o rechazar un valor especificado.

Ya que lo relacionado con las pruebas de hipótesis se experimenta en todas partes, se pueden considerar dos áreas: Pruebas de hipótesis sobre parámetros, para determinar si un parámetro de una distribución toma o no un determinado valor, y Pruebas de Bondad de Ajuste, para definir si un conjunto de datos se puede modelar mediante una determinada distribución.

Una hipótesis estadística es una proposición o conjetura con respecto a una o más poblaciones. Estas aseveraciones o suposiciones pueden ser con respecto a uno o varios parámetros, ó con respecto a la forma de las respectivas distribuciones de probabilidad. También es posible considerar una hipótesis estadística como una proposición sobre la distribución de probabilidad de una variable aleatoria ya que emplea distribuciones de probabilidad para representar poblaciones.

Por lo tanto surgiría la pregunta ¿Cuál es el rol de la sensibilidad y la potencia en una investigación? Definida de manera simple, la sensibilidad de un diseño investigativo reside en su capacidad de detectar diferencias o efectos donde los hay.

Es ahí, donde dada su importancia, la potencia estadística se manifiesta como una convención que vale la pena tener presente al momento de emprender una investigación cuantitativa.

III. OBJETIVO

Cuando se prueba una hipótesis hay dos posibilidades: rechazarla o no rechazarla (los estadísticos suelen decir que no rechazar una hipótesis no implica aceptarla, por lo que se utiliza la frase “no hay evidencia para rechazar la hipótesis nula”). Cualquiera que sea la decisión hay una probabilidad conocida de cometer un error.

Cuando la hipótesis nula se rechaza esa probabilidad es el error tipo I al que se ha llamado tradicionalmente “nivel de significancia”. Cuando la hipótesis nula no se rechaza la probabilidad de equivocarse se denomina error tipo II.

La consideración del primero está presente en todas las investigaciones en las que se realiza prueba de hipótesis, la consideración del segundo suele estar ausente.

Ya que nuestro objetivo es destacar la importancia de potencia de una prueba estadística, cómo ésta se manifiesta y su interpretación, tanto cuando la hipótesis que se desea probar está planteada en la hipótesis nula como cuando no lo está.

El objetivo fundamental es evidenciar la importancia del poder en una prueba estadística basados en el sentido de confiabilidad que se requiere en todo estudio científico.
Dado que en el mundo de negocios actual, un enfoque de procesos es esencial siendo que todas y cada una de las actividades, funciones o tareas dentro de toda organización pueden ser consideradas como un proceso, ahí la importancia de que los estudios inherentes a estos procesos sean confiables en cuanto a su alcance técnico y práctico.

IV. ANTECEDENTES

4.1 Conceptos de mejora contemporáneos

La mejora de la calidad tiene muchas acepciones aunque una de las más adecuadas podría ser "parte de la gestión orientada a mejorar su eficacia y eficiencia y, con ello, la satisfacción del cliente al menor coste". Considerando que es un sistema de calidad que está compuesta por un conjunto de normas interrelacionadas de la organización de forma ordenada la administración de la calidad de la misma (Escalante, 2006), respecto a la mejora continua.

Existen multitud de herramientas aplicables a entidades de formación que deseen mejorar su gestión y la calidad de las actividades formativas que ofrece.

- ❖Estructura de organización: se refiere a la forma en que se dividen, agrupan y coordinan las actividades de la organización en cuanto al organigrama de los sistemas de la empresa u organización, donde se jerarquizan los niveles directivos y de gestión.
- ❖Estructuras de responsabilidades: implica a personas y departamentos. La forma más sencilla de explicitar las responsabilidades de la calidad, es mediante un cuadro de doble entrada, donde mediante un eje se sitúan los diferentes departamentos y en el otro, las diversas funciones de la calidad.
- ❖Procedimientos: sucesión cronológica de operaciones concatenadas entre sí que se constituyen en una unidad de función para la realización de una actividad o tarea específica dentro de un ámbito predeterminado de aplicación.
- ❖Procesos: es la sucesión completa de una operación dirigiendo la realización de un objetivo específico.
- ❖Recursos: estos no solo dependen de lo financiero si no de conjunto de personas, bienes materiales y técnicos con que cuenta y utiliza una dependencia, entidad, u organización para alcanzar sus objetivos y producir los bienes o servicios que son de su competencia.

4.1.1 Gemba

Gemba introduce una nueva palabra a la cultura gerencial japonesa que significa lugar real ó el lugar donde usted trabaja donde se agrega valor (Evans, Lindsay, 2000). En el sistema manufacturero, el gemba es el piso de la fábrica, en cada industria es un lugar distinto. Se trabajo con un equipo compuesto por directivos, gerentes y operarios con base del mejoramiento continuo. Se debe de considerar los siguientes puntos ante un problema:

- Detectar el lugar donde se originó el problema (Gemba).
- Revisar objetivos relevantes (Gembutsu).
- Implementar medidas de contención inmediatas.
- Encontrar la causa raíz.
- Estandarizar para prevenir recurrencia.

Administrar el gemba como base de la mejora continua puede implementarse con facilidad, ya que se trata de una disciplina de sentido común. No se requiere que la empresa haya logrado niveles de calidad estable. Requiere trabajo intenso y constante de la organización.

4.1.2. Kaizen

El Kaizen se puede definir como aquella forma que buscan las empresas para realizar un mejoramiento continúo en base a pequeños cambios. El término Kaizen proviene del japonés, en el que “Kai” se traduce al español como “cambio”, y “Zen” que se entiende como “mejoramiento”.

La idea del Kaizen no es realizar grandes cambios, si no que a partir de pequeñas y simples modificaciones, poder mejorar la calidad y reducir los costos de producción. De este modo, se cambian todos aquellos aspectos que no permiten mejorar el servicio a los clientes ni mejorar la calidad de los productos. La idea es ir realizando mínimas modificaciones a diario, ya que, a fin de cuentas se habrán realizado más de 300 mejoras en solo un año.

Así, sin tener que realizar grandes esfuerzos, a través del Kaizen se favorecerá la reducción de costos y un mejoramiento integral de la empresa. Lo anterior se realiza llevando a cado las denominadas Ss. La primera de ellas es "Seire", que se entiende como organización. A través de su implementación, es posible poner cada cosa en su lugar y encontrar un lugar para cada cosa. De esta manera, hay que preguntarse si cada persona del personal tiene bien claras sus funciones y si se sabe en qué lugar se encuentran ubicadas las cosas, además es necesario verificar si se realizan reuniones para ver si las tareas se han cumplido.

La segunda S es "Seiton" y se refiere a reducir búsquedas, pudiendo así facilitar el movimiento de personas, servicios y cosas. La idea es no perder tiempo valioso de trabajo en buscarlos. En tercer lugar se encuentra "Seiketsu", para la simplificación de los procesos. En este momento es necesario cuestionarse si existe algo que se pueda hacer para mejorar los procesos tanto logísticos, como los de ventas y servicios. Preguntarse por cuánto tiempo se pierde en ellos, cuánto tiempo se utiliza en trámites innecesarios. La idea es eliminar todos aquellos procesos que no son más que verdaderas pérdidas de tiempo.
En cuarto lugar, la siguiente S, es "Shitsuke", para mejorar la disciplina y los buenos hábitos de trabajo. El objetivo es respetar las normas de trabajo, en todos los niveles de la jerarquía empresarial, a fin de que el ambiente laboral sea justo y equilibrado.

"Seiso" es la quinta S. A través de ésta se hace alusión a la necesidad de limpieza. En otras palabras, la limpieza se hace nuestra aliada a la hora de mantener el orden, por lo tanto, es necesario verificar la limpieza de los puestos de trabajo, los baños y la pulcritud de la entrega de productos al cliente. La idea a la base es que la limpieza nos lleva al orden, el que a su vez conduce a la reducción de costos.

Para lograr una efectiva implementación de Kaizen resulta indispensable contar con el apoyo de todo el personal, quien también debe tener una participación activa realizando sugerencias acerca de los cambios que es necesario realizar, ya que el compromiso es imprescindible si se quiere que las mejoras resulten.
Se debe de considerar en primera instancia un Equipo de Mejoras y Proyectos Estratégicos (EMPE). Constituido por los máximos responsables de la organización, serán los encargados no sólo de desarrollar las estrategias, sino además de apoyar las actividades para poner en práctica éstas. Una de sus funciones principales será aprobar los presupuestos y proyectos de mejoras que se le eleven.

El EMPE tendrá entre sus tareas fundamentales las de analizar, corregir y eliminar las Mudas (desperdicio en japonés) Estratégicos. ¿Cuáles son? ¿Por qué son estratégicos? ¿Cómo eliminarlos?

Los Mudas estratégicos están constituidos por:

1. Las capacidades de empleados desaprovechadas.
2. La falta de enfoque y posicionamiento
3. Tiempo
4. Información
5. Oportunidades del entorno
6. Fortalezas de la empresa
7. Clientes / Consumidores

4.1.3 Lean

Lean es un sistema y filosofía de mejoramiento de procesos de manufactura y servicios basado en la eliminación de desperdicios y actividades que no agregan valor al proceso. Permitiendo alcanzar resultados inmediatos en la productividad, competitividad y rentabilidad del negocio.

Los principios clave de Lean (Baena, 2008)

- Calidad perfecta a la primera: Detección de cero defectos y solución de los problemas en su origen.
- Minimización de desperdicios: Optimizar el uso de recursos y eliminación de actividades que no sean de valor agregado para el producto o proceso.
- Mejora continua: pretende mejorar los productos, servicios con el fin de mejorar la calidad, aumentando la productividad y reducción de costos.
- Sistema de Producción “Pull”: Consiste en mantener pequeñas cantidades de inventario y evitar la sobre producción.
- Flexibilidad: Producir rápidamente gran variedad de productos, sin suprimir la eficiencia debido a volúmenes menores de producción.
- ¿Quiénes participan? directivos, gerentes y operarios que promueven la mejora continua.

- ❖ El líder Lean puede ser el supervisor al cual se le han capacitado en las herramientas básicas de Lean.

4.1.4 Seis Sigma

El concepto de Seis Sigma provee una medición común, así como objetivos comunes, a la vez que inculca una visión común y sobre todo promueve el trabajo en equipo.

Adicionalmente combina objetivos agresivos con un método y un conjunto de herramientas, que se aplican a través de todo el ciclo de vida del proceso o servicio: Existe una alta correlación entre le mejora del tiempo de ciclo y la reducción de defectos y costos. Muchas empresas utilizan el concepto de Seis Sigma para establecer un parámetro de negociación durante los procesos de negociación Cliente – Proveedor Interno.

Han existido dos filosofías sobre la calidad, la primera de ellas la que llamaríamos la filosofía antigua, se basaba en cumplir con las especificaciones o requerimientos del cliente, un precursor de ello fue Crosby, con su teoría de que la "Calidad es Gratis" y la nueva filosofía la cual predica que las pérdidas de calidad están basadas en la desviación de la meta u objetivo de acuerdo a los requerimientos o especificaciones. Esto quiere decir que cualquier producto o servicio desviado del centro o meta, no cumple la norma de calidad, sobre ésta última es que se basa el concepto de Seis Sigma.

En el proceso de introducción del Seis Sigma, uno de los conceptos que más se aplica, son una serie de pasos conocidos por sus siglas DMAIC, con lo cual se busca establecer la fuente u origen de la variación.

DEFINIR

En la fase de medición se identifican los posibles proyectos Seis Sigma, que deben ser evaluados por la dirección para evitar mal gastar el dinero y recursos, antes de tiempo. Una vez seleccionado el proyecto se prepara su misión y se selecciona el equipo más adecuado, asignándole la prioridad necesaria (Gutiérrez y De La Vera, 2008).

MEDIR

La fase de medición es la que nos interesa, pues consiste en la caracterización del proceso identificando los requisitos clave de los clientes, las características clave del producto (o variables del resultado) y los parámetros (variables de entrada) que afectan al funcionamiento del proceso y a las características o variables clave (Miranda, 2006). A partir de ésta caracterización, se define el sistema de medida y se mide la capacidad del proceso.

ANALIZAR

En la tercera fase, el equipo analiza los datos de resultados actuales e históricos. Se desarrollan y comprueban hipótesis sobre posibles relaciones causa-efecto utilizando las herramientas estadísticas pertinentes. De ésta forma el equipo confirma los determinantes del proceso, es decir, las variables clava de entrada o “poco vitales” que afectan a las variables de salida del proceso (Miranda, 2006).

MEJORAR

En la fase de mejora el equipo trata de determinar la relación causa-efecto (relación matemática entre las variables de entrada y la variable de respuesta que interese) para predecir, mejorar y optimizar el funcionamiento del proceso (Brown and Morrison, 1991). Por último, se determina el rango operacional de los parámetros o variables de entrada del proceso.

CONTROLAR

La última fase consiste en diseñar y documentar los controles necesarios para asegurar que lo conseguido mediante el proyecto se mantenga una vez que se hayan implementado los cambios. Para prevenir que la solución sea temporal, se documenta el nuevo proceso y su plan de monitoreo (Miranda, 2006).

V. INTRODUCCIÓN

La potencia estadística se atribuye, a Neyman y Pearson (1928, 1933). A partir de entonces, aparecieron una serie de autores que tenían en cuenta en sus estudios la potencia estadística (Cox, 1948; McNemar, 1960; Sterling, 1959; Tukey, 1960; Tullock, 1959). Dentro de la literatura de las ciencias sociales, es de destacar Mosteller y Bush (1954). Sin embargo, no fue hasta 1962 cuando apareció por primera vez, en el ámbito de las ciencias sociales, un estudio sistemático de la potencia estadística.

Se trata del trabajo realizado por Cohen (1962) en el que se destacó, en primer lugar, la importancia de la potencia estadística dentro de la investigación experimental y, en segundo lugar, proporcionó una serie de pautas para llevar a cabo un análisis de potencia. Este estudio alentó a los investigadores a prestar mayor atención a la potencia de las pruebas y a no centrarse, exclusivamente, en el nivel de significación.

Posteriormente, Cohen (1965, 1969) publicó una serie de recomendaciones para tratar con eficacia el problema de la potencia y, más recientemente, elaboró unas tablas muy útiles para su cálculo (Cohen, 1988). Sin duda, el trabajo de Cohen (1988) ha inspirado varios estudios sobre la potencia y el tamaño del efecto en áreas de las ciencias sociales, así como diversos programas de ordenador (examinados en Goldstein, 1989).

Además, si se toman como evidencia manuales de estadística, se observa que con frecuencia la potencia no es tratada. No se sabe con certeza la razón por la que los investigadores dejan de lado el análisis de la potencia. Según Cohen (1992), una buena parte se debe a la poca importancia que se da al tamaño del efecto y a la mayor relevancia de los resultados estadísticos.

Chase y Tucker (1976) y Sedlmeier y Gigerenzer (1989) atribuyen este hecho al predominio de la teoría fisheriana. La escuela de Fisher considera las pruebas estadísticas como pruebas de significación (Fisher, 1935, 1950, 1955), mientras que la escuela de Neyman-Pearson las conceptúa como pruebas de decisión (Neyman y Pearson, 1928, 1933, 1936).
Otra distinción importante entre estas dos aproximaciones se refiere a si el nivel de significación empleado en el análisis debe fijarse anticipadamente.

Aunque Fisher (1935, 1950, 1955) daba prioridad a un nivel de significación de 0.05, nunca prescribió que tal nivel debiera mantenerse fijo o que debiera establecerse antes de llevar a cabo el experimento.
Por el contrario, la posición de Neyman-Pearson (1928, 1933, 1936) requiere que el nivel de significación se determine antes de cualquier análisis estadístico y que el investigador se adhiera a él en todas las decisiones estadísticas.

En este sentido, si el nivel de significación no se estableciera con anterioridad, se podría caer en el error de fijar un valor a basado más en los resultados que en una estimación de la probabilidad.

Según Sedlmeier y Gigerenzer (1989), este continuado desinterés por la potencia de las pruebas estadísticas sólo cambiará cuando los editores de las principales revistas exijan, dentro de su política editorial, que los autores estimen la potencia de sus pruebas de significación. De todos modos, en los últimos años han aparecido una serie de textos básicos sobre potencia estadística (Cohen, 1988; Kraemer y Thiemann, 1987; Lipsey, 1990, entre los más destacados).

Desafortunadamente, la mayor parte de estas fuentes aún no contienen mucha información sobre cómo calcular la potencia o están limitadas a la consideración de pruebas tradicionales. Por otro lado, están disponibles programas informáticos como por ejemplo el Power (Borenstein y Cohen, 1988) que calcula la potencia del análisis de la variancia unidireccional (Borenstein et al., 1990), de la regresión múltiple (Rothstein et al., 1990), y de correlaciones y diferencias entre proporciones (Borenstein y Cohen, 1988).

VI. MARCO TEÓRICO

La estadística es la disciplina que nos proporciona una metodología para recoger, organizar, resumir, analizar datos y hacer inferencias a partir de ellas.

Puede deducirse de la definición que hay dos ramas claramente diferenciadas dentro de la estadística: La Estadística Descriptiva y la Inferencia Estadística que es el punto a tratar en el presente trabajo.
La inferencia Estadística tiene como función inferir las características de un colectivo a partir de un subconjunto de éste.

