TEORÍA DE LAS MUESTRAS DE TRABAJO

Aportado por: **Dos Santos, Ma. Yolanda** yolanda_dossantos@yahoo.com

---

Republica Bolivariana de Venezuela
Ministerio de Educación Superior
Instituto Universitario Nuevas Profesiones
Semestre: 4to
Materia: Estadísticas II
Profesor: Aira Grover

INDICE

INTRODUCCIÓN ... 1
UNIDAD 1. TEORÍA DE MUESTRAS ... 1
2. IMPORTANCIA DEL MUESTREO ... 2
3. TAMAÑO DE LAS MUESTRAS ... 3
4. MUESTRAS PROBABILÍSTICAS ... 6
5. MUESTRA CON Y SIN REPOSICION ... 10
6. DISTRIBUCIONES DE PROBABILIDAD EN EL MUESTREO ... 10
7. DISTRIBUCIÓN MUESTRAL DE LA MEDIA ... 12
8. DISTRIBUCIÓN MUESTRAL DE PROPORCIONES ... 15
9. DIFERENCIAS DE MEDIAS Y DE PROPORCIONES ... 17
UNIDAD 2 TEORÍA DE PEQUEÑAS MUESTRAS ... 20
1. EL MODELO DE STUDENT ... 21
3. EL MODELO DE CHI-CUADRADO ... 30
4. EL MODELO DE FISHER ... 33
CONCLUSION ... 39

INTRODUCCIÓN

Una parte fundamental para realizar un estudio estadístico de cualquier tipo es obtener unos resultados confiables y que puedan ser aplicables. Como ya se comentó anteriormente, resulta casi imposible o impráctico llevar a cabo algunos estudios sobre toda una población, por lo que la solución es llevar a cabo el estudio basándose en un subconjunto de ésta denominada muestra.

Sin embargo, para que los estudios tengan la validez y confiabilidad buscada es necesario que tal subconjunto de datos, o muestra, posea algunas características específicas que permitan, al final, generalizar los resultados hacia la población en total. Esas características tienen que ver principalmente con el tamaño de la muestra y con la manera de obtenerla.

En las siguientes secciones de esta unidad lo comentaremos.

La Alumna

UNIDAD 1. TEORÍA DE MUESTRAS

En este capítulo se resume la Teoría de Muestras estadística, la cual trata el concepto de estudiar una población desconocida tomándole muestras, y a través del estudio de las mismas poder hacer inferencias acerca de toda la población. Primero se analiza el caso del muestreo aleatorio simple y estratificado, mostrando el manejo de una tabla de números aleatorios. Luego se ven los tipos no-aleatorios en el muestreo y se discute acerca de las ventajas y desventajas de cada uno de los métodos. Se explican los métodos usados en Bioquímica para lograr que las muestras extraídas a los pacientes cumplan los requisitos de aleatoriedad, aunque sea aproximadamente. Lo mismo para el caso de la industria Farmacéutica y para las simulaciones en los muestreos de mercadeo, usadas en el comercio en general. De manera tal de poder aplicar luego los modelos estadísticos que exigen tal requisito. En la Tabla 3, del fascículo de las tablas, se presenta la Tabla Aleatoria, más conocida como: "Random Numbers".

2. IMPORTANCIA DEL MUESTREO

A lo largo del curso se hacen uso de dos tipos de razonamiento: el **deductivo** y el **inductivo**. El primero está relacionado directamente con la teoría de probabilidad, que se aborda en la unidad 4, y que a partir de las características de la población se obtienen las posibles características de una muestra. El segundo tipo de razonamiento se relaciona con la denominada **inferencia estadística**: utilizar las características de un subconjunto de la población (la muestra) para hacer afirmaciones (inferir) sobre la población en general. Éste será el caso de esta unidad.

El muestro, como ya se mencionó, implica algo de incertidumbre que debe ser aceptada para poder realizar el trabajo, pues aparte de que estudiar una población resulta ser un trabajo en ocasiones demasiado grande, Wonnacott y Wonnacott ofrecen las siguientes razones extras:

- **Recursos limitados**. Es decir, no existen los recursos humanos, materiales o económicos para realizar el estudio sobre el total de la población. Es como cuando se compra un aparato, un automóvil usado (por ejemplo), que se prueba unos minutos (el encendido, una carrerita, etc.) para ver si funciona correctamente y luego se adquiere, pero no se espera a probarlo toda la vida (encendiéndolo y apagándolo o, simplemente, dejándolo encendida) antes de realizar la adquisición.
- **Escasez**. Es el caso en que se dispone de una sola muestra. Por ejemplo, para el estudio paleontológico de los dinosaurios (el T. Rex por ejemplo) sería muy bueno contar con, al menos, muchos restos fósiles y así realizar tales investigaciones; sin embargo, se cuenta sólo con una docena de esqueletos fosilizados (casi todos incompletos) de esas criaturas en todo el mundo.
- **Pruebas destructivas**. Es el caso en el que realizar el estudio sobre toda la población llevaría a la destrucción misma de la población. Por ejemplo, si se quisiese saber el conteo exacto de hemoglobina de una persona habría que extraerle **toda** la sangre.

- **El muestreo puede ser más exacto**. Esto es en el caso en el que el estudio sobre la población total puede causar errores por su tamaño o, en el caso de los censos, que sea necesario utilizar personal no lo suficientemente capacitado; mientras que, por otro lado, el estudio sobre una muestra podría ser realizada con menos personal pero más capacitado.

Ya que hemos mencionado la necesidad de realizar muestras, continuaremos con algunas características que deben tener éstas para que, realmente, se puedan realizar inferencias (inducciones) sobre ellas hacia la población total.