Referente al contraste de hipótesis, sabemos que un problema es investigable cuando existen dos o más soluciones alternativas y tenemos dudas acerca de cuál de ellas es la mejor. Esta situación permite formular una o más hipótesis de trabajo, ya que cada una de ellas destaca la conveniencia de una de las soluciones sobre las demás.

Si nuestro propósito es comprobar una teoría ella misma será la hipótesis del trabajo, pero es importante destacar que al formular dicha o dichas hipótesis no significa que ya esté resuelto el problema, al contrario, que nuestra duda nos impulsa a comprobar la verdad o falsedad de cada una de ellas.
La decisión final partirá de las decisiones previas de aceptar o rechazar las hipótesis de trabajo.

6.1 Conceptos básicos

Debido a que el proceso de obtener toda la información relevante a una población particular es difícil y en muchos casos imposibles de obtener, se utilizan muestras para estimar la información necesaria para la toma de decisiones. Con la información obtenida se determina qué tipo de estadística es la que se maneja si descriptiva o inferencial, para de ahí desglosar las pruebas de hipótesis y qué tipo de errores.

6.1.1 Estadística descriptiva

La estadística descriptiva trata del recuento, ordenación y clasificación de los datos obtenidos por las observaciones. Se construyen tablas y se representan gráficos que permiten simplificar la complejidad de los datos que intervienen en la distribución. Asimismo, se calculan parámetros estadísticos que caracterizan la distribución. No se hace uso del Cálculo de Probabilidades y únicamente se limita a realizar deducciones directamente a parámetros obtenidos por:

- Elementos

El primer paso en toda investigación estadística consiste en fijar el conjunto de elementos que queremos estudiar, que llamaremos población o universo. Cada elemento de la población se denomina individuo o unidad estadística. La población puede ser el conjunto de personas de una localidad, las llamadas telefónicas a una central. Llamaremos muestra a un subconjunto limitado extraído de la población, con objeto de reducir el número de experiencias.

Una vez fijada la población debemos indicar cuáles son las características o cualidades que nos interesan estudiar en esa población, estableciendo la forma en la que deben medirse, las unidades de medida.
Estas características observables en una población se clasifican en cualitativas, que son aquellas que no se pueden cuantificar, tales como el color de pelo, el gusto musical, grupo sanguíneo. Las características que no son cualitativas las llamamos cuantitativas, que son aquellas que sí se pueden cuantificar, como es la estatura, el número de hijos.

A su vez, las características cuantitativas se dividen en dos tipos, las discretas y las continuas. Las características cuantitativas discretas son aquellas que toman valores aislados, como es el número de televisores en una unidad familiar o el número de hijos de una pareja. Por el contrario, las variables continuas pueden tomar cualquier valor comprendido en un determinado rango o intervalo, aunque muchas veces la unidad de medida no nos permita tal hecho. Esto ocurre, por ejemplo, al estudiar la altura de una población, que aunque sabemos que es una variable continua, los aparatos de medida sólo nos permiten tomar éstas con una determinada aproximación.

Algunas veces también es preferible, en el caso de las variables discretas con un gran número de resultados, tratarlas como si fueran variables continuas y viceversa.

Una vez obtenida la información referente a la variable de estudio, ésta se organiza y resume en las llamadas distribuciones de frecuencias, que nos proporcionan el número de individuos que hay para cada uno de los valores de la variable. Estas distribuciones de frecuencias pueden ser de frecuencias absolutas, que nos dicen el número de individuos que presentan un determinado valor de la variable, o de frecuencias relativas, que nos dan el tanto por uno o por ciento de la población que presenta es determinado carácter. En ocasiones, también serán de utilidad las frecuencias acumuladas, en las que cada valor acumula los datos pertenecientes también a los que son menores que él.

- ❖ Medidas

Uno de los objetivos de la estadística descriptivas la de resumir toda la información recopilada en unos pocos valores numéricos, para poder sacar consecuencias de esa información. Dentro del conjunto de valores numéricos que resumen toda la información los hay de distinto tipo y que aportan distintas características como:

- Medidas de centralización: Considerando la media, moda, mediana, cuartiles, deciles y percentiles.
- Medidas de dispersión: Considerando la varianza, desviación típica, rango y rango intercuartílico.
- Medidas de forma: Considerando el Coeficiente de variación de Pearson, curtosis de Fisher.
- Relación entre variables: Con el coeficiente de correlación lineal y recta de regresión.

- ❖ Gráficos

Dentro de las técnicas que permiten resumir la información de una variable estadística, los gráficos ocupan un papel fundamental, debido a su facilidad de compresión incluso entre aquellas personas que no poseen conocimientos de estadísticas.

Los diagramas de barras, pictogramas, diagramas de sectores, histogramas, polígonos de frecuencias, diagramas de caja y bigote, pirámides de población, cartogramas, entre otras, ofrecen una información visual muy clara para comprender cómo está distribuida la característica que estamos estudiando en la población.

La introducción del ordenador ha permitido que estos gráficos se obtengan de forma sencilla y rápida con una gran calidad gráfica.

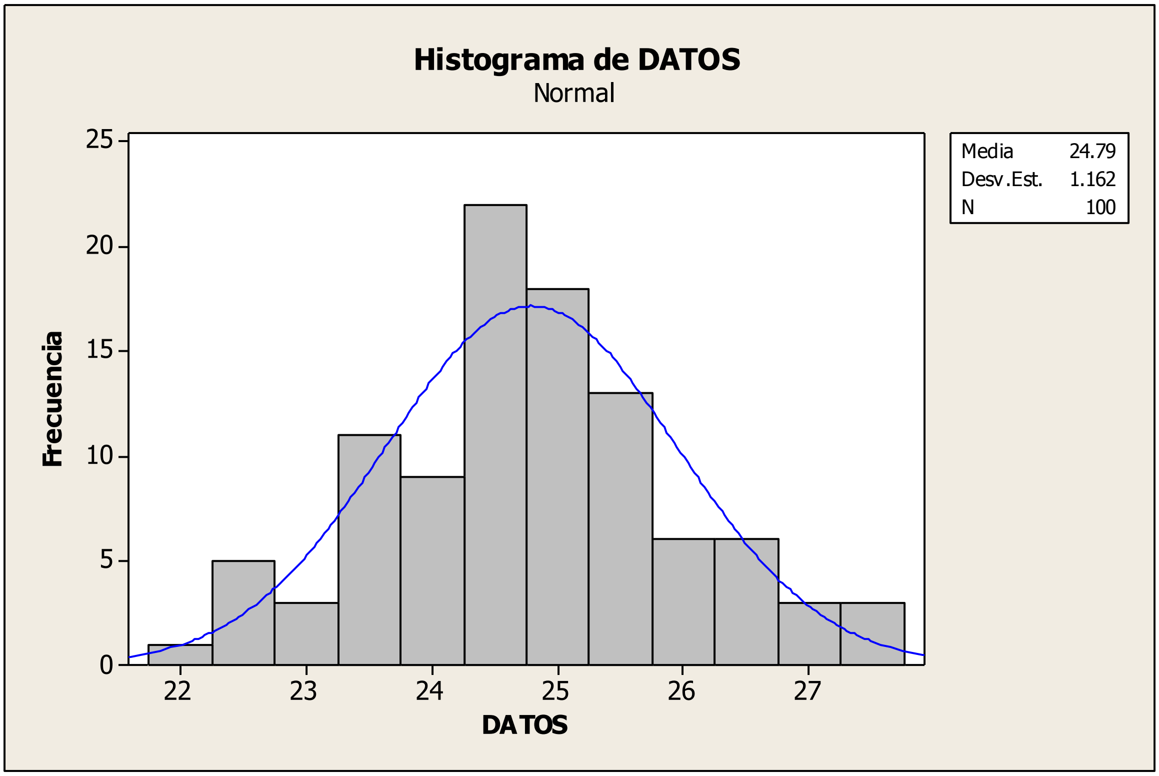

Figura 1. Ejemplo de Estadística Descriptiva

6.1.2 Estadística inferencial

La Estadística Inferencial o inductiva plantea y resuelve el problema de establecer previsiones y conclusiones generales sobre una población a partir de los resultados obtenidos de una muestra.

Los modelos estadísticos actúan de puente entre lo observado (muestra) y lo desconocido (población). Su construcción y estudio están basados en el Cálculo de Probabilidades.

Así pues, la Inferencia Estadística es la metodología tendente a hacer descripciones, predicciones, comparaciones y generalizaciones de una población estadística a partir de la información contenida en una muestra. Utiliza resultados obtenidos mediante la Estadística Descriptiva y se apoya fuertemente en el Cálculo de Probabilidades. Los procedimientos de Inferencia Estadística los podemos clasificar atendiendo a tres criterios:

a) Respecto al objetivo de estudio. Si el objetivo es describir una variable o las relaciones entre un conjunto de variables se utilizan técnicas de muestreo. Cuando el objetivo es contrastar relaciones entre las variables y predecir sus valores futuros se utilizan técnicas de diseño experimental.

b) Respecto al método utilizado, nos encontramos con los métodos paramétricos y los no paramétricos. Los métodos paramétricos suponen que los datos provienen de una distribución que puede caracterizarse por un pequeño número de parámetros que se estiman a partir de los datos. Los métodos no paramétricos suponen únicamente aspectos muy generales de la distribución y tratan de estimar su forma o contrastar su estructura.

c) Respecto a la información considerada. Aquí distinguimos el enfoque clásico y el enfoque bayesiano. El enfoque clásico supone que los parámetros son cantidades fijas desconocidas sobre las que no se dispone de información inicial relevante. Por el contrario, el enfoque bayesiano considera a los parámetros del modelo como variables aleatorias y permite introducir información inicial sobre sus valores mediante una distribución de probabilidades a priori.

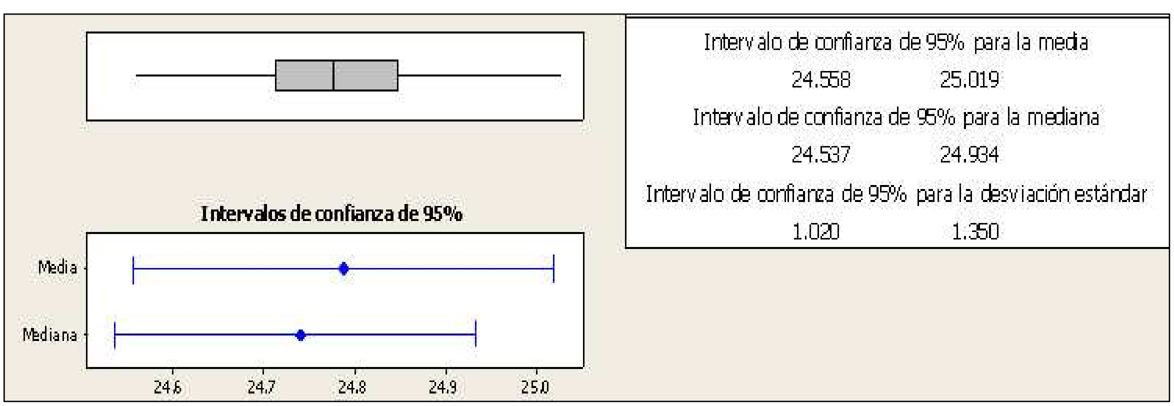

Figura 2. Ejemplo de Estadística Inferencial.

6.1.3 Pensamiento estadístico

El pensamiento estadístico es una filosofía de aprendizaje y acción basada en tres principios inevitables que son:

1. Todo trabajo ocurre en un sistema de procesos interconectados;
2. Variación hay en todos los procesos; y
3. Entender y reducir la variación son las claves del éxito.

El primer principio se relaciona con la primera parte del Sistema de Conocimiento Profundo, Apreciación de un sistema. El segundo y tercer principios, se unen a la segunda parte del Sistema de Conocimiento Profundo, Conocimiento sobre la variación.

En la ASQC (1996*b*) se plantea primero que el pensamiento estadístico es una filosofía, o sea es una forma de pensar, o un proceso de pensamiento, en lugar de cálculos que a algunos confunden. Ésta es la diferencia clave entre el pensamiento estadístico y las técnicas estadísticas, y cuya integración se ilustra en la Figura 3.

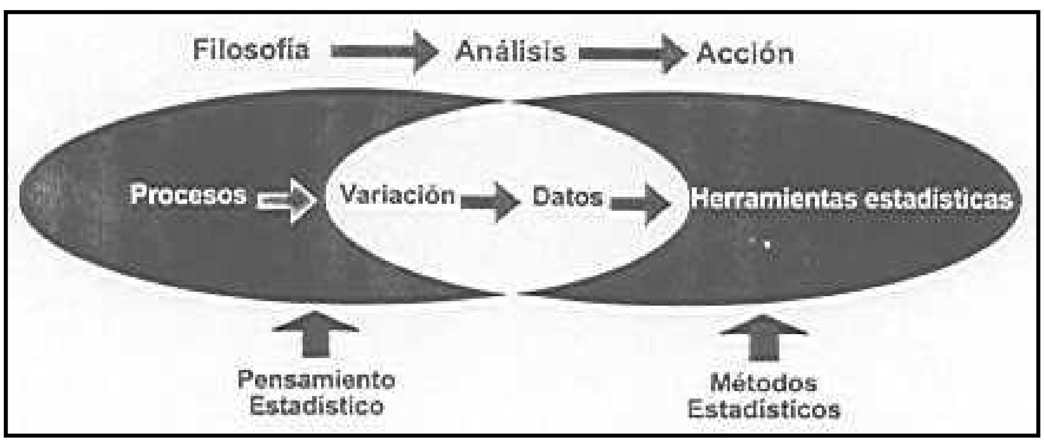

Figura 3. Relación entre pensamiento estadístico y métodos estadísticos

Es necesario examinar los principios fundamentales del pensamiento estadístico, porque para aplicar la filosofía no puede hablarse de pensamiento estadístico a menos que se incorporen los tres (ASQC, 1996*b*), tomando en cuenta los siguientes puntos:

• **Todo trabajo ocurre en un sistema de procesos interconectados**. Es decir, toda actividad es un proceso, el cual es un conjunto de actividades conectadas donde las entradas son transformadas en salidas para un propósito específico (ASQC, 1996b; López, 2004).

Este principio se refiere al entendimiento de toda actividad y tiene que ver con el aprendizaje, el mejoramiento potencial y las fuentes de variación que se mencionan en el segundo y tercer principios.

• **Variación hay en todo proceso**. Este principio proporciona el enfoque para el mejoramiento del trabajo. La variación es la clave. La variación resulta cuando dos o más cosas son diferentes y pudiéramos pensar que deberían ser exactamente iguales.

La presencia de la variación crea la necesidad del pensamiento estadístico (Dransfield*,* 1999). Si no hubiera variación, los procesos podrían ejecutarse mejor, los productos tendrían la calidad deseada, el servicio podría ser más consistente y los directivos podrían administrar mejor (Hoerl y Snee, 2002). Enfocarse en la variación es una estrategia clave para el logro de la mejora.

• **Entender y reducir la variación son las claves del éxito.** La calidad de los procesos determina inevitablemente la calidad de los productos que se obtienen, de ahí que la primera gran meta de todo proceso de mejora debe ser: controlar y reducir la variabilidad, de forma tal que los procesos sean estables, consistentes y predecibles.

Diversos negocios valoran tener diversidad de pensamiento y perspectiva porque es importante una diversidad de pensamiento entre los miembros del ejecutivo. El desarrollo de un nuevo producto requiere experimentación con nuevas ideas y acercamientos al mercado. La aplicación de la estadística inicia sobre la base de que hay variación en los datos, de lo contrario no se aplicaría. Por lo que la variación debe cuantificarse para comprender mejor las actividades o procesos y mejorar la toma de decisiones (ASQC, 1996*b*).

La estadística ha demostrado que las variaciones de un producto o proceso pueden medirse, con lo cual se puede determinar el comportamiento del proceso, o el lote de productos, o los tiempos de atención en el servicio.

Deming manifestó el mensaje de toda la vida, reduzca la variación. En la mayoría de las situaciones, reducir variación es calificativo de mejor. Reduciendo variación en el producto se satisface a los clientes, mientras que reduciendo la variación del proceso y de las entradas se reduce la de los productos, y comúnmente se reducen costos en los negocios en general (ASQC, 1996*b*).
En resumen, debiéramos siempre entender la variación y siempre tratar de reducirla, así el pensamiento estadístico sería pro activo.

Wild y Pfannkuch (1999) describieron en un sentido más amplio los procesos de pensamiento que tienen lugar en la solución de un problema estadístico. Desde la formulación del problema hasta las conclusiones. Ellos dicen: …*el pensamiento estadístico es la encarnación del sentido común...*, *Nosotros lo reconocemos en cuanto lo vemos*, o mejor *...su ausencia es a menudo claramente obvia* y *...para la mayoría de nosotros es producto de la experiencia*. Evidentemente es necesario pensar estadísticamente desde la identificación del problema hasta sus conclusiones.