3. TAMAÑO DE LAS MUESTRAS

Para calcular el tamaño de una muestra hay que tomar en cuenta tres factores:

1. El porcentaje de confianza con el cual se quiere generalizar los datos desde la muestra hacia la población total.
2. El porcentaje de error que se pretende aceptar al momento de hacer la generalización.
3. El nivel de variabilidad que se calcula para comprobar la hipótesis.

◆ La **confianza** o el **porcentaje de confianza** es el porcentaje de seguridad que existe para generalizar los resultados obtenidos. Esto quiere decir que un porcentaje del 100% equivale a decir que no existe ninguna duda para generalizar tales resultados, pero también implica estudiar a la totalidad de los casos de la población.

Para evitar un costo muy alto para el estudio o debido a que en ocasiones llega a ser prácticamente imposible el estudio de todos los casos, entonces se busca un porcentaje de confianza menor. Comúnmente en las investigaciones sociales se busca un 95%.

◆ El **error** o **porcentaje de error** equivale a elegir una probabilidad de aceptar una hipótesis que sea falsa como si fuera verdadera, o la inversa: rechazar a hipótesis verdadera por considerarla falsa. Al igual que en el caso de la confianza, si se quiere eliminar el riesgo del error y considerarlo como 0%, entonces la muestra es del mismo tamaño que la población, por lo que conviene correr un cierto riesgo de equivocarse.

Comúnmente se aceptan entre el 4% y el 6% como error, tomando en cuenta de que **no** son complementarios la confianza y el error.

◆ La **variabilidad** es la probabilidad (o porcentaje) con el que se aceptó y se rechazó la hipótesis que se quiere investigar en alguna investigación anterior o en un ensayo previo a la investigación actual. El porcentaje con que se aceptó tal hipótesis se denomina **variabilidad**

positiva y se denota por *p*, y el porcentaje con el que se rechazó se la hipótesis es la **variabilidad negativa**, denotada por *q*.

Hay que considerar que *p* y *q* son complementarios, es decir, que su suma es igual a la unidad: $p+q=1$. Además, cuando se habla de la máxima variabilidad, en el caso de no existir antecedentes sobre la investigación (no hay otras o no se pudo aplicar una prueba previa), entonces los valores de variabilidad es $p=q=0.5$.

Una vez que se han determinado estos tres factores, entonces se puede calcular el tamaño de la muestra como a continuación se expone.

Hablando de una población de alrededor de 10,000 casos, o mínimamente esa cantidad, podemos pensar en la manera de calcular el tamaño de la muestra a través de las siguientes fórmulas. Hay que mencionar que estas fórmulas se pueden aplicar de manera aceptable pensando en instrumentos que no incluyan preguntas abiertas y que sean un total de alrededor de 30.

Vamos a presentar dos fórmulas, siendo la primera la que se aplica en el caso de que **no se conozca con precisión el tamaño de la población**, y es:

$$n = \frac{Z^2 p q}{E^2}$$

Donde:

n es el tamaño de la muestra;
Z es el nivel de confianza;
p es la variabilidad positiva;
q es la variabilidad negativa;
E es la precisión o error.

Hay que tomar nota de que debido a que la variabilidad y el error se pueden expresar por medio de porcentajes, hay que convertir todos esos valores a proporciones en el caso necesario.

También hay que tomar en cuenta que el nivel de confianza no es ni un porcentaje, ni la proporción que le correspondería, a pesar de que se expresa en términos de porcentajes. El nivel de confianza se obtiene a partir de la distribución normal estándar, pues la proporción correspondiente al porcentaje de confianza es el área simétrica bajo la curva normal que se toma como la confianza, y la intención es buscar el valor *Z* de la variable aleatoria que corresponda a tal área.

Por ejemplo: Si se quiere un porcentaje de confianza del 95%, entonces hay que considerar la proporción correspondiente, que es 0.95. Lo que se buscaría en seguida es el valor *Z* para la variable aleatoria *z* tal que el área simétrica bajo la curva normal desde -*Z* hasta *Z* sea igual a 0.95, es decir, $P(-Z<z<Z)=0.95$.

Utilizando las tablas, o la función DISTR.NORM.ESTAND.INV() del *Excel*, se puede calcular el valor de *Z*, que sería 1.96 (con una aproximación a dos decimales).

Esto quiere decir que $P(-1.96<z<1.96)=0.95$.

En el caso de que **sí se conozca el tamaño de la población** entonces se aplica la siguiente fórmula:

$$n = \frac{Z^2 pq\, N}{NE^2 + Z^2 pq}$$

Donde

n es el tamaño de la muestra;
Z es el nivel de confianza;
p es la variabilidad positiva;
q es la variabilidad negativa;
N es el tamaño de la población;
E es la precisión o el error.

La ventaja sobre la primera fórmula es que al conocer exactamente el tamaño de la población, el tamaño de la muestra resulta con mayor precisión y se pueden incluso ahorrarse recursos y tiempo para la aplicación y desarrollo de una investigación.

Por ejemplo: En el *Colegio de Bachilleres*, una institución de nivel medio superior, se desea realizar una investigación sobre los alumnos inscritos en primer y segundo años, para lo cual se aplicará un cuestionario de manera aleatoria a una muestra, pues los recursos económicos y el tiempo para procesar la información resultaría insuficiente en el caso de aplicársele a la población estudiantil completa.

En primera instancia, suponiendo que no se conoce el tamaño exacto de la población, pero con la seguridad de que ésta se encuentra cerca a los diez millares, se aplicará la primera fórmula.