Respecto a lo anterior puede pensarse que el pensamiento estadístico es un método que consiste en la identificación de los procesos más importantes (pueden ser *procesos clave o de apoyo*) en una empresa, la determinación de variables, la obtención de datos y finalmente el análisis de los mismos utilizando herramientas estadísticas, desde las más simples hasta quizás otras muy complejas.

6.2 Distribuciones de probabilidad

Una distribución de probabilidad indica toda la gama de valores que pueden representarse como resultado de un experimento si éste se llevase a cabo. Es decir, describe la probabilidad de que un evento se realice en el futuro, constituye una herramienta fundamental para la prospectiva, puesto que puede diseñar un escenario de acontecimientos futuros considerando las tendencias actuales de diversos fenómenos naturales.

Toda distribución de probabilidad es generada por una variable (por que puede tomar diferentes valores) aleatoria x (por que el valor tomado es totalmente al azar), y puede ser de dos tipos:

1. Variable Aleatoria Discreta (x)

Porque solo puede tomar valores enteros y un número finito de ellos. El cálculo de probabilidades es por:

*Uniforme discreta * Geométrica

*Binomial * Binomial Negativa

*Multinomial * Poisson

* Hipergeométrica

Distribución uniforme discreta

Describe el comportamiento de una variable discreta que puede tomar n valores distintos con la misma probabilidad cada uno de ellos. Un caso particular de esta distribución, ocurre cuando los valores son enteros consecutivos.

Esta distribución asigna igual probabilidad a todos los valores enteros entre el límite inferior y el límite superior que definen el recorrido de la variable. Si la variable puede tomar valores entre a y b, debe ocurrir que b sea mayor que a, y la variable toma los valores enteros empezando por a, a+1, a+2, etc. hasta el valor máximo b.

Por ejemplo, cuando se observa el número obtenido tras el lanzamiento de un dado perfecto, los valores posibles 4siguen una distribución uniforme discreta en {1, 2, 3, 4, 5, 6}, y la probabilidad de cada cara es 1/6.

Valores: x: a, a+1, a+2,..., b, números enteros

Parámetros:

a: mínimo, a entero

b: máximo, b entero con a < b

Distribución binomial

La distribución binomial es una distribución discreta muy importante que surge en muchas aplicaciones bioestadísticas.

Esta distribución aparece de forma natural al realizar repeticiones independientes de un experimento que tenga respuesta binaria, generalmente clasificada como “éxito” o “fracaso”.

Por ejemplo, esa respuesta puede ser el hábito de fumar (sí/no), si un paciente hospitalizado desarrolla o no una infección, o si un artículo de un lote es o no defectuoso. La variable discreta que cuenta el número de éxitos en n pruebas independientes de ese experimento, cada una de ellas con la misma probabilidad de “éxito” igual a p, sigue una distribución binomial de parámetros n y p.

Este modelo se aplica a poblaciones finitas de las que se toma elementos al azar con reemplazo, y también a poblaciones conceptualmente infinitas, como por ejemplo las piezas que produce una máquina, siempre que el proceso de producción sea estable (la proporción de piezas defectuosas se mantiene constante a largo plazo) y sin memoria (el resultado de cada pieza no depende de las anteriores).

Un ejemplo de variable binomial puede ser el número de pacientes ingresados en una unidad hospitalaria que desarrollan una infección nosocomial.

Un caso particular se tiene cuando n=1, que da lugar a la distribución de Bernoulli.
Valores: x: 0, 1, 2,..., n
Parámetros:
n: número de pruebas, n > 0 entero
p: probabilidad de éxito, 0 < p < 1

Distribución multinomial

Generaliza la distribución binomial al caso en que la población se divida en m>2 grupos mutuamente exclusivos y exhaustivos.

Se supone un proceso estable y sin memoria que genera elementos que pueden clasificarse en m clases distintas. Supóngase que se toma una muestra de n elementos y se definen m variables aleatorias X_i=número de elementos de la clase i (i=1,…., m) entonces el vector de m-variables es una variable aleatoria m-dimensional que sigue una distribución multinomial de parámetros n, p1,…, pm, donde p_i= (i=1,…, m) es la probabilidad de la clase i.

Por ejemplo: de acuerdo con la teoría de la genética, un cierto cruce de conejillo de indias resultará en una descendencia roja, negra y blanca en la relación 8:4:4. Si se tienen 8 descendientes, el vector de variables (x_1, x_2, x_3) donde:

x_1 =N° de descendientes rojos
x_2 =N° de descendientes negros
x_3 =N° de descendientes blancos

Sigue una distribución multinomial con parámetros n=8; p_1=8/16 =0.5; p_2=4/16 =0.25 y p_3=4/16 =0.25.

Una situación muy común en la práctica se da cuando se conoce el tamaño de muestra n y se quiere estimar las probabilidades p_i a partir de los valores observados. Pero también hay situaciones en las que se debe estimar el tamaño de muestra n, además de las probabilidades $p_{i.}$

Valores: xi=0, 1,2,…, (i=1,…, m)

Parámetros:
n: número de pruebas, *n*>0 entero
m: número de clases, *m*>0 entero
p_i: probabilidad de la clase i, 0< p_i <1 (i=1,…, m), donde $\sum_{i=1}^{m} p1 = 1$

Distribución hipergeométrica

La distribución hipergeométrica suele aparecer en procesos muéstrales sin reemplazo, en los que se investiga la presencia o ausencia de cierta característica. Piénsese, por ejemplo, en un procedimiento de control de calidad en una empresa farmacéutica, durante el cual se extraen muestras de las cápsulas fabricadas y se someten a análisis para determinar su composición.

Durante las pruebas, las cápsulas son destruidas y no pueden ser devueltas al lote del que provienen. En esta situación, la variable que cuenta el número de cápsulas que no cumplen los criterios de calidad establecidos sigue una distribución hipergeométrica. Por tanto, esta distribución es la equivalente a la binomial, pero cuando el muestreo se hace sin reemplazo.

Esta distribución se puede ilustrar del modo siguiente: se tiene una población finita con N elementos, de los cuales R tienen una determinada característica que se llama "éxito" (diabetes, obesidad, hábito de fumar, etc.). El número de "éxitos" en una muestra aleatoria de tamaño n, extraída sin reemplazo de la población, es una variable aleatoria con distribución hipergeométrica de parámetros N, R y n.

Cuando el tamaño de la población es grande, los muestreos con y sin reemplazo son equivalentes, por lo que la distribución hipergeométrica se aproxima en tal caso a la binomial.

Valores: x: max {0, n-(N-R)},..., min {R, n}, donde max {0, n-(N-R)} indica el valor máximo entre 0 y n-(N-R) y min {R, n} indica el valor mínimo entre R y n.

Parámetros:
N: tamaño de la población, N>0 entero
R: número de éxitos en la población, R≥0 entero
n: número de pruebas, n>0 entero

Distribución geométrica

Supóngase, que se efectúa repetidamente un experimento o prueba, que las repeticiones son independientes y que se está interesado en la ocurrencia o no de un suceso al que se refiere como "éxito", siendo la probabilidad de este suceso p.

La distribución geométrica permite calcular la probabilidad de que tenga que realizarse un número k de repeticiones hasta obtener un éxito por primera vez. Así pues, se diferencia de la distribución binomial en que el número de repeticiones no está predeterminado, sino que es la variable aleatoria que se mide y, por otra parte, el conjunto de valores posibles de la variable es ilimitado.

Para ilustrar el empleo de esta distribución, se supone que cierto medicamento opera exitosamente ante la enfermedad para la cual fue concebido en el 80% de los casos a los que se aplica; la variable aleatoria "intentos fallidos en la aplicación del medicamento antes del primer éxito" sigue una distribución geométrica de parámetro p=0,8. Otro ejemplo de variable geométrica es el número de hijos hasta el nacimiento de la primera niña.

La distribución geométrica se utiliza en la distribución de tiempos de espera, de manera que si los ensayos se realizan a intervalos regulares de tiempo, esta variable aleatoria proporciona el tiempo transcurrido hasta el primer éxito.

Esta distribución presenta la denominada "propiedad de Harkov" o de falta de memoria, que implica que la probabilidad de tener que esperar un tiempo t no depende del tiempo que ya haya transcurrido.
Valores: x: 0, 1, 2,...

Parámetros:
p: probabilidad de éxito, $0<p<1$

Distribución binomial negativa

Una generalización obvia de la distribución geométrica aparece si se supone que un experimento se continúa hasta que un determinado suceso, de probabilidad p, ocurre por r-ésima vez. La variable aleatoria que proporciona la probabilidad de que se produzcan k fracasos antes de obtener el r-ésimo éxito sigue una distribución binomial negativa de parámetros r y p, BN(r,p). La distribución geométrica corresponde al caso particular en que r=1. Un ejemplo es el número de lanzamientos fallidos de un dado antes de obtener un 6 en tres ocasiones, que sigue una BN (3,1/6).

En el caso de que los sucesos ocurran a intervalos regulares de tiempo, esta variable proporciona el tiempo total para que ocurran r éxitos, por lo que también se denomina "distribución binomial de tiempo de espera".

La distribución binomial negativa fue propuesta, originalmente, como una alternativa a la distribución de Poisson para modelar el número de ocurrencias de un suceso cuando los datos presentan lo que se conoce como variación extra-Poisson o sobredispersión. En estas situaciones, la varianza es mayor que la media, por lo que se incumple la propiedad que caracteriza a una distribución de Poisson, según la cual la media es igual a la varianza. La primera aplicación en bioestadística la realizó Student (William S. Gosset) a principios de siglo cuando propuso esta distribución para modelar el número de glóbulos rojos en una gota de sangre. En este caso, la variabilidad extra se debe al hecho de que esas células no están uniformemente distribuidas en la gota, es decir, la tasa de intensidad no es homogénea.

Por ejemplo, la distribución binomial negativa es más adecuada que la de Poisson para modelar el número de accidentes laborales ocurridos en un determinado lapso. La distribución de Poisson asume que todos los individuos tienen la misma probabilidad de sufrir un accidente y que ésta permanece constante durante el período de estudio; sin embargo, es más plausible la hipótesis de que los individuos tienen probabilidades constantes en el tiempo, pero que varían de unos sujetos a otros; esto es lo que se conoce en la literatura como la propensión a los accidentes. Esta hipótesis se traduce en una distribución de Poisson mixta, o de efectos aleatorios, en la que se supone que:

Las probabilidades varían entre individuos de acuerdo a una distribución gamma y esto resulta en una distribución binomial negativa para el número de accidentes.

Valores: x: 0, 1, 2,...
Parámetros:
p: probabilidad de éxito, $0<p<1$
r: número de éxitos, $r\geq 0$

Distribución Poisson

La distribución de Poisson se puede utilizar como una aproximación de la binomial, si el número de pruebas *n* es grande, pero la probabilidad de éxito *p* es pequeña; una regla es que la aproximación Poisson-binomial es "buena" si $n \geq 20$ y $p \leq 0{,}05$ y "muy buena" si $n \geq 100$ y $p \leq 0{,}01$.

La distribución de Poisson también surge cuando un evento o suceso "raro" ocurre aleatoriamente en el espacio o el tiempo. La variable asociada es el número de ocurrencias del evento en un intervalo o espacio continuo, por tanto, es una variable aleatoria discreta que toma valores enteros de 0 en adelante (0, 1, 2,...). Así, el número de pacientes que llegan a un consultorio en un lapso dado, el número de llamadas que recibe un servicio de atención a urgencias durante 1 hora, el número de células anormales en una superficie histológica o el número de glóbulos blancos en un milímetro cúbico de sangre son ejemplos de variables que siguen una distribución de Poisson. En general, es una distribución muy utilizada en diversas áreas de la investigación médica y, en particular, en epidemiología.

El concepto de evento "raro" o poco frecuente debe ser entendido en el sentido de que la probabilidad de observar k eventos decrece rápidamente a medida que k aumenta. Supóngase, por ejemplo, que el número de reacciones adversas tras la administración de un fármaco sigue una distribución de Poisson de media *lambda* = 2.

Si se administra este fármaco a 1.000 individuos, la probabilidad de que se produzca una reacción adversa (k=1) es 0,27; los valores de dicha probabilidad para k=2, 3, 4, 5, 6 reacciones, respectivamente, son: 0,27; 0,18; 0,09; 0,03 y 0,01. Para k=10 o mayor, la probabilidad es virtualmente 0.

El rápido descenso de la probabilidad de que se produzcan k reacciones adversas a medida que k aumenta puede observarse claramente en el gráfico de la función de densidad:

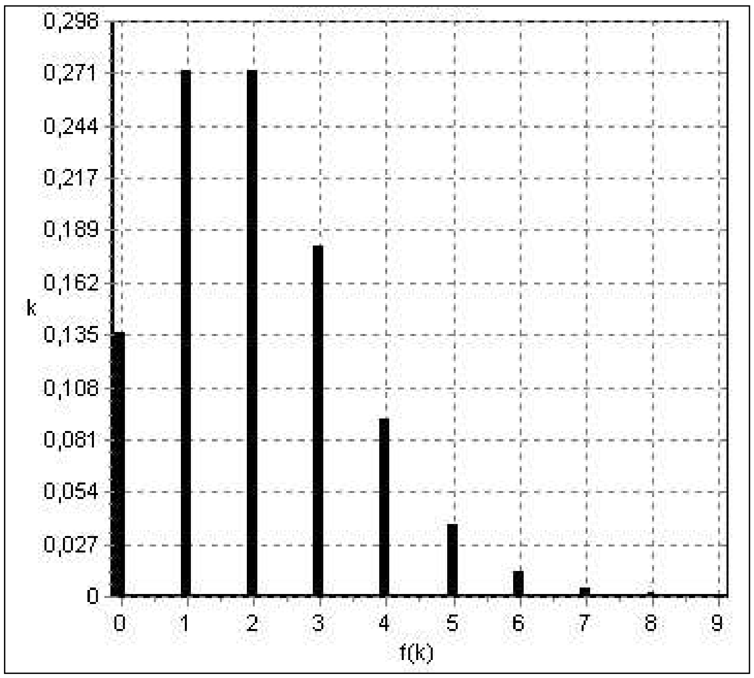

Figura 4. Función de densidad

Para que una variable recuento siga una distribución de Poisson deben cumplirse varias condiciones:

1. En un intervalo muy pequeño (p. e. de un milisegundo) la probabilidad de que ocurra un evento es proporcional al tamaño del intervalo.
2. La probabilidad de que ocurran dos o más eventos en un intervalo muy pequeño están reducida que, a efectos prácticos, se puede considerar nula.
3. El número de ocurrencias en un intervalo pequeño no depende de lo que ocurra en cualquier otro intervalo pequeño que no se solape con aquél.

Estas propiedades pueden resumirse en que el proceso que genera una distribución de Poisson es estable (produce, a largo plazo, un número medio de sucesos constante por unidad de observación) y no tiene memoria (conocer el número de sucesos en un intervalo no ayuda a predecir el número de sucesos en el siguiente).

El parámetro de la distribución, *lambda*, representa el número promedio de eventos esperados por unidad de tiempo o de espacio, por lo que también se suele hablar de *lambda* como “la tasa de ocurrencia” del fenómeno que se observa.

A veces se usan variables de Poisson con "intervalos" que no son espaciales ni temporales, sino de otro tipo. Por ejemplo, para medir la frecuencia de una enfermedad se puede contar, en un período dado, el número de enfermos en cierta población, dividida en "intervalos" de, por ejemplo, 10.000 habitantes. Al número de personas enfermas en una población de tamaño prefijado, en un instante dado, se le denomina prevalencia de la enfermedad en ese instante y es una variable que sigue una distribución de Poisson. Otra medida para la frecuencia de una enfermedad es la incidencia, que es el número de personas que enferman en una población en un periodo determinado.

En este caso, el intervalo es de personas tiempo, habitualmente personas-año, y es también una variable con distribución de Poisson. Habitualmente, ambas medidas se expresan para intervalos de tamaño unidad o, dicho de otro modo, en lugar de la variable número de enfermos, se usa el parámetro *lambda* (el riesgo, en el caso de la prevalencia, y la densidad de incidencia, en el de incidencia).

La distribución de Poisson tiene iguales la media y la varianza. Si la variación de los casos observados en una población excede a la variación esperada por la Poisson, se está ante la presencia de un problema conocido como sobre dispersión y, en tal caso, la distribución binomial negativa es más adecuada.

Valores: x: 0, 1, 2,...
Parámetros:
Lambda: media de la distribución, *lambda* > 0

2. Variable aleatoria continua (x)

Porque puede tomar tanto valores enteros como fraccionarios y un número infinito de ellos dentro de un mismo intervalo. El cual el cálculo de probabilidades es por:

*Uniforme	* Gamma
*Normal	* Exponencial
*Log normal	* Ji-cuadrado
*Logística	* t de Student
*Beta	* F de Snedecor

Distribución uniforme

La distribución uniforme es útil para describir una variable aleatoria con probabilidad constante sobre el intervalo [a, b] en el que está definida. Esta distribución presenta una peculiaridad importante: la probabilidad de un suceso dependerá exclusivamente de la amplitud del intervalo considerado y no de su posición en el campo de variación de la variable.

Cualquiera sea la distribución F de cierta variable X, la variable transformada Y=F(X) sigue una distribución uniforme en el intervalo [0,1]. Esta propiedad es fundamental por ser la base para la generación de números aleatorios de cualquier distribución en las técnicas de simulación.