Se considerará una confianza del 95%, un porcentaje de error del 5% y la máxima variabilidad por no existir antecedentes en la institución sobre la investigación y porque no se puede aplicar una prueba previa.

Primero habrá que obtener el valor de Z de tal forma que la confianza sea del 95%, es decir, buscar un valor de Z tal que $P(-Z<z<Z)=0.95$. Utilizando las tablas o las funciones de *Excel* se pueden obtener, o viendo (en este caso) el ejemplo anterior, resulta que $Z=1.96$.

De esta manera se realiza la sustitución y se obtiene:

$$n=\frac{(1.96^2)(0.5)(0.5)}{0.05^2}=\frac{(3.8416)(0.25)}{0.0025}=\frac{0.9604}{0.0025}=384.16$$

Esto quiere decir que el tamaño de la muestra es de 385 alumnos.

Supongamos ahora que sí se conoce el tamaño de la población estudiantil y es de 9,408, entonces se aplicará la segunda fórmula. Utilizando los mismos parámetros la sustitución queda como:

$$n=\frac{(1.96^2)(0.5)(0.5)(9408)}{(9408)(0.05^2)+(1.96^2)(0.5)(0.5)}=\frac{9035.4432}{24.4804}=369.088\ldots$$

4. MUESTRAS PROBABILÍSTICAS

Las técnicas de **muestreo probabilística** son aquellas en las que se determina al azar los individuos que constituirán la muestra. Estas técnicas nos sirven cuando se desean generalizar los resultados que se obtienen a partir de la muestra hacia toda la población. Lo anterior se dice dado que se supone que el proceso aleatorio permitirá la obtención de una muestra **representativa** de la población.

Los muestreos probabilísticas pueden ser **con** o **sin** reemplazo.

Los **muestreos con reemplazo** son aquellos en los que una vez que ha sido seleccionado un individuo (y estudiado) se le toma en cuenta nuevamente al elegir el siguiente individuo a ser estudiado. En este caso cada una de las observaciones permanece independiente de las demás, pero con poblaciones pequeñas (un grupo de escuela de 30 alumnos, por ejemplo) tal procedimiento debe ser considerado ante la posibilidad de repetir observaciones. En el caso de poblaciones grandes no importa tal proceder, pues no afecta sustancialmente una repetición a las frecuencias relativas.

Los **muestreos sin reemplazo** son los que una vez que se ha tomado en cuenta un individuo para formar parte de la muestra, no se le vuelve a tomar en cuenta nuevamente. En este caso, y hablando específicamente para el caso de poblaciones pequeñas, las observaciones son dependientes entre sí, pues al no tomar en cuenta nuevamente el individuo se altera la probabilidad para la selección de otro individuo de la población. Para el caso de las poblaciones grandes (por ejemplo la población de un país) dicha probabilidad para la selección de un individuo se mantiene prácticamente igual, por lo que se puede decir que existe independencia en las observaciones.

Las técnicas de muestreo probabilística que mencionaremos serán básicamente tres: el aleatorio simple, el aleatorio estratificado y el sistemático.

4.2 Muestreo aleatorio simple

Podemos aquí mencionar que para el caso de que se estuviese estudiando un propoción dentro de la población (una elección de candidato, la aceptación o rechazo de una propuesta en una comunidad, la presencia o ausencia de una característica hereditaria), y el en caso de un muestreo aleatorio simple, la estimación que se puede hacer de la proporción buscada a partir de la proporción hallada en la muestra se obtiene mediante la construcción de un intervalo de confianza:

$\pi = P \pm$ tolerancia de la muestra

Donde π es la proporción buscada en la población y P es la proporción presente en la muestra.

Por otro lado, la **tolerancia de la muestra** está relacionada directamente con el nivel de confianza y se obtiene a partir de la distribución normal al igual que como se obtuvo para el cálculo del tamaño de las muestras. La representaremos con Z para obtener la fórmula:

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

4.3 Muestras aletorias

Para que las conclusiones de la teoría del muestreo y de la inferencia estadística sean validas, las muestras deben escogerse representativas de la población. El análisis de los métodos de muestreo y problemas relacionados se llaman el diseño del experimento.

4.2 Muestras no aleatorias

Cuando el método de extracción de las muestras no asegure a cada individuo de la población o del estrato, igual probabilidad de ser elegido, entonces la muestra obtenida no es aleatoria.

A veces, esto se hace por razones de practicidad en el sentido del costo o del tiempo. Si se desea tomar una muestra probabilística de la población argentina no parece razonable usar a cada individuo como unidad de muestreo. Lo mismo cuando se desea hacer un muestreo a los escolares de una provincia, es muy difícil empadronar a todos primero para luego sortear, y se tardaría demasiado para ubicarlos uno por uno hasta terminar el trabajo.

- En el *muestreo de etapas múltiples* se utiliza para el caso de grandes poblaciones humanas.

Acá, la unidad de muestreo en la primera etapa son los departamentos de cada provincia. Se los lista y se hace un primer sorteo para la selección. En una segunda etapa se distingue la población rural de la urbana, subdividiendo en fracciones (diferentes superficies con densidad de población semejante). Otra vez se sortea para elegir, y se continúa con otra división en radios dentro de las fracciones, segmentos dentro de radios, y así sucesivamente. La razón es repartir equitativamente el trabajo del encuestador.