Campo de variación: $a \leq x \leq b$

Parámetros:

a: mínimo del recorrido

b: máximo del recorrido

Distribución normal

La distribución normal es, sin duda, la distribución de probabilidad más importante del Cálculo de probabilidades y de la Estadística. Fue descubierta por De Moivre (1773), como aproximación de la distribución binomial.

De todas formas, la importancia de la distribución normal queda totalmente consolidada por ser la distribución límite de numerosas variables aleatorias, discretas y continuas, como se demuestra a través de los teoremas centrales del límite.

Las consecuencias de estos teoremas implican la casi universal presencia de la distribución normal en todos los campos de las ciencias empíricas: biología, medicina, psicología, física, economía, etc. En particular, muchas medidas de datos continuos en medicina y en biología (talla, presión arterial, etc.) se aproximan a la distribución normal.

Junto a lo anterior, no es menos importante el interés que supone la simplicidad de sus características y de que de ella derivan, entre otras, tres distribuciones (Ji-cuadrado, t y F) que se mencionarán más adelante, de importancia clave en el campo de la contrastación de hipótesis estadísticas.

La distribución normal queda totalmente definida mediante dos parámetros: la media (*Mu*) y la desviación estándar (*Sigma*).

Campo de variación: $-\infty < x < \infty$
Parámetros:
Mu: media de la distribución, $-\infty <$ *Mu* $< \infty$
Sigma: desviación estándar de la distribución, *Sigma* > 0

Distribución lognormal

La variable resultante al aplicar la función exponencial a una variable que se distribuye normal con media *Mu* y desviación estándar *Sigma*, sigue una distribución lognormal con parámetros *Mu* (escala) y *Sigma* (forma). Dicho de otro modo, si una variable X se distribuye normalmente, la variable *ln* X, sigue una distribución lognormal.

La distribución lognormal es útil para modelar datos de numerosos estudios médicos tales como el período de incubación de una enfermedad, los títulos de anticuerpo a un virus, el tiempo de supervivencia en pacientes con cáncer o SIDA, el tiempo hasta la seroconversión de VIH+, etc.

Campo de variación: $0 < x < \infty$
Parámetros:
Mu: parámetro de escala, $-\infty < Mu < \infty$
Sigma: parámetro de forma, *Sigma* > 0

Distribución logística

La distribución logística se utiliza en el estudio del crecimiento temporal de variables, en particular, demográficas. En biología se ha aplicado, por ejemplo, para modelar el crecimiento de células de levadura, y para representar curvas de dosis-respuesta en bioensayos.

La más conocida y generalizada aplicación de la distribución logística en Ciencias de la Salud se fundamenta en la siguiente propiedad: si U es una variable uniformemente distribuida en el intervalo [0,1], entonces la variable X= ln $(\frac{U}{1-U})$ sigue una distribución logística. Esta transformación, denominada *log it*, se utiliza para modelar datos de respuesta binaria, especialmente en el contexto de la regresión logística.

Campo de variación: $- \infty < x < \infty$

Parámetros:

a: parámetro de posición, $-\infty < a < \infty$

b: parámetro de escala, $b > 0$

Distribución beta

La distribución beta es posible para una variable aleatoria continua que toma valores en el intervalo [0,1], lo que la hace muy apropiada para modelar proporciones. En la inferencia bayesiana, por ejemplo, es muy utilizada como distribución a priori cuando las observaciones tienen una distribución binomial.

Uno de los principales recurso de esta distribución es el ajuste a una gran variedad de distribuciones empíricas, pues adopta formas muy diversas dependiendo de cuáles sean los valores de los parámetros de forma p y q, mediante los que vienen definida la distribución.

Un caso particular de la distribución beta es la distribución uniforme en [0,1], que se corresponde con una beta de parámetros p=1 y q=1, denotada Beta (1,1).

Campo de variación: $0 \leq x \leq 1$

Parámetros:

p: parámetro de forma, $p > 0$

q: parámetro de forma, $q > 0$

Distribución gamma

La distribución gamma se puede caracterizar del modo siguiente: si está interesado en la ocurrencia de un evento generado por un proceso de Poisson de media lambda, la variable que mide el tiempo transcurrido hasta obtener n ocurrencias del evento sigue una distribución gamma con parámetros a=n*lambda (escala) y p=n (forma). Se denota Gamma (a, p).

Por ejemplo, la distribución gamma aparece cuando se realiza el estudio de la duración de elementos físicos (tiempo de vida).
Esta distribución presenta como propiedad interesante la “falta de memoria”. Por esta razón, es muy utilizada en las teorías de fiabilidad, mantenimiento y fenómenos de espera (por ejemplo en una consulta médica “tiempo que transcurre hasta la llegada del segundo paciente”).

Campo de variación: $0 < x < \infty$

Parámetros:

a: parámetro de escala, $a > 0$

p: parámetro de forma, $p > 0$

Distribución exponencial

La distribución exponencial es el equivalente continuo de la distribución geométrica discreta. Esta ley de distribución describe procesos en los que interesa saber el tiempo hasta que ocurre determinado evento; en particular, se utiliza para modelar tiempos de supervivencia. Un ejemplo es el tiempo que tarda una partícula radiactiva en desintegrarse. El conocimiento de la ley que sigue este evento se utiliza, por ejemplo, para la datación de fósiles o cualquier materia orgánica mediante la técnica del carbono 14.

Una característica importante de esta distribución es la propiedad conocida como “falta de memoria”. Esto significa, por ejemplo, que la probabilidad de que un individuo de edad *t* sobreviva *x* años más, hasta la edad *x+t*, es la misma que tiene un recién nacido de sobrevivir hasta la edad *x*. Dicho de manera más general, el tiempo transcurrido desde cualquier instante dado *t*0 hasta que ocurre el evento, no depende de lo que haya ocurrido antes del instante *t0*.

La distribución exponencial se puede caracterizar como la distribución del tiempo entre sucesos consecutivos generados por un proceso de Poisson; por ejemplo, el tiempo que transcurre entre dos heridas graves sufridas por una persona. La media de la distribución de Poisson, *lambda*, que representa la tasa de ocurrencia del evento por unidad de tiempo, es el parámetro de la distribución exponencial, y su inversa es el valor medio de la distribución. También se puede ver como un caso particular de la distribución gamma(*a*, *p*), con a=*lambda* y *p*=1.

El uso de la distribución exponencial ha sido limitado en bioestadística, debido a la propiedad de falta de memoria que la hace demasiado restrictiva para la mayoría de los problemas.

Campo de variación: $0 < x < \infty$
Parámetros:
Lambda: tasa, *lambda* > 0

Distribución Ji-cuadrado

Un caso especial, muy importante, de la distribución Gamma se obtiene cuando $a=1/2$ y $p=n/2$. La distribución resultante se conoce con el nombre de Ji-cuadrado con *n* grados de libertad. Es la distribución que sigue la suma de los cuadrados de n variables independientes N (0,1).

La Ji-cuadrado es una distribución fundamental en inferencia estadística y en los test estadísticos de bondad de ajuste. Se emplea, entre muchas otras aplicaciones, para determinar los límites de confianza de la varianza de una población normal, para contrastar la hipótesis de homogeneidad o de independencia en una tabla de contingencia y para pruebas de bondad de ajuste. La distribución Ji-cuadrado queda totalmente definida mediante sus grados de libertad *n.*

Campo de variación: $0 \leq x < \infty$
Parámetros:
n: grados de libertad, $n>0$

Distribución t de Student

La distribución t de Student se construye como un cociente entre una normal y la raíz de una Ji-cuadrado independientes. Esta distribución desempeña un papel importante en la inferencia estadística asociada a la teoría de muestras pequeñas. Se usa habitualmente en el contraste de hipótesis para la media de una población, o para comparar las medias de dos poblaciones, y viene definida por sus grados de libertad *n.*

A medida que aumentan los grados de libertad, la distribución t de Student se aproxima a una normal de media 0 y varianza 1 (normal estándar).

Campo de variación: $-\infty < x < \infty$
Parámetros:
n: grados de libertad, $n>0$

Distribución F de Snedecor

Otra de las distribuciones importantes asociadas a la normal es la que se define como el cociente de dos variables con distribución Ji-cuadrado divididas por sus respectivos grados de libertad, *n* y *m*. En este caso la variable aleatoria sigue una distribución F de Snedecor de parámetros *n* y *m*. Hay muchas aplicaciones de la F en estadística y, en particular, tiene un papel importante en las técnicas del análisis de la varianza y del diseño de experimentos.

Campo de variación: $0 \leq x < \infty$
Parámetros:
n: grados de libertad del numerador, $n>0$
m: grados de libertad del denominador, $m>0$

6.3 Pruebas de hipótesis

Una hipótesis estadística es una proposición o supuesto sobre los parámetros de una o más poblaciones, el cual el estudio de la inferencia estadística nos muestra la forma de usar la prueba de hipótesis para determinar si una afirmación acerca del valor de un parámetro poblacional debe o no ser rechazado (David R. Anderson, Dennis J. Sweeney).

Basado en esto, se pueden denotar etapas en la construcción de pruebas de hipótesis.

Etapa 1. Planear la hipótesis nula y la hipótesis alternativa. La hipótesis nula (H0) es el valor hipotético del parámetro que se compra con el resultado muestral resulta muy poco probable cuando la hipótesis es cierta.

Etapa 2. Especificar el nivel de significancia que se va a utilizar. El nivel de significancia del 5%, entonces se rechaza la hipótesis nula solamente si el resultado muestral es tan diferente del valor hipotético que una diferencia de esa magnitud o mayor, pudiera ocurrir aleatoria mente con una probabilidad de 1.05 o menos.

Etapa 3. Elegir la estadística de prueba. La estadística de prueba puede ser la estadística muestral (el estimador no segado del parámetro que se prueba) o una versión transformada de esa estadística muestral. Por ejemplo, para probar el valor hipotético de una media poblacional, se toma la media de una muestra aleatoria de esa distribución normal, entonces es común que se transforme la media en un valor z el cual, a su vez, sirve como estadística de prueba.

Etapa 4. Establecer el valor o valores críticos de la estadística de prueba. Habiendo especificado la hipótesis nula, el nivel de significancia y la estadística de prueba que se van a utilizar, se produce a establecer el o los valores críticos de estadística de prueba. Puede haber uno o más de esos valores, dependiendo de si se va a realizar una prueba de uno o dos extremos.

Etapa 5. Determinar el valor real de la estadística de prueba. Por ejemplo, al probar un valor hipotético de la media poblacional, se toma una muestra aleatoria y se determina el valor de la media muestral. Si el valor crítico que se establece es un valor de z, entonces se transforma la media muestral en un valor de z.

Etapa 6. Tomar la decisión. Se compara el valor observado de la estadística muestral con el valor (o valores) críticos de la estadística de prueba. Después se acepta o se rechaza la hipótesis nula. Si se rechaza ésta, se acepta la alternativa; a su vez, esta decisión tendrá efecto sobre otras decisiones de los administradores operativos, como por ejemplo, mantener o no un estándar de desempeño o cuál de dos estrategias de mercadotecnia utilizar.

6.3.1 Importancia

Cuando se hace una prueba de hipótesis se empieza por hacer una suposición tentativa acerca del parámetro poblacional. A esta suposición tentativa se le llama hipótesis nula y se denota por H_0. Después se define otra hipótesis alternativa, que dice lo contrario de lo que establece la hipótesis nula. La hipótesis alternativa se denota H_1 (David R. Anderson, Dennis J. Sweeney 2004).

La distribución apropiada de la prueba estadística se divide en dos regiones: una región de rechazo y una de no rechazo. Si la prueba estadística cae en esta última región no se puede rechazar la hipótesis nula y se llega a la conclusión de que el proceso funciona correctamente.

Al tomar la decisión con respecto a la hipótesis nula, se debe determinar el valor crítico en la distribución estadística que divide la región del rechazo (en la cual la hipótesis nula no se puede rechazar) de la región de rechazo. A hora bien el valor crítico depende del tamaño de la región de rechazo.

6.3.2 Hipótesis nula

La hipótesis nula es una afirmación o enunciado tentativo que se realiza acerca del valor de un parámetro poblacional. Por lo común en una afirmación de que el parámetro de población tiene valor especifico.
Una hipótesis nula es importante por varias razones:

- Es una hipótesis que se acepta o se rechaza según el resultado de la investigación.
- El hecho de contar con una hipótesis nula ayuda a determinar si existe una diferencia entre los grupos, si esta diferencia es significativa, y si no se debió al azar (Mark L. Berenson 2006).

La hipótesis nula se refiere siempre a un valor especificado del parámetro de población, no a una estadística de muestra. El planteamiento de la hipótesis nula siempre contiene un signo de igualdad con respecto al valor especificado del parámetro. (Torino, 1997)

6.3.3 Hipótesis alternativa

La hipótesis alternativa representa la conclusión a la que se llegaría si hubiera suficiente evidencia de la información de la muestra para decidir que es improbable que la hipótesis nula sea verdadera y, por tanto se rechaza (Mark L. Berenson 2006).

La hipótesis alternativa es cualquier hipótesis que difiera de la hipótesis nula. El planteamiento de la hipótesis alternativa nunca contiene un signo de igualdad con respecto al valor especificado del parámetro. (Torino, 1997).

6.4 Significancia estadística

La significancia estadística (Evans, Lindsay 2000) es el grado de riesgo que estamos dispuestos a asumir, de que rechazaremos una hipótesis nula cuando en realidad es cierta. Considerando que si se disminuye el valor de α también se disminuye el poder de la prueba. Es decir, si disminuimos la probabilidad de cometer un error de tipo I aumentamos simultáneamente la probabilidad de un error de tipo II, por lo que se trata de encontrar un punto de "equilibrio" entre ambas. Habitualmente se trabaja con un nivel de significación del 95% ($\alpha = 0.05$), por lo que el equilibrio hay que en encontrarlo finalmente entre el tamaño de la muestra que es posible estudiar y el poder que se quiere para el estudio.

La distribución de muestreo de la estadística de prueba se divide en dos regiones, una región de rechazo (conocida como región crítica) y una región de no rechazo (aceptación). Si la estadística de prueba cae dentro de la región de aceptación, no se puede rechazar la hipótesis nula.

La región de rechazo puede considerarse como el conjunto de valores de la estadística de prueba que no tienen posibilidad de presentarse si la hipótesis nula es verdadera. Por otro lado, estos valores no son tan improbables de presentarse si la hipótesis nula es falsa. El valor crítico separa la región de no rechazo de la de rechazo. (Torino, 1997).

Un error tipo I se presenta si la hipótesis nula es rechazada cuando de hecho es verdadera y debía ser aceptada. Un error tipo II se presenta si la hipótesis nula es aceptada cuando de hecho es falsa y debía ser rechazada. En cualquiera de los dos casos se comete un error al tomar una decisión equivocada.

Para que cualquier ensayo de hipótesis sea bueno, debe diseñarse de forma que minimice los errores de decisión. En la práctica un tipo de error puede tener más importancia que el otro, y así se tiene a conseguir poner una limitación al error de mayor importancia. La única forma de reducir ambos tipos de errores es incrementar el tamaño de la muestra, lo cual puede ser o no ser posible.

La probabilidad de cometer un error tipo I denotada con la letra griega alfa, se conoce como nivel de significación de la prueba estadística. Está bajo control directo del individuo que lleva a cabo la prueba. Ya que se ha especificado el valor de alfa, se conoce el tamaño de la región de rechazo, puesto que alfa es la probabilidad de un rechazo de la hipótesis nula.

El complemento $(1 - \alpha)$ de la probabilidad de cometer un error de tipo I se conoce como coeficiente de confianza. El coeficiente de confianza es la probabilidad de que la hipótesis nula no sea rechazada cuando de hecho es verdadera y debería ser aceptada.

La probabilidad de cometer un error de tipo II denotada con la letra griega beta, se conoce como nivel de riesgo del consumidor. La probabilidad de cometer un error de tipo II depende de la diferencia entre los valores supuesto y real del parámetro de la población. Como es más fácil encontrar diferencias grandes, si la diferencia entre la estadística de muestra y el correspondiente parámetro de población es grande, la probabilidad de cometer un error de tipo II, probablemente sea pequeña.

El estudio y las conclusiones que obtengamos para una población cualquiera, se habrán apoyado exclusivamente en el análisis de una parte de ésta. (Torino, 1997).

6.4.1 Error tipo I

El error tipo I, conocido también como erro tipo alfa (α), se comete cuando el investigador rechaza la hipótesis nula (H_0), siendo ésta verdadera en la población. Es equivalente a encontrar un resultado falso positivo ya que el investigador concluye que hay diferencia, cuando en realidad no existe. (Wilfredo Caballero Armas, 1975).

La significación estadística es por tanto una condición resultante del rechazo de una hipótesis nula mediante la aplicación de una prueba estadística de significación.

El nivel de significación es el riesgo o la probabilidad que voluntariamente asume el investigador de equivocarse al rechazar la hipótesis nula, cuando en realidad es cierta. Este riesgo se establece normalmente en 0.05 (95%) ó 0.01 (99%).