- En el *muestreo por conglomerados* se eligen conjuntos donde naturalmente se agrupan los individuos. Es, por ejemplo, el caso de las escuelas para hacer un muestreo alumnos en el sistema educativo, o las facultades para los universitarios. Si se trata de estudiar las condiciones laborales de los empleados de comercio que trabajan en supermercados, primero se empadronan a los lugares naturales de trabajo (supermercados), y luego se sortea entre estos conglomerados para elegir a uno. Luego se entrevista a todos los empleados del supermercado elegido, y se acepta esto como una muestra representativa del sector.

- El *muestreo sistemático* se usa para el caso de sucesiones de elementos. Por ejemplo, el caso de las historias clínicas de pacientes, certificados de nacimiento, tarjetas de catálogo en una biblioteca, etc. Son los casos donde la información está en archivos y hay que trabajar con estos para obtenerlas. Se elige una cifra entera, razonable, tomando en cuenta el tamaño de la muestra y el de la población. Por ejemplo, hay que tomar una muestra de tamaño 25 de un archivo que contiene 488 fichas; luego, el cociente entre población y muestra es 488 /25, aproximadamente 19. Notar que si se elige 20 el tamaño muestral no llega a 25. Entonces, se cuentan las fichas y a llegar a la décimo novena se la extrae, se sigue hasta la número 38 que será la segunda escogida, y así sucesivamente hasta tener las 25 fichas necesarias. Es también el caso de los soldados que se numeran de 1 en adelante y cada 5 (u otro número cualquiera) dan un paso al frente. Es un método sencillo y rápido de selección.

Hay otros casos de muestreo no aleatorios de uso común. Como el de tomar una lista de nombres, cerrar los ojos y con la punta de un lápiz marcar a uno de ellos, para escogerlo. Son los casos de los programas de TV donde se toma la guía telefónica, se la abre en una página cualquiera y se escoge a uno de los números que figuran, para luego hacer un llamado con premio. En los juegos infantiles se hace una ronda con los participantes, se vendan los ojos del que va a elegir, se le hace dar varias vueltas con los ojos vendados para que marque a alguno. Todos los casos vistos tienen una característica común: *los individuos de la población no son equiprobables*.

4.4 Números Aletoricos

Una forma para obtener una muestra *representativa* es mediante el muestreo aleatorio, de acuerdo con el cual, cada miembro de la población tiene la misma probabilidad de ser incluido en la muestra. Un método para lograrlo es asignarle a cada uno un número, escribir cada número en una papeleta, y realizar en una urna un soporte justo en ella. Un método alternativo consiste en recurrir una tabla de *números aleatorios.*

4.5 Sistemático

Es análogo al anterior, aunque resulta más cómoda la elección de los elementos. Si hemos de elegir 40 elementos de un grupo de 600, se comienza por calcular el cociente 600/40 que nos dice que existen 40 grupos de 15 elementos entre los 600. Se elige un elemento de salida entre los 15 primeros, y suponiendo que sea el k-simo, el resto de los elementos serán los k-simos de cada grupo. En concreto, si el elemento de partida es el número 6, los restantes serán los que tengan los números: 15+6 ,2x15+6,......,39x15+6

Este procedimiento simplifica enormemente la elección de elementos, pero puede dar al traste con la representatividad de la muestra, cuando los elementos se hayan numerados por algún criterio concreto, y los k-simos tienen todos una determinada característica, que haga conformarse una muestra no representativa.

4.6 Estratificado

A veces nos interesa, cuando las poblaciones son muy grandes, dividir éstas en sub-poblaciones o estratos, sin elementos comunes, y que cubran toda la población.

Una vez hecho esto podemos elegir, por muestreo aleatorio simple, de cada estrato, un número de elementos igual o proporcional al tamaño del estrato.

Este procedimiento tiene la gran ventaja de que se puede obtener una mayor precisión en poblaciones no homogéneas (aunque en este curso no estudiaremos los métodos necesarios)

Si decidiéramos hacer una encuesta sobre la incidencia del tabaco en nuestro centro, podríamos razonar de la siguiente forma:

Nuestro centro tiene 2000 alumnos, 720 en 3º de ESO, 700 en 4º de ESO, 340 en 1º de Bachillerato, y 240 en 2º de Bachillerato.

Si deseamos tomar una muestra de 100 alumnos, para analizar la incidencia del tabaco en la adolescencia, bastaría tomar un número igual de alumnos de cada estrato, es decir 25.

Si embargo, si lo que se quiere es hacer una encuesta para conocer la opinión que tiene el alumnado sobre una medida que ha tomado el Consejo Escolar, es más representativo elegir de cada estrato, y en número proporcional a su tamaño, los elementos que compondrán la muestra. Si 3º de ESO representa al 36% del alumnado, el 36% de la muestra (es decir 36 alumnos) se elegirán de este estrato por muestreo aleatorio simple, 35 para 4º de ESO, y así hasta completar los 100 elementos de la muestra.

5. MUESTRA CON Y SIN REPOSICION

Si sacamos el número de una urna, podemos volverlos en ella o no, antes de la siguiente extracción. En el primer caso, ese número puede salir de nuevo mas veces, mientras que en el segundo pueda salir cada numero una vez. Estos dos tipos de muestras se llaman, respectivamente, Muestras con reposición y muestra sin reposición

Las poblaciones son finitas o infinitas. Si por ejemplo, sacamos 10 bolas sucesivamente, sin reposición, de una urna que contiene 100 bolas, estamos tomando muestra de población finita; mientras que si lanzamos 50 veces una moneda contamos el numero de caras, estamos ante una muestra población infinita.

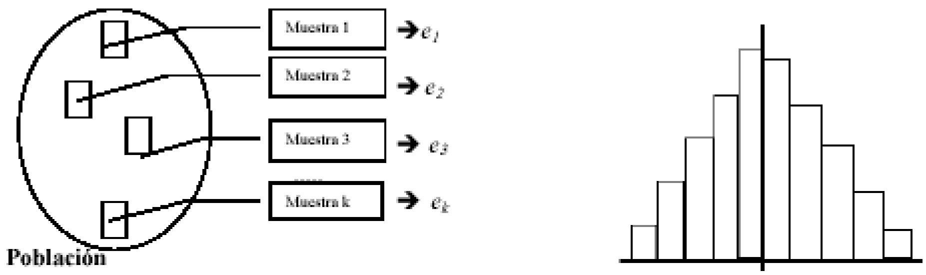
Una población finita en la que se efectúa muestra con reposición, puede considerarse infinita teóricamente, ya que puede tomar cualquier numero de muestras sin agotarla. Para muchos efectos prácticos, una población muy grande se puede considerar como si fuera infinita.

6. DISTRIBUCIONES DE PROBABILIDAD EN EL MUESTREO

También llamadas *distribuciones muestrales* de un estadígrafo cualquiera obtenido a través de la muestra. La idea es la siguiente: si se toman *k* muestras, *todas las posibles* de tamaño n (con o sin reemplazamiento) de una población de tamaño *NP*, y a cada muestra se le calcula un estadígrafo *e* (media, mediana, varianza, etc.), se obtienen una serie de *k* valores: *e1, e2, e3 , ...*, *ek*

Estos valores pueden agruparse mediante un histograma de frecuencias para poder apreciar la forma de la distribución de los mismos. En la Figura 10.1 se esquematiza esta situación:

Figura 6.1 Distribuciones muestrales.



Población

De una población cualquiera se extraen k muestras; cada una permite calcular k estadígrafos con los cuales se puede hacer un histograma como el de la derecha de la Figura 10.1. Se aprecia que este histograma adquiere forma de campana si se suavizan los escalones, al achicar los intervalos.

Esta curva obtenida a partir de datos muestrales, observados a través del muestreo, tiende asintóticamente a otra curva teórica a medida que k aumenta, y los intervalos se hacen infinitesimales.

Dicha curva teórica es la función de Gauss de acuerdo con el Teorema Central del Límite, el principal de la Estadística.

El Teorema Central del Límite permite establecer que, en condiciones muy generales, si la muestra es lo *suficientemente grande*, la distribución teórica de los k valores obtenidos es aproximadamente la función de Gauss. Esta es la base de la *Teoría de las Grandes Muestras*. Las principales *Distribuciones Muestrales* son funciones de Gauss identificadas en forma unívoca con sus dos parámetros μ y SE. En la Tabla 10.1 se presentan estos dos valores para cada uno de los estadígrafos más usuales. En la primer columna de la tabla se muestra cada estadígrafo, en la segunda columna se da la fórmula para el cálculo del error típico de estimación SE. Finalmente en la tercera columna se muestra la estimación puntual para obtener el valor esperado del estadígrafo μe, con las aclaraciones respecto al tamaño muestral requerido para que tal estimación sea considerada aceptable.

7. DISTRIBUCIÓN MUESTRAL DE LA MEDIA

Si el estadígrafo elegido es la media, se tendrán .X1 ,.X2 ,.X 3 ,...,.Xk medias muestrales; estas se distribuyen normalmente si *k* es muy grande. En la práctica, 30 o más valores son suficientes.

En la teoría, cuando $k \to \infty$ entonces la distribución muestral de la media es *asintóticamente normal* y coincidirá con la función de Gauss. Esta distribución tendrá un valor esperado y una varianza que permitirán estimar los respectivos valores poblacionales. O sea,

$$\mu x = \mu \quad \sigma 2 x = \sigma 2 / n = SE2(x) = VAR(x)$$

Tabla 7.1. Errores típicos para algunas distribuciones muestrales

Distribución muestral	Error típico	Notas especiales
Medias	$SE(\overline{x}) = \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$	Se cumple para muestras grandes o pequeñas. La distribución muestral de medias se ajusta mucho a una normal para $n \geq 30$ incluso para poblaciones no normales. $\mu_{\overline{x}} = \mu$ en todos los casos.
Proporciones	$SE(\pi) = \sigma_{\pi} = \sqrt{\frac{\pi(1-\pi)}{n}}$	Las notas anteriores para medias son igualmente aplicables aquí con $n > 25$. $\mu_P = \pi$ en todos los casos.
Desviaciones típicas	(*1*) $SE(DS) = \sigma_{DS} = \frac{\sigma}{\sqrt{2n}}$ (*2*) $SE(DS) = \sigma_{DS} = \sqrt{\frac{\mu_4 - \mu_2^2}{4n\mu_2}}$	Para $n \geq 100$, la distribución muestral de *s* es muy próxima a una normal. σ_s esta dada por (*1*) solamente cuando la población es normal (o aproximadamente normal). Si la población no es normal, puede utilizarse (*2*). Nótese que (*2*) pasa a ser (*1*) cuando $\mu_2 = \sigma^2$ y $\mu_4 = 3\sigma^4$, lo que se cumple para poblaciones normales. Para $n \geq 100$, $\mu_s = \sigma$ con gran aproximación.
Medianas	$SE(Med.) = \sigma_{med} = \frac{1{,}2533\ \sigma}{\sqrt{n}}$	Para $n \geq 30$, *la* distribución muestral de la mediana es muy próxima a una normal. Los resultados dados son válidos solamente si la población es normal (o aproximadamente normal) $\mu_{med} = \mu$
Cuartiles primero y tercero	$SE(Q2) = \sigma_{Q2} = \sigma_{Q3} = \frac{1{,}3626\ \sigma}{\sqrt{n}}$	Las notas anteriores para medianas son igualmente aplicables aquí. μ_{Q1} y μ_{Q3} son casi iguales al primero y tercer cuartil de la población. Nótese que $\sigma_{Q2} = \sigma_{med}$
Varianzas	(*1*) $SE(Var) = \sigma_{Var} = \sigma^2\sqrt{\frac{2}{N}}$ (*2*) $SE(Var) = \sigma_{Var} = \sqrt{\frac{\mu_4 - \mu_2^2}{n}}$	Las notas para desviaciones típicas son igualmente aplicables aquí. Nótese que (2) pasa a ser (1) en caso de que la población sea normal. $\mu_s^2 = \sigma^2 (n-1) / n$ que es casi igual a σ^2 para valores grandes de n.
Coeficiente de variación	$SE(CV) = \sigma_{CV} = \frac{CV}{\sqrt{2n}}\sqrt{1+2(CV)^2}$	Aquí $CV = \sigma / \mu$ es el coeficiente de variación de la población. Los resultados dados son válidos para poblaciones normales (o aproximadamente normales) y $n \geq 100$.