- Si $p < 0.05$ se considera significativo, en cuyo caso se rechaza la hipótesis nula.
- Si $p > 0.05$ se considera no significativo en cuyo caso no se rechaza la hipótesis nula.

6.4.2 Error tipo II

El error tipo II ó beta (β) se comete en la situación contraria: cuando el investigador no rechaza la hipótesis nula (H_0), siendo ésta falsa en la población. Es equivalente a un resultado falso negativo, ya que el investigador concluye que ha sido incapaz de encontrar una diferencia que existe en la realidad. (Wilfredo Caballero Armas).

El error tipo II o falso negativo, se comete cuando el investigador no rechaza la hipótesis nula siendo ésta falsa en la población. Es equivalente a la probabilidad de un resultado falso negativo, ya que el investigador llega a la conclusión de que ha sido incapaz de encontrar una diferencia que existe en la realidad. Se acepta en un estudio que el valor del error beta esté entre el 5 y el 20%.

Contrariamente al error tipo I, en la mayoría de los casos no es posible calcular la probabilidad del error tipo II. La razón de esto se encuentra en la manera en que se formulan las hipótesis en una prueba estadística.

Mientras que la hipótesis nula representa siempre una afirmación enérgica (como por ejemplo H_0: «Promedio μ = 0») la hipótesis alternativa, debido a que engloba todas las otras posibilidades, es generalmente de naturaleza global (por ejemplo H_1: «Promedio μ ≠ 0»). El gráfico ilustra la probabilidad del error tipo II (rojo) en dependencia del promedio μ desconocido.

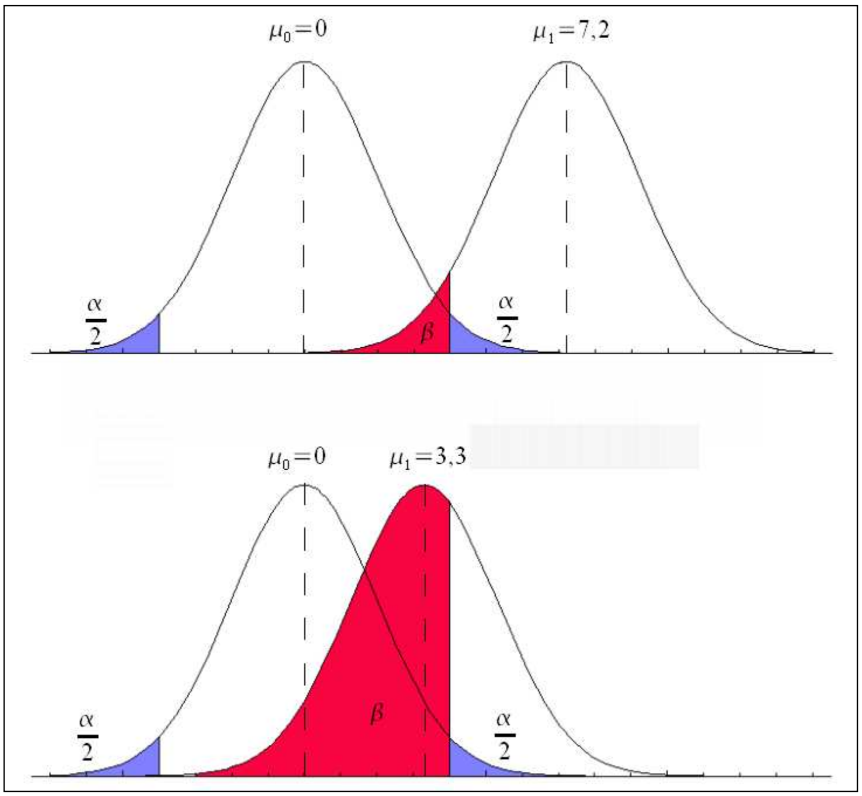

FIGURA 5. Representación de los valores posibles de la probabilidad de un error tipo II (rojo) en el ejemplo de un test de significancia estadística para el parámetro μ.

6.5 Potencia estadística

Potencia estadística es la probabilidad de que el efecto que se pretende encontrar sea detectado por el investigador, suponiendo que este efecto esté presente. La potencia de un análisis estadístico es en parte una función de la prueba estadística empleada. La potencia de una prueba se define como la probabilidad que existe de rechazar H0 cuando ésta es realmente falsa.

Matemáticamente, la potencia se representa como la complementaria del error tipo II o error β (1 – β), siendo β la probabilidad de aceptar H0 cuando ésta es falsa (y, por tanto, H1 es cierta).

En situaciones comparables, las pruebas paramétricas son más potentes que las no paramétricas. Esto se comprende fácilmente si consideramos que las pruebas paramétricas necesitan datos medidos al menos en escala de intervalos.

Estos datos pueden reducirse a una escala inferior, de orden, y emplear así una prueba no paramétrica. Pero se observa enseguida que se produce una pérdida de información, pues la escala ordinal considera sólo el orden de las observaciones, y no la cuantía de su separación.

Habitualmente, la mayoría de los investigadores desean rechazar las hipótesis de nulidad que plantearon al inicio de su proyecto. Por ello, lo que más les interesa es utilizar una prueba estadística con una potencia más bien alta.

Sin embargo, se les ha prestado poca atención a este punto y con frecuencia se realizan experimentos en los que existen muchas probabilidades de cometer errores de tipo II, sin saberlo el investigador, y se abandonan líneas de investigación que pueden ser prometedoras. El análisis de la potencia de la prueba aclararía este error.

La potencia de una prueba depende de tres factores fundamentales: α, n y γ (Welkowitz 1981). El nivel de significación α está fijado convencionalmente en 0.05 ó 0.01 para la mayoría de los casos, aunque el investigador pueda cambiarla, mientras que el tamaño de muestra n suele ser fijado de antemano por el investigador.

Desafortunadamente, γ suele quedar casi siempre fuera del control del investigador. La falta de control sobre γ suele ser el verdadero problema para el cálculo de la potencia de una prueba estadística.

6.6 Factores que inciden en la potencia estadística

El poder estadístico de un estudio depende de diferentes factores, como:

- *El tamaño del efecto a detectar*, es decir, la magnitud mínima de la diferencia o asociación entre los grupos que se considera clínicamente relevante. Cuanto mayor sea el tamaño del efecto que se desea detectar, mayor será la probabilidad de obtener hallazgos significativos y, por lo tanto, mayor será el poder estadístico.
- La *variabilidad* de la respuesta estudiada. Así, cuanto mayor sea la variabilidad en la respuesta, más difícil será detectar diferencias entre los grupos que se comparan y menor será el poder estadístico de la investigación. De ahí que sea recomendable estudiar grupos lo más homogéneos posibles.
- El *tamaño de la muestra* a estudiar. Cuanto mayor sea el tamaño muestral, mayor será la potencia estadística de un estudio. Es por ello que en los estudios con muestras muy grandes se detectan como significativas diferencias poco relevantes, y en los estudios con muestras menores es más fácil obtener resultados falsamente negativos.
- El *nivel de significación estadística*. Si se disminuye el valor de α también se disminuye el poder de la prueba. Es decir, si disminuimos la probabilidad de cometer un error de tipo I aumentamos simultáneamente la probabilidad de un error de tipo II, por lo que se trata de encontrar un punto de "equilibrio" entre ambas.

Habitualmente se trabaja con un nivel de significación del 95% ($\alpha = 0.05$), por lo que el equilibrio hay que en encontrarlo finalmente entre el tamaño de la muestra que es posible estudiar y el poder que se quiere para el estudio.

Los cuatro factores anteriores, junto con el poder estadístico, forman un sistema cerrado. De este modo, una vez fijados tres de ellos, el cuarto queda completamente determinado.

6.6.1 Tamaño de muestra

Cuando todos los demás factores se mantienen constantes, el investigador puede aumentar el poder del contraste H0 aumentando el tamaño de la muestra n. El motivo de ellos es que la exactitud de la mayoría de los valores estadísticos depende del tamaño de n, pues casi todos estos valores tienen alguna función de n en el denominador. Al aumentar n disminuye el error y se incrementa la potencia del test.

Por ello, en principio, cuanto mayor sea el tamaño de muestra escogido, mayor será la potencia del test estadístico y menor la probabilidad de error. Este fenómeno se puede apreciar esquemáticamente en la siguiente figura, que muestra cómo aumenta la potencia del test al aumentar el tamaño de la muestra. Estas muestras se tomaron de poblaciones normales con varianza σ.

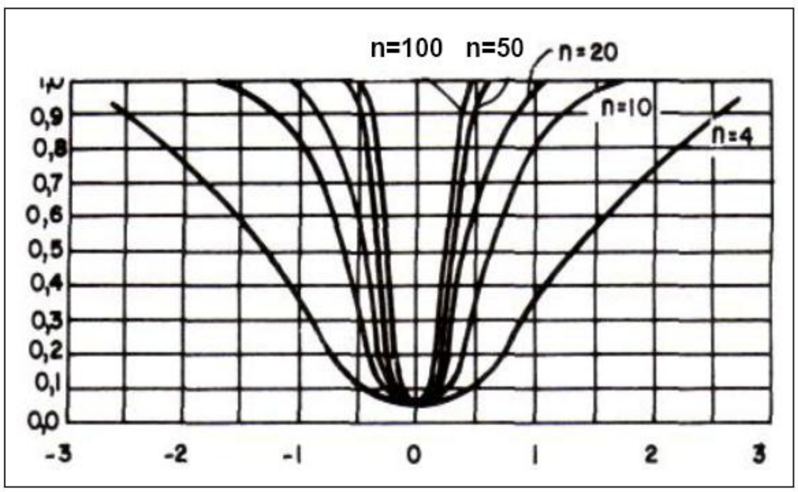

FIGURA 6. Curvas de potencia de una prueba de dos colas con α = 0,05 y distintos valores de n.

6.6.2 α (Probabilidad del error tipo I)

El nivel de significación α de una prueba estadística representa la probabilidad de rechazar la hipótesis nula H0 cuando ésta es verdadera, es decir, la probabilidad de que las diferencias halladas en nuestros resultados se deban puramente al azar. De hecho, y si los demás factores se mantienen constantes, cuanto más alto sea el nivel de significación, mas difícil será rechazar H0 cuando ésta es verdadera.

Sin embargo, todo aumento del nivel de significación se asocia a una disminución de la potencia del test estadístico utilizado y, por tanto, a un aumento de la probabilidad de error β. Por ejemplo, si se decide utilizar un nivel de significancia α = 0,05 en lugar de un nivel α = 0,01, la potencia aumenta. El problema es que la manipulación de α no suele ser una técnica eficaz porque, para unos datos determinados, dicha manipulación suele tener efectos opuestos en los dos tipos de errores α y β. Por ese motivo, la mayoría de los investigadores utilizan niveles de significación estándar fijados en 0,05 (o lo que es lo mismo, 5%) o en 0,01 (1%). Un ejemplo de este efecto de α sobre β y de β sobre α se muestra en la siguiente figura. En ella, se ilustra la interacción que se produce entre α y β, al analizar los datos de un diseño pretest-postest de grupo único.

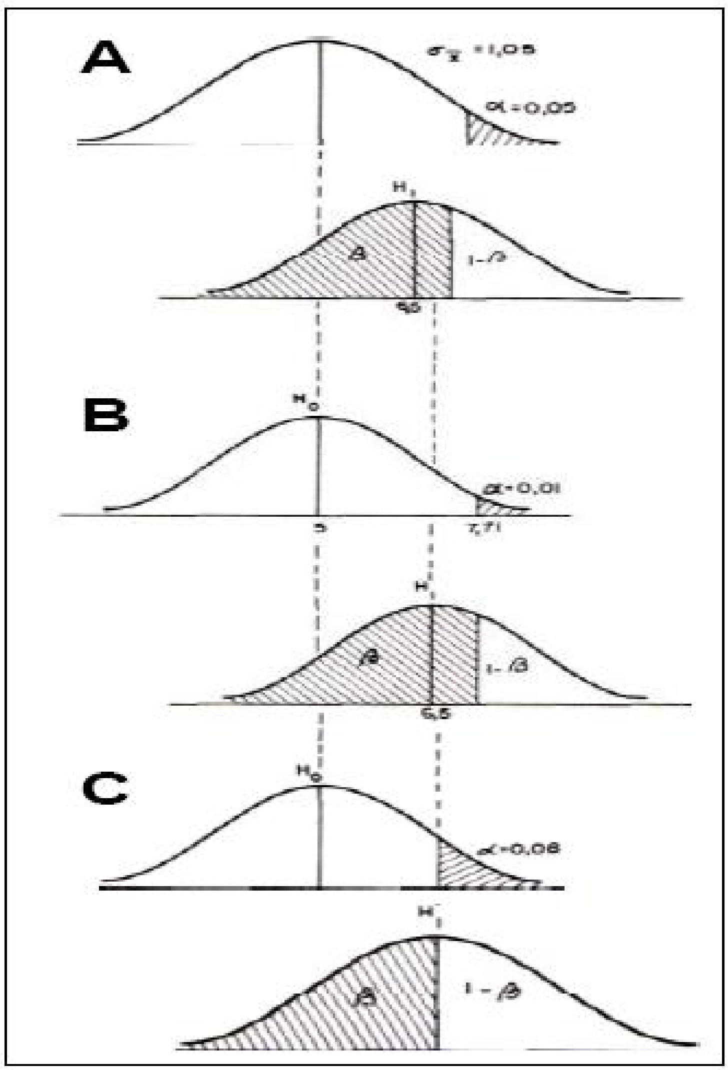

FIGURA 7. Poder de contraste de H0: μ=5 comparado con H1: μ=6.5 para distintos valores de α.

6.6.3 σ (Desviación estándar)

Para conocer con detalle un conjunto de datos, no basta con conocer las medidas de tendencia central, sino que necesitamos conocer también la desviación que representan los datos en su distribución respecto de la media aritmética de dicha distribución, con objeto de tener una visión de los mismos más acorde con la realidad al momento de describirlos e interpretarlos para la toma de decisiones.
La varianza representa la media aritmética de las desviaciones con respecto a la media que son elevadas al cuadrado. Si atendemos a la colección completa de datos (la población en su totalidad) obtenemos la varianza poblacional; y si por el contrario prestamos atención sólo a una muestra de la población, obtenemos en su lugar la varianza muestral.

La desviación estándar puede ser interpretada como una medida de incertidumbre. La desviación estándar de un grupo repetido de medidas nos da la precisión de éstas. Cuando se va a determinar si un grupo de medidas está de acuerdo con el modelo teórico, la desviación estándar de esas medidas es de vital importancia: si la media de las medidas está demasiado alejada de la predicción (con la distancia medida en desviaciones estándar), entonces consideramos que las medidas contradicen la teoría. Esto es coherente, ya que las mediciones caen fuera del rango de valores en el cual sería razonable esperar que ocurrieran si el modelo teórico fuera correcto.

6.6.4 Efecto de la magnitud de la población.

Para la determinación del tamaño de la muestra hay que decidir previamente que potencia se desea. Welkowitz recomienda el de 0,80, que fija en 0,20 la probabilidad de un error tipo II. El sugerir una probabilidad mayor para los errores tipo II que para los de tipo I se debe a que en la mayoría de los problemas que se investigan resultan menos perjudiciales los falsos negativos (rechazar H0 cuando es falsa o error tipo II) que los falsos positivos (rechazar H0 cuando es verdadera o error tipo I).

Por otra parte, si se fija convencionalmente una potencia de 0,95 ó de 0,99 la muestra resultante no estará al alcance de muchos investigadores (Jiménez Fernández, 2000)

6.7 Cálculo de la potencia estadística.

Al momento de diseñar una investigación, es importante determinar si dicho estudio alcanzará una precisión suficiente. Generalmente, se suele trabajar con un poder en torno al 80% o al 90%. Con frecuencia, sin embargo, las condiciones en las que se lleva a cabo una investigación son diferentes de las que se habían previsto en un principio. En consecuencia, y a la vista de hallazgos no significativos, es recomendable evaluar de nuevo a posteriori su potencia con el fin de discernir si el estudio carece del poder necesario para detectar una diferencia relevante o bien si realmente puede no existir tal diferencia.

6.7.1 ANOVA

El análisis de la varianza (o ANOVA: Analysis of Variance) es un método para comparar medias, que es necesario porque cuando se quiere comparar más de dos medias es incorrecto utilizar repetidamente el contraste basado en la t de Student. Por dos motivos:

En primer lugar, y como se realizarían simultánea e independientemente varios contrastes de hipótesis, la probabilidad de encontrar alguno significativo por azar aumentaría. En cada contraste se rechaza la H0 si la t supera el nivel crítico, para lo que, en la hipótesis nula, hay una probabilidadα.

Si se realizan m contrastes independientes, la probabilidad de que, en la hipótesis nula, ningún estadístico supere el valor crítico es (1 - α) m, por lo tanto, la probabilidad de que alguno lo supere es 1 - (1 - α) m, que para valores de α próximos a 0 es aproximadamente igual a αm.

Una primera solución, denominada método de Bonferroni, consiste en bajar el valor deα, usando en su lugar α/m, aunque resulta un método muy conservador.