Esto es: la media aritmética de las k medias muestrales obtenidas es aproximadamente igual a la media poblacional (o valor verdadero). Sin embargo, esta aproximación tiene un error de estimación denominado *error típico* o *error estándar* de estimación que en el caso de la media

es: σ x . En la bibliografía clínica la nomenclatura más empleada es SE(x). En la Tabla 10.1 se muestran los valores anteriores para el caso de la media aritmética.

Las relaciones anteriores son válidas solo si la población es infinita, o si es finita, pero el muestreo es con reemplazamiento. Caso contrario, cuando la población es finita y se realizan muestreos sin reposición, entonces se deben ajustar dichas relaciones con:

$$\mu x = \mu \sigma 2 x = (\sigma 2 / n) [(NP - n) / (NP - 1)] = SE2(x) = VAR(x)$$

En el Cuadro 10.1 siguiente se presenta un problema de aplicación para un caso donde se conoce a toda la población, los parámetros poblacionales se calculan directamente aplicando las fórmulas vistas en el Tema 4 resultando: $\mu = 4,5$ y $\sigma 2 = 1,25$. Se pueden verificar las relaciones anteriores de dos maneras. En la primera se toman las seis muestras posibles de tamaño 2, para un muestreo sin reemplazamiento. A cada muestra se le calcula su media respectiva, luego con estos 6 promedios se pueden calcular: el promedio y la varianza de esas seis muestras. Ahora, el promedio de todas las medias muestrales da exactamente igual al valor medio poblacional y la varianza de las medias muestrales verifica la relación anterior, si se aplica el factor de corrección para muestras de tamaño finito. La segunda manera (*Bootstrap procedure*) es tomando muestras con reposición, primero se toman todas las 16 muestras posibles con reemplazamiento de tamaño 2. Luego, con las dieciséis muestras se calculan las 16 medias respectivas. Por último, se calcula el promedio y la varianza de estos 16 valores, verificando de nuevo las relaciones vistas más arriba, para el caso de muestras con reposición. Para el otro problema, se suponen conocidos los valores poblacionales, y tomando 50 muestras de tamaño 3 hay que determinar la cantidad de casos donde el resultado esté comprendido en un intervalo (6 ; 7,796). La forma de proceder es calculando primero las probabilidades de obtener esos resultados límites, luego por diferencia calcular

la probabilidad gaussiana asociada al intervalo y entonces, multiplicando dicha probabilidad por el tamaño muestral, se puede contestar la pregunta efectuada.

CUADRO 7.2: Distribución Muestral de Medias.

Caso 1) Un jefe de laboratorio quiere determinar cuántos Calcios realiza su personal en una hora de trabajo. Tiene a cuatro bioquímicos trabajando y los resultados obtenidos son:

Bioquímico	*Cantidad*	
A	4	Si considera a estos datos como su población
B	6	esta se representa con el conjunto (3 , 4 , 5 , 6)
C	5	Donde los parámetros poblacionales serán
D	3	$\mu = 4{,}5$ y $\sigma^2 = 1{,}25$ ($\sigma = 1{,}118$)

Si se toman todas la muestras posibles, sin reemplazamiento, de tamaño 2, se tendrán 6 pares en total, que son las combinaciones de 4 elementos tomados de a 2 (C = 4! / 2! . 2!) O sea,

Combinaciones	*Resultados*	*Media muestral*
(A , B)	(4 , 6)	5
(A , C)	(4 , 5)	4,5
(A , D)	(4 , 3)	3,5
(B , C)	(6 , 5)	5,5
(B , D)	(6 , 3)	4,5
(C , D)	(5 , 3)	4

El promedio de estas medias es

$$\frac{5 + 4{,}5 + 3{,}5 + 5{,}5 + 4{,}5 + 4}{6} = 4{,}5$$

valor igual al poblacional

Esto es $\mu_{\overline{X}} = 4{,}5 = \mu$

Y la varianza de estos valores es $\sigma^2_{\overline{X}} = 0{,}417 = (1{,}25 / 2)\,[(4\text{-}2) / (4\text{-}1)]$ verificando la relación vista más arriba. Si se toman todas las muestras con reemplazamiento, hay 16 casos:

(AA) (AB) (AC) (AD) (BA) (BB) (BC) (BD) (CA) (CB) (CC) (CD) (DA) (DB) (DC) (DD)

O sea, las muestras son:

(4;4) (4;6) (4;5) (4;3) (6;4) (6;6) (6;5) (6;3) (5;4) (5;6) (5;5) (5;3) (3;4) (3;6) (3;5) (3;3)

Y las medias de las muestras respectivas son:

4 ; 5 ; 4,5 ; 3,5 ; 5 ; 6 ; 6,5 ; 4,5 ; 4,5 ; 5,5 ; 5 ; 4 ; 3,5 ; 4,5 ; 4 ; 3

Luego el promedio de este conjunto de 16 medias es $\mu_{\overline{X}} = 4{,}5 = \mu$. Y la varianza de estos valores es $\sigma^2_{\overline{X}} = 0{,}625 = 1{,}25 / 2 = \sigma^2 / n$. Las relaciones se cumplen para ambos casos de muestreo, tanto si es con o sin reposición.

El método del *jacknife* consiste en sacar secuencialmente cada dato, calcular la media del grupo y luego reinsertarlo a los restantes para formar cada muestra. Para la población (4, 6, 5, 3) sería:

Muestras	*Medias*	*Promedio general*
(6, 5, 3)	4,6666	
(4, 5, 3)	4	4,5
(4, 6, 3)	4,3333	
(4, 6, 5)	5	

Notar que se puede "mirar" la estabilidad de la media que acá varia entre 4 y 5.

Ejemplo: Suponiendo que las proteínas totales se distribuyen normalmente entre los 2.500 varones sanos de una cierta población, con edades de 14 a 50 años, con un valor promedio de 7 g/dl y un desvío de 0,9 g/dl. Un investigador decide tomar 50 muestras aleatorias de tamaño 3, con reemplazamiento. Desea saber en cuántas muestras cabe esperar una media entre 6 y 7,796 g/dl.

El primer paso para resolver este problema es encontrar los parámetros de la distribución muestral de medias, usando los valores poblacionales conocidos. Entonces:

μ_x . 7 g/dl = μ ; σ^2_x . $\sigma^2 / n = (0,9)^2 / 3 = 0,27$; $\sigma_x = SE(x) = 0,52$

Cabe esperar que la distribución muestral de medias sea una curva de Gauss con tales parámetros.

Para calcular la probabilidad de que una muestra tenga una media entre 6 y 7,796 g/dl, se tipifican estos valores con:

$$z_6 = (x - \mu_x) / \sigma_x = (6 - 7) / 0,52 = - 1,92$$

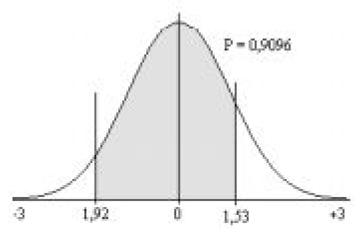
$$z_{7,6} = (x - \mu_x) / \sigma_x = (7,796 - 7) / 0,52 = 1,53$$

La probabilidad bajo la curva de Gauss tipificada dentro del intervalo - 1,92 ; + 1,53 se obtiene calculando las probabilidades:

P (-1,92 ; 0) = 0,4726

P (0 ; 1,53) = 0,4370

P (-1,92 ; 1,53) = 0,9096



Entonces, el número esperado de muestras buscado será la proporción anterior, multiplicada por el número total de muestras extraídas. Esto es, (0,9096) 50 = 45,48 muestras.

Es de esperar encontrar 45 muestras con valores de proteínas totales entre 6 y 7,6 g/dl.

8. DISTRIBUCIÓN MUESTRAL DE PROPORCIONES

Cuando el estadígrafo obtenido de cada muestra de una población sea la proporción de ocurrencia de un suceso dado, un “éxito”, se trata de una *distribución muestral de proporciones*.

Por ejemplo, el caso del lanzamiento de una moneda es uno de población infinita porque se puede realizar la prueba cuantas veces se desee. La probabilidad de sacar un éxito, una cara es p =

0,5, la misma para todos los casos posibles, mientras que la probabilidad de “fracaso” al lanzar la moneda y salir seca es q = 1 - p = 0,5. Si el experimento consiste en realizar **n** lanzamientos de la moneda, una cantidad lo suficientemente grande, luego la frecuencia relativa del suceso cara, tendrá una distribución aproximadamente normal de parámetros:

$\mu_p = \pi$ y σ^2

$\pi = [\pi .(1- \pi)] / n = SE^2(\pi)$

O bien

$\mu_r = r = n . \pi$ y σ^2

$r = n . [\pi .(1- \pi)] = SE^2 (r)$

CUADRO 8.1: **Distribución muestral de proporciones.**

Ejemplo) En un laboratorio farmacéutico, históricamente se ha encontrado que el 3% de las ampolletas fabricadas de cierto medicamento salen falladas. Se desea saber cuánto vale la probabilidad de encontrar r = 20 defectuosas o menos, en una partida de n = 500 producidas.

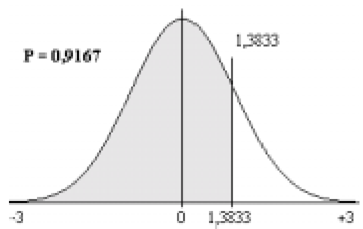
(a) La probabilidad exacta se puede calcular con la fórmula de la Binomial acumulada:

$$P_{Bi}(r \leq 20) = P_{Bi}(r = 0) + P_{Bi}(r = 1) + \ldots + P_{Bi}(r = 20) = 0{,}9202$$

(b) El segundo método es usando la aproximación normal a la binomial con la relaciones siguientes: $\mu_r = n . \pi = 15$; $\sigma^2_r = n\,[\pi .(1- \pi)] = (500\ .0{,}03.\ 0{,}97) = 14{,}55$; $SE(r) = \sigma_r = 3{,}8144$

Como se trata de una variable discreta y la función de Gauss es continua, hay que efectuar una corrección por continuidad de Yates o de Molenaar:

$z_{20} = [(4\ .\ 20 + 3)(1-0{,}03)]^{1/2} - [(4\ .\ 500 - 4\ .\ 20 - 1)(0{,}03)]^{1/2} = 1{,}3833$ O sea $P = 0{,}9167$



La probabilidad pedida es el área a la izquierda de $Z_{20} = 1{,}3883$ en la curva Gauss tipificada.