Por otro lado, en cada comparación la hipótesis nula es que las dos muestras provienen de la misma población, por lo tanto, cuando se hayan realizado todas las comparaciones, la hipótesis nula es que todas las muestras provienen de la misma población y, sin embargo, para cada comparación, la estimación de la varianza necesaria para el contraste es distinta, pues se ha hecho en base a muestras distintas.

El método que resuelve ambos problemas es el ANOVA, aunque es algo más que esto: es un método que permite comparar varias medias en diversas situaciones; muy ligado, por tanto, al diseño de experimentos y, de alguna manera, es la base del análisis multivariable.

6.7.2 Prueba de normalidad

La prueba de normalidad es una hipótesis de cierta muestra para determinar si la población de la cual extrajo su muestra es normal. Muchos procedimientos estadísticos dependen de la normalidad de la población, de modo que recurrir a una prueba de normalidad para determinar si se rechaza este supuesto pudiera ser un paso importante en su análisis. La hipótesis nula para una prueba de normalidad establece que la población es normal. La hipótesis alternativa establece que la población es no normal. Para determinar si los datos de su muestra provienen de una población no normal, usted puede elegir entre cuatro pruebas.

Técnica gráfica

Se puede evaluar la normalidad de una población con una gráfica de probabilidad normal, la cual genera de manera gráfica valores de datos ordenados en comparación con los valores que se espera sean cercanos a los primeros, si la población de la muestra está normalmente distribuida. Si la población es normal, los puntos de la gráfica conformarán una línea aproximadamente derecha.

Datos normales Datos no normales

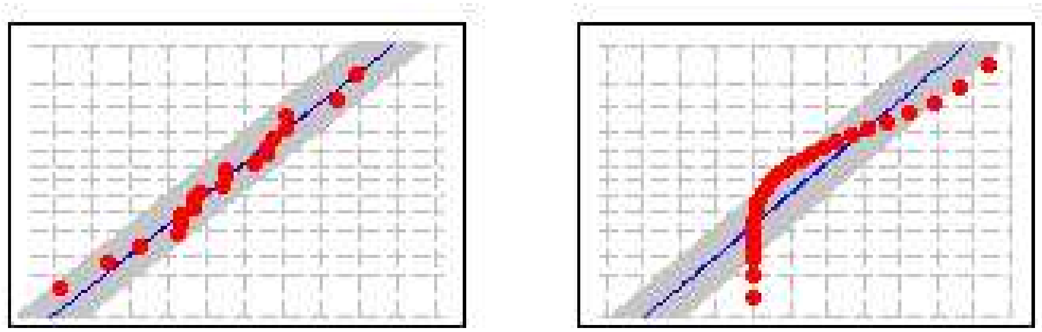

FIGURA 8. Ejemplo de gráficas de probabilidad de datos.

Prueba de Anderson-Darling

Esta prueba compara la función de distribución acumulada empírica de los datos de una muestra con la distribución esperada si los datos son normales. Si esta diferencia observada es suficientemente grande, la prueba rechazará la hipótesis nula de normalidad en la población.

Prueba de normalidad de Ryan-Joiner

Esta prueba evalúa la normalidad calculando la correlación entre un grupo de datos y las puntuaciones normales de estos. Si el coeficiente de correlación se encuentra cerca de 1, es probable que la población sea normal.

La estadística de Ryan-Joiner evalúa la solidez de esta correlación; si se encuentra por debajo del valor crítico apropiado, se rechazará la hipótesis nula de normalidad en la población.

6.7.3 Z de 1 muestra

Prueba si una media de población normal es igual a un valor objetivo. Esta prueba requiere que se conozca la varianza de población, si no la conoce, debe utilizarse la prueba t de una muestra en su lugar. La prueba Z de 1 muestra utiliza la hipótesis nula que establece que la media de la población (m) es igual a un valor hipotético (H_0: $m = m_0$) y lo prueba en una hipótesis alternativa que puede tener cola hacia la izquierda, ($m < m_0$), hacia la derecha ($m > m_0$) o dos colas ($m \neq m_0$).

Por ejemplo, los resultados de las pruebas de IQ y otras pruebas estandarizadas con frecuencia tienen una varianza de población conocida.

Supongamos que millones de estudiantes toman un tipo de prueba estandarizada todos los años. En dicha prueba, la puntuación oscila entre 0 y 100. Un grupo de 40 estudiantes toma la prueba y obtiene una media de promedio de 85. Usted desea determinar si esta puntuación es estadísticamente diferente del promedio nacional de 80. Los resultados históricos indican que la varianza de población será de 100. Debido a que se conoce la varianza de la población, la prueba Z de 1 muestra es apropiada.

6.7.4 T de 1 muestra

Compara la media de la muestra con un valor hipotético. Por ejemplo, una compañía de refrescos afirma que en promedio la lata de soda de 12 onzas contiene 20 calorías. Usted puede utilizar la prueba t de 1 muestra para evaluar si la afirmación de la compañía no es cierta.

La prueba t de 1 muestra compara la media de la muestra con el valor hipotético en relación con la variabilidad en la muestra. Los resultados indican si la diferencia entre los valores es estadísticamente significativa.

Utilice la prueba t de 1 muestra cuando:

- No se conoce la desviación estándar de la población
- La muestra se obtiene de una población de forma aleatoria

6.7.5 T de 2 muestras

Prueba de hipótesis para las medias de dos poblaciones, cuya finalidad es determinar si éstas son significativamente diferentes. Este procedimiento utiliza la hipótesis nula de que la diferencia entre las medias de las dos poblaciones es igual al valor hipotético (H_0: m_1 - m_2 = m_0), y la prueba comparándola con la hipótesis alternativa, la cual puede ser de cola izquierda (m_1- m_2< m_0), de cola derecha (m_1- m_2> m_0) o de dos colas (m_1- m_2≠ m_0).

Por ejemplo, supongamos que desea realizar una comparación entre dos fabricantes de automóviles – Compañía A y Compañía B – para determinar cuál de los dos fabrica los cinturones de seguridad más fuertes. Usted toma una muestra de los cinturones de seguridad de ambas compañías y mide la media de la magnitud de la fuerza necesaria para que se rompan.

La prueba t de 2 muestras analiza la diferencia entre las dos medias para determinar si la diferencia es estadísticamente significativa. La hipótesis de una prueba de dos colas sería:

- H_0:m_1 - m_2= 0 (la resistencia de los cinturones de seguridad producidos por ambas compañías es igual)
- H_1:m_1 - m_2≠ 0 (la resistencia de los cinturones de seguridad producidos por ambas compañías es diferente)

Si el valor p de la prueba es menor que el nivel de significancia que usted ha elegido, debe rechazar la hipótesis nula.

Para realizar una prueba t de 2 muestras, las dos poblaciones deben ser independientes; en otras palabras, las observaciones de la primera muestra no deben tener ninguna relación con las observaciones de la segunda muestra.

Por ejemplo, las puntuaciones en las pruebas de dos grupos separados de estudiantes son independientes, pero las mediciones realizadas antes y después en el mismo grupo de estudiantes no son independientes, aunque ambos ejemplos poseen dos muestras. Si usted no puede sustentar el supuesto de independencia entre las muestras, reconstruya su experimento a fin de utilizar una prueba t pareada para poblaciones dependientes.

6.7.6 1 proporción

Prueba de hipótesis para determinar si la proporción de ensayos que produce cierto evento es igual a un valor objetivo. Este procedimiento utiliza la hipótesis nula de que la diferencia entre las proporciones de dos poblaciones es igual a un valor hipotético (H_0: $p = p_0$). La hipótesis alternativa puede ser de cola izquierda ($p < p_0$), de cola derecha ($p > p_0$), o de dos colas ($p \neq p_0$). Los siguientes son ejemplos de datos de proporción:

- La proporción de maníes en un recipiente lleno con diferentes clases de frutos secos
- La proporción de votantes que apoyan al Candidato A en un proceso de elecciones
- La proporción de bienes manufacturados que pasan una inspección de seguridad

Supongamos que usted es el gerente del departamento de reclamaciones de garantía de una empresa que fabrica televisores. Usted desea saber si la proporción de televisores defectuosos que produce su compañía se encuentra por debajo del promedio de la industria, que es de 0.045. Usted toma una muestra de 1000 televisores y observa que hay 30 televisores defectuosos, o una proporción de 0.03. Usted utiliza una prueba de 1 proporción de una cola, con las siguientes hipótesis nula y alternativa:

H_0: $p = 0.045$

H_1: $p < 0.045$

La prueba produce un valor p de 0.010, de manera que usted rechaza la hipótesis nula y concluye que la proporción de televisores defectuosos que produce su compañía es menor que el promedio de la industria.

6.7.7 2 proporciones

Es una prueba de hipótesis para proporciones de dos poblaciones a fin de determinar si son significativamente diferentes.

Este procedimiento utiliza la hipótesis nula de que la diferencia entre las proporciones de dos poblaciones es igual a un valor hipotético (H_0: $p_1 - p_2 = P_0$), y la prueba comparándola con una hipótesis alternativa, la cual puede ser indistintamente de cola izquierda ($p_1 - p_2 < P_0$), de cola derecha ($p_1 - p_2 > P_0$), o de dos colas ($p_1 - p_2 \neq P_0$).

Por ejemplo, supongamos que el candidato A participa en una elección local. Usted desea determinar si el sexo de la persona influye en su voto, de manera que toma dos muestras – una de hombres y otra de mujeres – y observa que 55% de los hombres y 45% de las mujeres apoyan al candidato A.

Se puede utilizar la prueba de 2 proporciones para determinar si la diferencia observada es estadísticamente significativa. La hipótesis de una prueba de dos colas sería:

- ❖ H_0: $p_m - p_w = 0$ (la proporción de hombres que apoyan al candidato A es igual a la proporción de mujeres que apoyan al mismo candidato).
- ❖ H_1: $p_m - p_w \neq 0$ (la proporción de hombres que apoyan al candidato A no es igual a la proporción de mujeres que lo apoyan).

6.7.8 ANOVA de 1 solo factor

El análisis de varianza de un factor sirve para comparar varios grupos en una variable cuantitativa. Se trata, por tanto, de una generalización de la prueba T de dos muestras independientes al caso de diseños con más de dos muestras.

A la variable categórica (nominal u ordinal) que define los grupos que se desean comparar se le llama independiente o factor y generalmente se representa por VI. A la variable cuantitativa (de intervalo o razón) en la que se desea comprar los grupos se le llama dependiente y se representa por VD.

Si se quiere, por ejemplo, averiguar cuál de tres programas distintos de incentivos aumenta de forma eficaz el rendimiento de un determinado colectivo, se pueden seleccionar tres muestras aleatorias de ese colectivo y aplicar a cada una de ellas uno de los tres programas. Después, se puede medir el rendimiento de cada grupo y averiguar si existen o no diferencias entre ellos.

Se tendrá una VI categórica (el tipo de programa de incentivos) cuyos niveles se desea comprara entre sí, y una VD cuantitativa (la medida del rendimiento), en la cual se quiere comparar los tres programas. El ANOVA de un factor permite obtener información sobre el resultado de esa comparación. Es decir, permite concluir si los sujetos sometidos a distintos programas difieren la medida de rendimiento utilizada.

La hipótesis que se pone a prueba en el ANOVA de un factor es que las medias poblacionales (las medias de la VD en cada nivel de la VI) son iguales. Si las medias poblacionales son iguales, eso significa que los grupos no difieren en la VD y que, en consecuencia, la VI o factor es independiente de la VD.

6.7.9 Diseño factorial de 2 niveles

Este tipo de diseño permite asignar o cambiar el nombre a los factores y asignar valores para niveles de factores. Si los factores pueden ser continuos, utiliza niveles numéricos; si los factores son categóricos, utiliza niveles de texto.

Las variables continuas pueden tomar cualquier valor en la escala de medición que se está utilizando (por ejemplo, longitud del tiempo de reacción). Las variables categóricas sólo pueden presuponer un número limitado de valores posibles (por ejemplo, tipo de catalizador).

6.7.10 Diseño de Plackett-Burman

Diseños de experimentos factoriales de dos niveles con resolución clase III, que permiten investigar muchos factores a bajo costo. Los diseños de Plackett-Burman se utilizan para identificar los factores más importantes en las primeras etapas de la experimentación. Por lo general, estos diseños se utilizan con ocho o más factores (hasta 47).

En los diseños de Plackett-Burman, los efectos principales presentan una complicada y confusa relación con las interacciones de dos factores. Por lo tanto, debe utilizar estos diseños para estudiar los efectos principales, cuando esté dispuesto a partir del supuesto de que las interacciones de dos factores son insignificantes.

En los diseños de Plackett-Burman el número de corridas siempre es un múltiplo de cuatro (12 a 48). El número de factores debe ser menor que el número de corridas. Por ejemplo, un diseño con 12 corridas permite estimar los efectos principales para un máximo de 11 factores.

Supongamos que se están observando los diversos factores que inciden en la textura de un helado: contenido de grasa, temperatura de pasteurización, proceso de homogeneización, velocidad de mezclado y velocidad de enfriamiento. Se puede aplicar un experimento de Plackett-Burman para identificar los efectos principales más importantes, utilizar diseños fraccionales o factoriales completos para estudiar los efectos en mayor profundidad, y luego utilizar los diseños de superficie de respuesta para optimizar su proceso.

6.8 Uso de software estadístico

La innovación y mejora constante experimentada en el campo de los ordenadores personales, junto con el desarrollo de las redes informáticas (en especial de Internet) y del *software* especializado, han causado una revolución en los campos asociados a la transmisión y tratamiento de datos.

Así, resulta cada vez más frecuente encontrarse con la necesidad de analizar estadísticamente grandes volúmenes de datos con la finalidad de obtener información y, eventualmente, conocimiento. En dicho análisis, resulta fundamental el uso de un paquete estadístico.

Minitab es un paquete estadístico que abarca todos los aspectos necesarios para el aprendizaje y la aplicación de la Estadística en general. El programa incorpora opciones vinculadas a las principales técnicas de análisis estadístico (análisis descriptivo, contrastes de hipótesis, regresión lineal y no lineal, series temporales, análisis de tiempos de fallo, control de calidad, análisis factorial, ANOVA, etc.), además de proporcionar un potente entorno gráfico y de ofrecer total compatibilidad con los editores de texto, hojas de cálculo y bases de datos más usuales.

Con la finalidad de mostrar de manera práctica la aplicación de las pruebas estadísticas antes mencionadas, se llevará a cabo una corrida experimental con datos reales, para lo cual se utilizará un Minitab.

6.8.1 Corrida experimental con Minitab

En una compañía de fabricación de componentes para equipo agrícola y de construcción, se cambió la operación de corte en cizalla a corte en lasser, por lo cual los ingenieros de proceso deben validar nuevamente las especificaciones de las piezas.

Se desea conocer cuál es el tamaño de muestra óptimo para saber si el peso medio de las piezas que cortan en lasser es de 500 grs., sabiendo que se tiene una desviación típica de 10 grs.

Para éste problema, es necesario plantearse con antelación qué diferencia mínima se está dispuesto a asumir y con qué probabilidad, es decir, el tamaño de la muestra dependerá de si se desea detectar una diferencia en la media de 10, 15 ó 20 unidades, según sea el caso y la complejidad para obtener las piezas (costo, tiempo de fabricación, etc.). Para esto, MINITAB obliga a introducir la diferencia y la probabilidad mínima que se desea detectar.

Así, si se quiere tomar una muestra para comprobar que el peso medio no es de 500 grs., evidenciando diferencias de 5 grs. por lo menos el 80% de las veces, se deberá introducir como valor de potencia 0.8 y como valor de diferencia mínima 5.

ESCENARIO 1

- Se selecciona en el menú de la barra de herramientas ESTADISTICAS>POTENCIA Y TAMAÑO DE LA MUESTRA>Z DE 1 MUESTRA

- En el cuadro diálogo: TAMAÑOS DE MUESTRA: en blanco, ya que es el parámetro que se desea conocer; DIFERENCIAS: 5; VALORES DE POTENCIA: 0.8; DESVIACIÓN ESTÁNDAR: 10.

- En Opciones: HIPÓTESIS ALTERNA: No es igual a; NIVEL DE SIGNIFICANCIA: 0.95. Los resultados son los siguientes:

Prueba Z de 1 muestra

Probando la media = nula (no vs. = nula)

Calculando la potencia para la media = nulo + diferencia

Alfa = 0.05 Desviación estándar asumida = 10

Diferencia	*Tamaño De la Muestra*	*Potencia Objetivo*	*Potencia Real*
5	*32*	*0.8*	*0.807430*

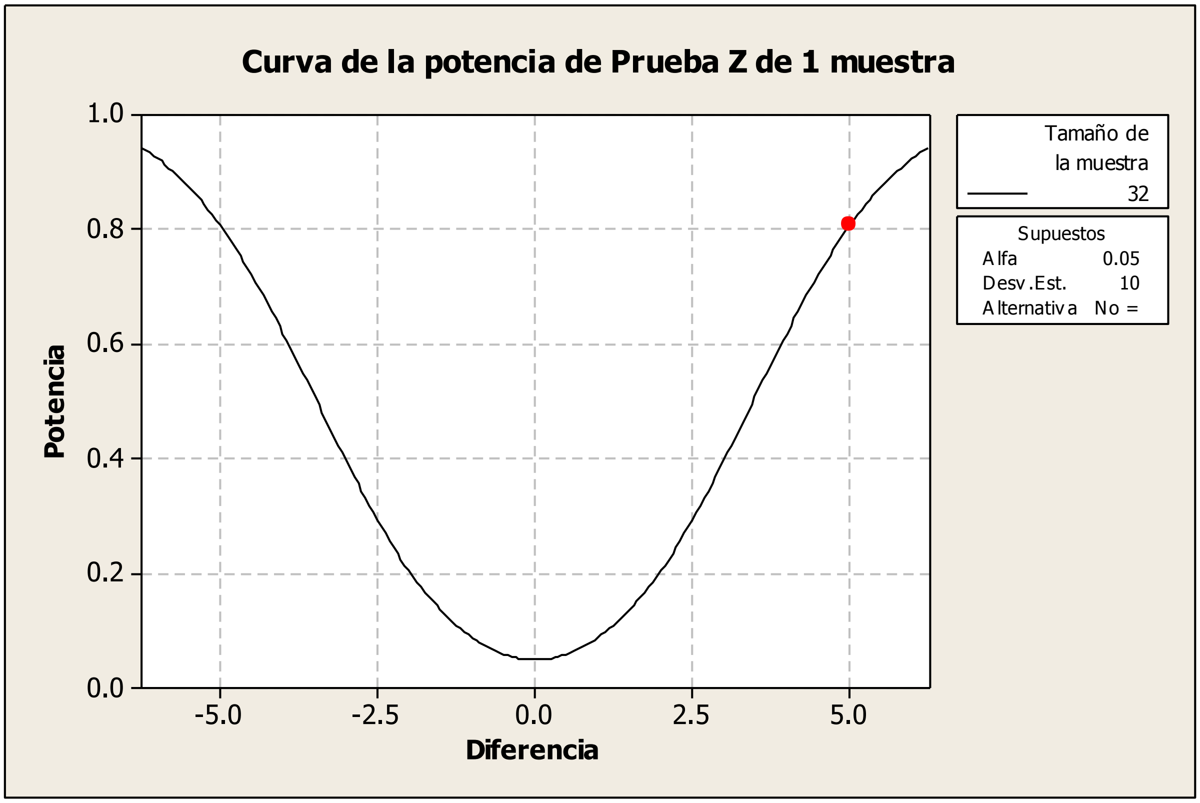

FIGURA 9. Curva de potencia de 80%

Para comprobar la robustez del proceso, los ingenieros también desean saber hasta cuantos gramos de desviación se pueden detectar en el peso de las piezas, así como que porcentajes de potencia pueden ser obtenidos variando dichas diferencias. Esto con la finalidad de hacer una proyección del impacto monetario necesario para utilizar en piezas prototipo. Las instrucciones son las siguientes:

ESCENARIO 2

- Se selecciona en el menú de la barra de herramientas ESTADISTICAS>POTENCIA Y TAMAÑO DE LA MUESTRA>Z DE 1 MUESTRA

- En el cuadro de diálogo: TAMAÑOS DE MUESTRA: se deja en blanco; DIFERENCIAS: 5 10 15; VALORES DE POTENCIA: 0.8; DESVIACIÓN ESTÁNDAR: 10.

- En Opciones: HIPÓTESIS ALTERNA: No es igual a; NIVEL DE SIGNIFICANCIA: 0.95.

Prueba Z de 1 muestra

Probando la media = nula (no vs. = nula)

Calculando la potencia para la media = nulo + diferencia

Alfa = 0.05 Desviación estándar asumida = 10

Diferencia	*Tamaño De la Muestra*	*Potencia Objetivo*	*Potencia Real*
5	*32*	*0.8*	*0.807430*
10	*8*	*0.8*	*0.807430*
15	*4*	*0.8*	*0.850839*

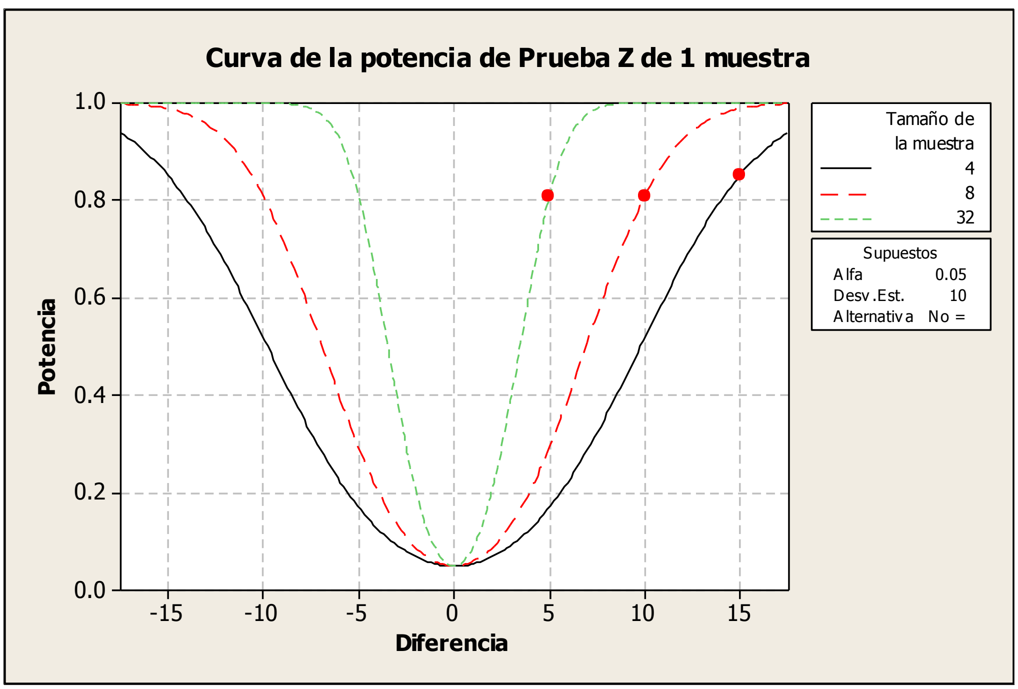

FIGURA 10. Curvas de potencia para diferentes niveles de detección de diferencias

Así puede observarse más directamente la importante relación entre el poder de la prueba y los tamaños de muestra necesarios para alcanzarlo. Este es, en la mayoría de los casos, el punto clave de la realización de una prueba estadística exitosa, ya que en la industria muchas veces se carece de la libertad económica de gastar dinero en piezas para las pruebas prototipo.

Si lo que se busca es variar los parámetros de potencia de la prueba, ahora bastará con comparar diferentes porcentajes, para ello pueden variarse individualmente observando su comportamiento puntual, ó en grupos, observando las comparaciones entre cada uno de ellos.

Estas variaciones son generalmente con el fin de observar hasta que nivel de confianza se le puede ofrecer a un cliente al momento de realizar la prueba estadística de un control de proceso. No es lo mismo ajustar un proceso para detectar diferencias hasta en el 80% de los casos, que robustecerlo para lograr detecciones por encima del 95% ó 99%. El impacto monetario aumenta con cada punto porcentual según sea el caso particular. Por lo tanto, las instrucciones serían las siguientes.

ESCENARIO 3

- Se selecciona en el menú de la barra de herramientas ESTADISTICAS>POTENCIA Y TAMAÑO DE LA MUESTRA>Z DE 1 MUESTRA

- En el cuadro de diálogo: TAMAÑOS DE MUESTRA: se deja en blanco; DIFERENCIAS: 5; VALORES DE POTENCIA: 0.8, 0.9, 0.95; DESVIACIÓN ESTÁNDAR: 10.

- En Opciones: HIPÓTESIS ALTERNA: No es igual a; NIVEL DE SIGNIFICANCIA: 0.95.

Prueba Z de 1 muestra

Probando la media = nula (no vs. = nula)
Calculando la potencia para la media = nulo + diferencia
Alfa = 0.05 Desviación estándar asumida = 10

	Tamaño De la	Potencia	Potencia
Diferencia	Muestra	Objetivo	Real
5	32	0.80	0.807430
5	43	0.90	0.906375
5	52	0.95	0.950076

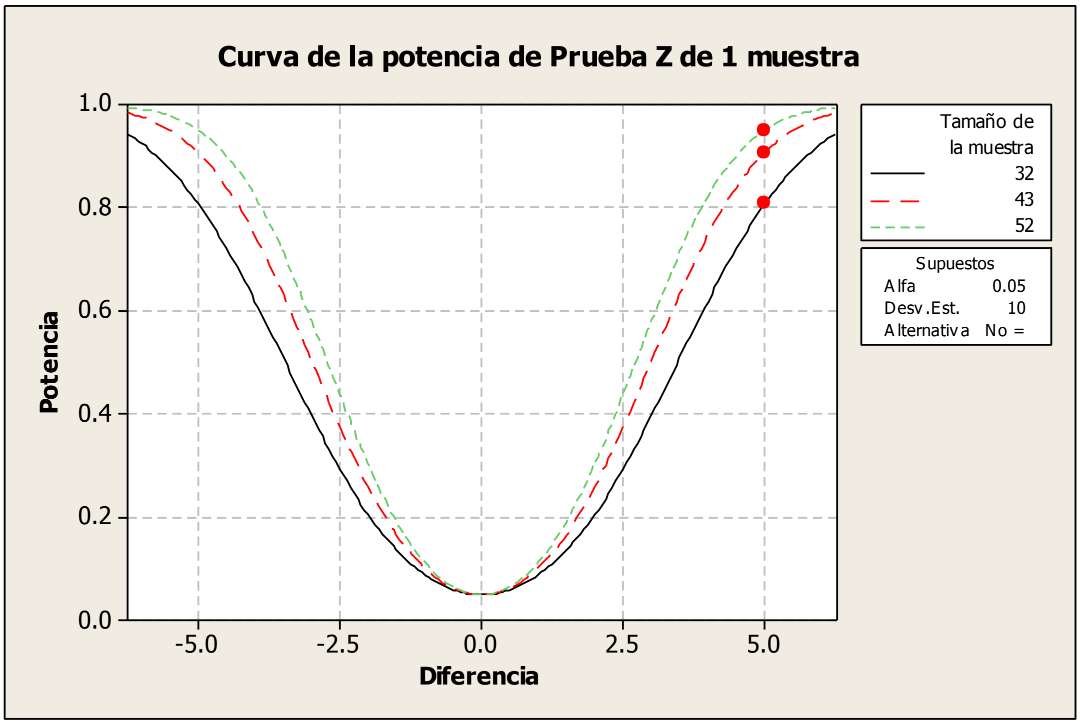

FIGURA 11. Curvas de potencia para diferentes niveles de poder en la prueba.

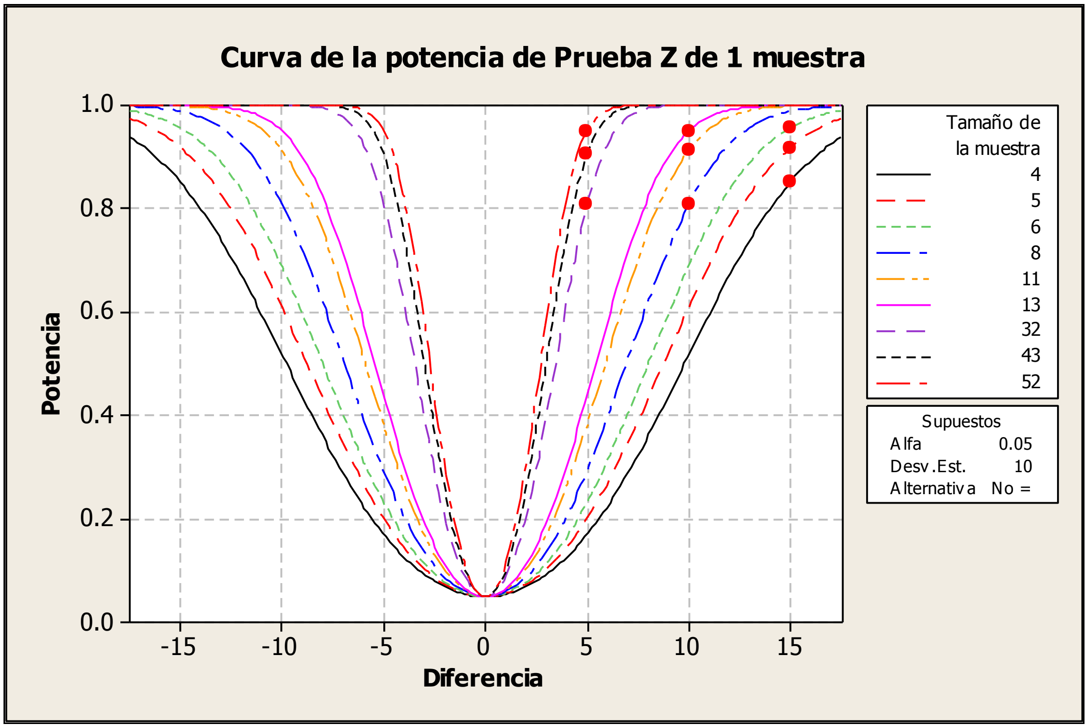

FIGURA 12. Curvas de potencia para la variación de poder y nivel de detección.

Esta misma ventana podrá ser utilizada si, una vez tomada la muestra, se quisiera conocer qué diferencias son las que detecta con mayor probabilidad.

Para esta respuesta se deberá introducir en el cuadro de diálogo el tamaño de muestra tomado y la probabilidad:

ESCENARIO 4

- Se selecciona en el menú de la barra de herramientas ESTADISTICAS>POTENCIA Y TAMAÑO DE LA MUESTRA>Z DE 1 MUESTRA

- En el cuadro de diálogo: TAMAÑOS DE MUESTRA: 15; DIFERENCIAS: en blanco, para observar la variación; VALORES DE POTENCIA: 0.90 0.95; DESVIACIÓN ESTÁNDAR: 10.

- En Opciones: HIPÓTESIS ALTERNA: No es igual a; NIVEL DE SIGNIFICANCIA: 0.95.

Prueba Z de 1 muestra

Probando la media = nula (no vs. = nula)

Calculando la potencia para la media = nulo + diferencia

Alfa = 0.05 Desviación estándar asumida = 10

Tamaño De la Muestra	Potencia	Diferencia
15	0.90	8.36956
15	0.95	9.30760

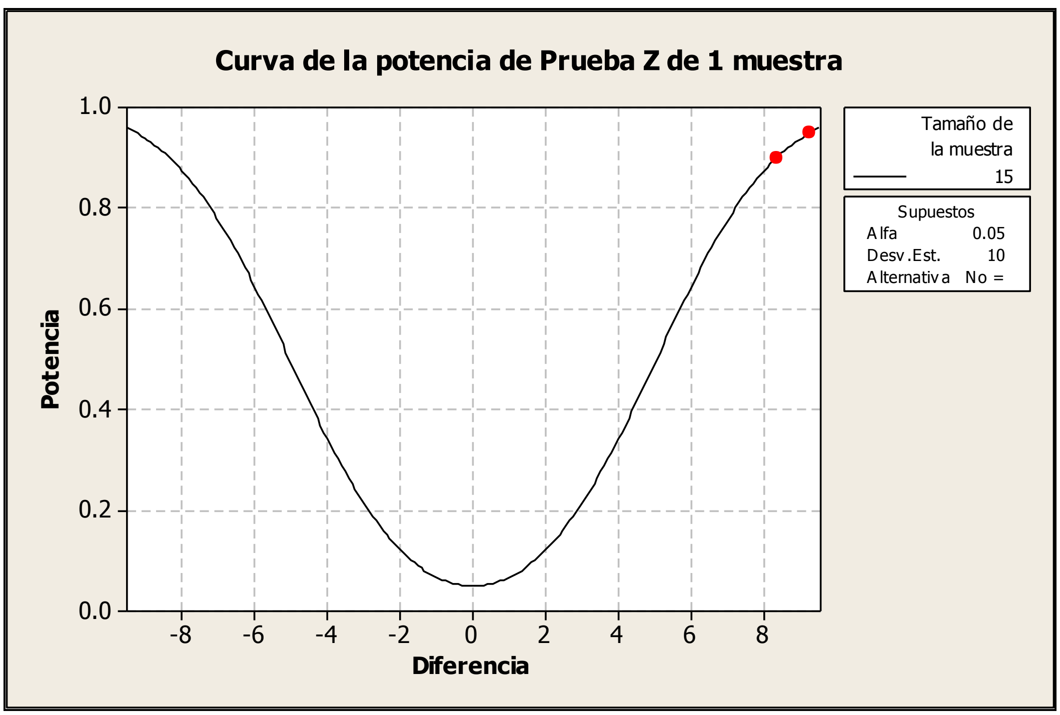

FIGURA 13. Curva de potencia para diferentes niveles de detección de diferencias a 90% y 95%.

Finalmente, si se desea saber la potencia en base a diferencias a detectar dadas y tamaños de muestra fijos, las instrucciones serían las siguientes.

ESCENARIO 5

- Se selecciona en el menú de la barra de herramientas ESTADISTICAS>POTENCIA Y TAMAÑO DE LA MUESTRA>Z DE 1 MUESTRA

- En el cuadro de diálogo: TAMAÑOS DE MUESTRA: 10 20 30; DIFERENCIAS: 5 10 15; VALORES DE POTENCIA: en blanco, para observar la variación; DESVIACIÓN ESTÁNDAR: 10.

- En Opciones: HIPÓTESIS ALTERNA: No es igual a; NIVEL DE SIGNIFICANCIA: 0.95.

Prueba Z de 1 muestra

Probando la media = nula (no vs. = nula)

Calculando la potencia para la media = nulo + diferencia

Alfa = 0.05 Desviación estándar asumida = 10

Diferencia	*Tamaño De la Muestra*	*Potencia*
5	*10*	*0.35261*
5	*20*	*0.60878*
5	*30*	*0.78191*
10	*10*	*0.88538*
10	*20*	*0.99400*
10	*30*	*0.99978*
15	*10*	*0.99731*
15	*20*	*1.00000*
15	*30*	*1.00000*

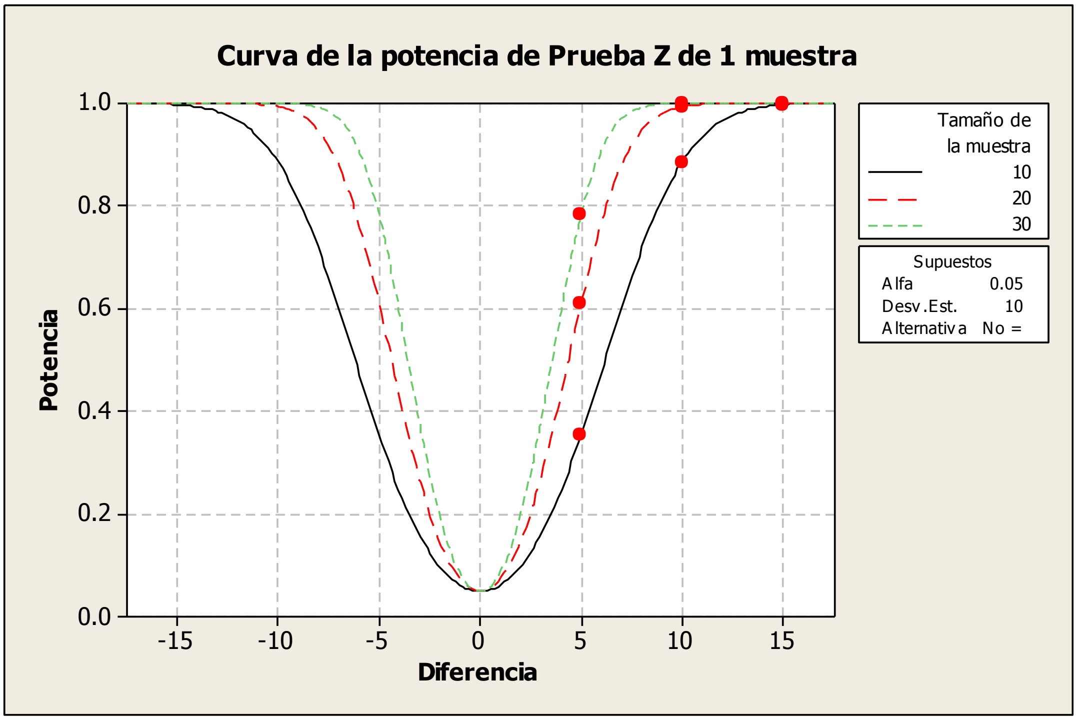

FIGURA 14. Curva de potencia para diferentes niveles de detección de diferencias a 90% y 95%.

6.8.2 Análisis de Resultados

Las pruebas realizadas demostraron que los tres factores (diferencia, tamaño de la muestra y potencia) tienen una alta correlación en la efectividad de la prueba estadística. La variación de un parámetro impacta directamente en el desempeño de los demás.

- Escenario 1.

El resultado dice que bastará tomar los pesos de 32 piezas elegidas al azar de la producción de la máquina lasser para poder detectar con una probabilidad del 80% una desviación de la media superior o igual a 5 gramos.

En términos generales, una potencia de 80% como la manejada en este primer escenario es relativamente baja, comparada con los estándares de la industria, sin embargo, dadas las circunstancias del problema, es suficiente para lograr observar las diferencias que se desean evidenciar.

- Escenario 2.

El resultado dice que bastará tomar los pesos de 4 y 8 piezas elegidas al azar de la producción de la máquina lasser para poder detectar con una probabilidad del 80% una desviación de la media superior o igual a 15 y 10 gramos respectivamente.

En éste escenario, se es más flexible con las diferencias a detectar, lo cual impacta directamente en la disminución de las piezas que deben tomarse para realizar las pruebas, sin embargo se admite la posibilidad de dejar pasar gramos adicionales en cada pieza.

- Escenario 3.

Como se observa en estos resultados, el tamaño de muestra necesario para detectar diferencias superiores a 10 gramos es bastante inferior y en concreto de tomar por ejemplo 7 datos a tomar 8, aumenta un 10% la probabilidad de detectar una mejora de 10 gramos en el peso medio.

Este escenario hace evidente que si la toma de datos no es excesivamente costosa, esta sería la mejor elección.

- Escenario 4.

Para una muestra de 15 piezas, se podría detectar muy probablemente (en el 90% de los casos) una variación de las piezas de hasta 8.3 gramos respecto de la media nominal. A la par, con el mismo número de piezas, pero en el 95% de los casos, se podrían detectar variaciones de 9.3 gramos en los pesos de las piezas.

- Escenario 5.

Para una muestra de hasta 30 piezas podríamos detectar solo en el 78% de los casos variaciones en los pesos de hasta 5 gramos. Sin embargo, en el 100% de los casos para muestras de solo 10 piezas, se podrían detectar variaciones de hasta 15 gramos.

En este último escenario se hace evidente que para lograr una potencia mayor en la prueba, será determinante la variación que se desee detectar. Si la variación es muy pequeña, se necesitará un mayor número de piezas. Si por el contrario, es holgada, se necesitará un menor número de piezas. Los parámetros varían proporcionalmente.

VII. CONCLUSION

El análisis adecuado del poder estadístico de una investigación, que es en definitiva la capacidad que tiene el estudio para encontrar diferencias si es que realmente las hay, es un paso fundamental tanto en la fase de diseño como en la interpretación y discusión de sus resultados.

A la hora del diseño, por tanto, debe establecerse la magnitud mínima de la diferencia o asociación que se considere de relevancia, así como el poder estadístico que se desea para el estudio y, de acuerdo con ello, calcular el tamaño de la muestra necesaria.

Tras realizar el análisis estadístico, deberá cuestionarse antes que nada si la ausencia de significación estadística indica realmente que no existe una diferencia o asociación relevante, o simplemente que no se dispone de suficiente número de muestras para obtener hallazgos significativos.

Este tipo de decisiones se basan generalmente en la negociación de las especificaciones con el cliente, o bien, en el precio de fabricación de piezas prototipo. Mientras que en algunas industrias puede contarse fácilmente con un buen número de piezas prototipo, en otras la obtención de una sola pieza para hacer pruebas resulta sumamente costosa.

Como se vio hasta aquí, entonces, las posibles aplicaciones prácticas de la potencia son varias:

- Se puede determinar un tamaño muestral adecuado antes de efectuar una investigación.
- Se puede determinar la viabilidad o inviabilidad de una investigación dadas ciertas limitantes (habitualmente relacionadas con el tamaño muestral).
- Se controla el riesgo de efectuar errores Tipo II, que por lo general no son tenidos en cuenta en las investigaciones cuantitativas.
- Se puede obtener un índice de potencia real observada para una investigación. Cuando éste nivel es lo suficientemente alto, la investigación gana en rigor y en posibilidades de aceptación.

Algunos autores, Cohen (1992) por ejemplo, consideran alta una potencia igual o superior a 80%, lo que se debe tener en cuenta al diseñar un estudio, particularmente, al definir el tamaño de muestra.

La potencia estadística, como ya se mencionó reiteradamente durante éste trabajo, depende de varios factores, entre ellos el tamaño del efecto a detectar o la magnitud mínima de la diferencia que se considera importante, la variabilidad de la respuesta estudiada, el nivel de significancia escogido (α) y el tamaño de muestra.

En la figura 14 se muestran los tamaños de muestra necesarios para alcanzar distintos niveles de potencia en función de efectos de magnitud variable, expresados en términos de desviación estándar de la variable respuesta de interés, para un nivel de significancia, α = 0.05, al comparar los promedios de un grupo tratado y un grupo de control.

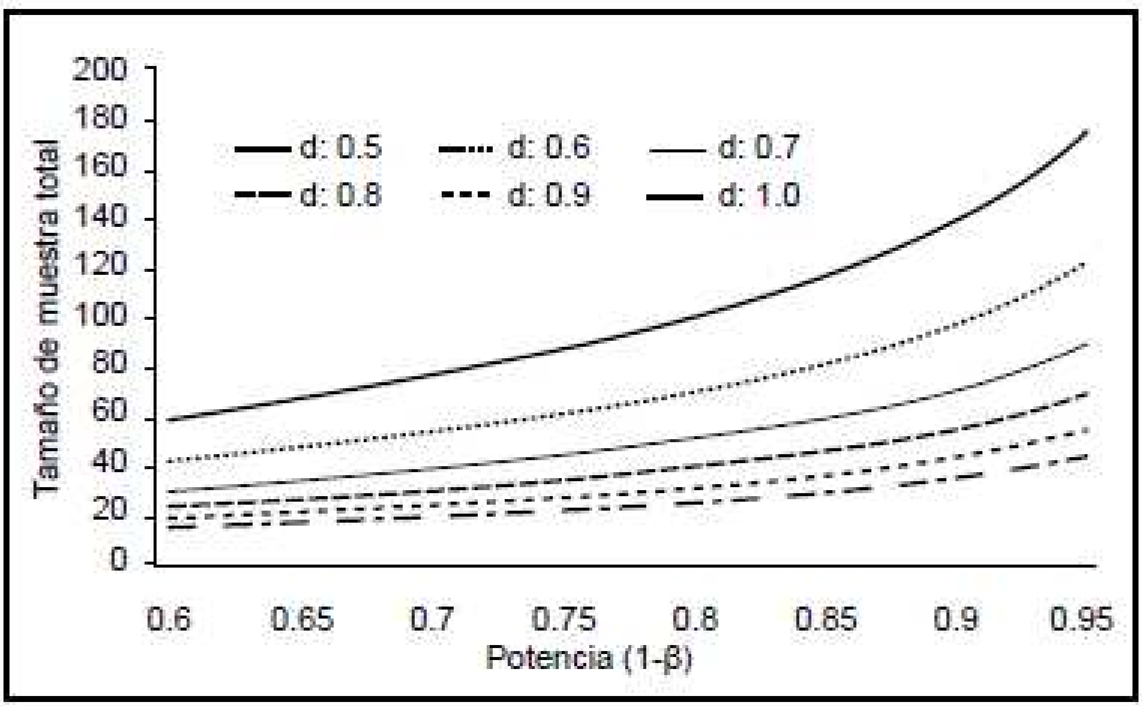

FIGURA 15. Relación entre tamaño de muestra, tamaño de efecto (d) y potencia, con un nivel de α de 0.05 al comparar 2 promedios.

Nótese que todas las líneas muestran una tendencia creciente en función de la potencia, es decir, cualquiera sea la magnitud del efecto, al incrementarse el nivel de potencia deseado se incrementa el tamaño de muestra. Por otra parte, para un nivel de potencia determinado, conforme se reduce la magnitud del efecto se incrementa el tamaño de muestra.

En un caso clínico de aplicación, Ramírez-Figueroa y Sotres-Ramos (2005) proponen un ejemplo en el que se compara la eficacia antihipertensiva de monoxodin y captotril. Se estableció como efecto significativo una diferencia de 8 mmHg y se estimó una desviación estándar de 10 mmHg, por lo que la magnitud del efecto estandarizado fue d: 8/10 = 0.8. Se estableció además un nivel de significancia, α, de 0.05. ¿Qué tamaño de muestra se requiere para alcanzar una potencia de 0.9? En la figura ese valor se obtiene proyectando al eje vertical el punto en que una línea perpendicular al valor de potencia 0.9 corta la curva correspondiente al tamaño de efecto 0.8, como lo indican las flechas. Ese valor corresponde a un tamaño de muestra de 55. Si se quisiera un tamaño de efecto menor, por ejemplo 0.6, el tamaño de muestra sería cercano a 95.

En resumen, es importante tener en cuenta la potencia estadística cuando se diseña un estudio, de manera que el tamaño de muestra utilizado garantice una elevada probabilidad de detectar diferencias si éstas realmente existen. Llevar a cabo estudios con baja potencia estadística no es éticamente aceptable, puede conducir a resultados de dudosa validez científica y al uso inadecuado de recursos de investigación.

VIII. GLOSARIO

- *Desviación estándar:* Es una medida de centralización o dispersión para variables de razón y de intervalo, que representa el promedio aritmético de fluctuación de una serie de datos respecto a su media.

- *Error:* Es la diferencia entre el valor de un estimador y el del parámetro correspondiente.

- *Estadística descriptiva:* Tiene por objeto fundamental describir y analizar las características de un conjunto de datos, obteniéndose de esa manera conclusiones sobre las características de dicho conjunto y sobre las relaciones existentes con otras poblaciones, a fin de compararlas.

- *Estadística Inferencial:* Está fundamentada en los resultados obtenidos del análisis de una muestra significativa de población, con el fin de inducir o inferir el comportamiento o característica de dicha población.

- *Hipótesis alterna:* Es cualquier hipótesis que difiera de la hipótesis nula. Es una afirmación que se acepta si los datos muestrales proporcionan evidencia suficiente de que la hipótesis nula es falsa.

- *Hipótesis nula:* Se refiere siempre a un valor especificado del parámetro de población, no a una estadística de muestra. El planteamiento de la hipótesis nula siempre contiene un signo de igualdad con respecto al valor especificado del parámetro.

- *Muestra:* Se llama muestra a una parte de la población a estudiar qué sirve para representarla.

- *Potencia estadística:* Se define como la probabilidad de tomar la decisión de rechazar la hipótesis nula cuando ésta es verdadera.

- *Significancia estadística:* Un resultado se denomina estadísticamente significativo cuando no es probable que haya sido debido al azar.

- *Variable:* También suelen ser llamados caracteres cuantitativos, son aquellos que pueden ser expresados mediante números. Son caracteres susceptibles de medición.

IX. BIBLIOGRAFÍA

Anderson, D., Sweeney, D. 2004. Estadística para administración y economía. Cengage. pp. 302-350.

Argimon, J. 2002. El intervalo de confianza: algo más que un valor de significación estadística. Med. Clin. (Barc), Cap. 118. pp. 382-384.

ASQC. 1996b. "Statistical Thinking", Special Edition, Quality Information Center, Statistics Division, Spring. (www.revistaespacios.com/a04v25n03/04250321).

Baena, C. 2008. Lean manufacturing: Guía. Capítulo 1.

Berenson, M., Levine, D. 2006. Estadística Básica en Administración, Conceptos y Aplicaciones. Capítulo 6. pp. 258-300.

Braitman, L. 1991. Confidence intervals assess both clinical significance and statistical significance. Ann Intern. Med. Cap. 114. pp. 515-517.

Brown, S., Morrison, G. 1991. The Introduction to Six-Sigma Methodology. Capítulo 4.

Caballero Armas, W., 1975. Introducción a la Estadística. Orton / CATIE. pp. 150-175.

Cohen, J. 1992. A power premier. Psychological Bulletin 112. pp. 155-159.

Dransfield, S., Fisher, N., Vogel, N. 1999. Using statistics and statistical thinking to improve organizational performance. International Statistical Review. pp. 99-150. (www.revistaespacios.com/a04v25n03/04250321).

Evans, J., Lindsay, W. 2000. Administración y Control de la Calidad. pp. 599-610.

Escalante, E. 2006. Análisis y mejoramiento de la calidad. Capítulo 7, p. 315.

Hoerl, R., Snee, D. 2002. Statistical Thinking - Improving Business Performance Duxbury Press, United States of America. (www.revistaespacios.com/a04v25n03/04250321).

Gutiérrez, H., De La Vera, R. 2008. Control Estadístico de Calidad y Seis Sigma. Capítulo 11.

López, L. L. y Sánchez, S. P. 2004. Pensamiento estadístico para los empresarios del Siglo XXI. Industrial. XXV, 1. pp. 3-9. (www.revistaespacios.com/a04v25n03/04250321).

Miranda, L. 2006. Seis Sigma: Guía para principiantes. pp. 55-68.

Ramírez-Figueroa, C., Sotres-Ramos, D. 2005. Tablas de potencia y tamaño de muestra de la prueba Uniformemente Más Potente Invariante (UMPI) para demostrar equivalencia de las medias de dos distribuciones normales. Agrociencia Cap. 39. pp. 657-666.

Torino, H. 1997. Resumen del libro de estadística de Berenson y Levine. Monografía en internet. www.monografías.com/trabajos/beren/bern.shtml.

Welkowitz, J. 1981. Estadística aplicada a las ciencias de la educación. Aula XXI. pp. 207-232.