Con la corrección de Yates sería: $z = [(20 - 15) - 1/2] / 3{,}8144 = 1{,}1797$ y $P = 0{,}881$ O sea una aproximación peor al valor exacto dado en (a). Notar que si no se hubiese hecho la corrección por continuidad, la estimación de la probabilidad pedida sería peor aún: $P = 0{,}8168$.

Las que se obtienen de la relación vista más arriba con reemplazamiento de $\pi = r / n$. Por muestra lo suficientemente grande, se entiende en la práctica $n > 25$. Por otra parte, cuando la población es finita y el muestreo se realiza sin reposición, entonces las ecuaciones anteriores se modifican con:

$$\mu_p = \pi \text{ y } \sigma^2_\pi = [\pi .(1- \pi) / n]\ [(NP - n) / (NP - 1)] = SE^2(\pi)$$

O bien

$$\mu_r = r = n . \pi \text{ y } \sigma^2_r = n\ [\pi .(1- \pi)]\ [(NP - n) / (NP - 1)] = SE^2(r)$$

Usando la distribución de Gauss, se pueden calcular las probabilidades asociadas

Tanto la distribución de medias como la de proporciones se usan muy frecuentemente en la práctica. El problema básico es conocer los valores poblacionales porque es muy raro saberlos de antemano, salvo el caso de la moneda y otros juegos de azar; lo común es ignorar estos parámetros. Entonces, usando los valores obtenidos del muestreo aleatorio se pueden aproximar los valores desconocidos poblacionales, como se verá en el tema *inferencia estadística*.

9. DIFERENCIAS DE MEDIAS Y DE PROPORCIONES

Muchas veces es necesario comparar dos poblaciones, como cuando se testea un tratamiento en pacientes contra un control o blanco; o bien, cuando se comparan dos sistemas de medición entre sí (estos sistemas pueden ser: técnicas de laboratorio, marcas comerciales, instrumentos de diferente marca, etc.). El problema es ver si hay diferencias entre ambos casos. El método es encontrar un estadístico para la diferencia (o suma) de ambos estadísticos muestrales. Se tienen dos poblaciones, se efectúa un muestreo aleatorio de tamaño n1 en la primera de ellas, y se obtienen los valores de un estadígrafo cualquiera: *e1*. La población gaussiana de la distribución muestral tendrá los parámetros μ_{e1} y σ_{e1}. Análogamente, se procede con la segunda población y

se obtienen los parámetros μ_{e2} y σ_{e2}. Se pueden combinar ambas poblaciones a través de un estadígrafo que sea la diferencia de ambos: $e = e1 - e2$. Este nuevo índice también tendrá una distribución gaussiana, pues la diferencia de dos funciones de Gauss es otra función gaussiana, siempre y cuando las muestras no dependan unas de otras; esto es, sean *independientes*. Sus parámetros son:

$$\mu_{e1-e2} = \mu_{e1} - \mu_{e2} \qquad \text{y} \qquad \sigma^2_{e1-e2} = \sigma^2_{e1} + \sigma^2_{e2} = SE^2\,(e_1 - e_2)$$

Si el estadígrafo es la media muestral $e = .x$ entonces, la *distribución muestral de la diferencia de medias* para poblaciones infinitas, de parámetros (μ_1 ; σ_1) y (μ_2 ; σ_2). Entonces:

$\mu_{\overline{x}_1-\overline{x}_2} = \mu_1 - \mu_2$ Es el valor esperado de la diferencia de medias y su error de estimación es:

$$SE^2\,(\overline{x}_1-\overline{x}_2) = \sigma^2_{\overline{x}_1-\overline{x}_2} = (\sigma^2_{\overline{x}_1} + \sigma^2_{\overline{x}_2}) = [(\sigma^2_1 / n_1) + (\sigma^2_2 / n_2)]$$

Si el estadígrafo es una proporción $e = \pi$ entonces, la *distribución muestral de la diferencia de proporciones* para poblaciones infinitas,de parámetros (μ_1 ; σ_1) y (μ_2 ; σ_2). Luego:

$$\mu_{P1 - P2} = \mu_1 - \mu_2 = \pi_1 \,.\, \pi_2$$

es el valor esperado de la diferencia de proporciones poblacionales.

$$\mu_{e1-e2} = \mu_{e1} - \mu_{e2} \text{ y } \sigma^2_{e1-e2} = \sigma^2_{e1} + \sigma^2_{e2} = SE^2\,(e1 - e2)$$

Es el error estándar al cuadrado de tal estimación para la diferencia de proporciones.

En el Cuadro 10.3 se presentan dos ejemplos de aplicación de lo visto. En el primer caso se trata de una diferencia de medias usando la vida útil de dos medicamentos, mientras que en el segundo caso se trata de una diferencia de Sensibilidades de dos métodos clínicos (tomando como proporciones a las Sensibilidades) para poder compararlos. En ambos casos las probabilidades encontradas son muy chicas lo que indica diferencias entre las medias muestrales. Para poder decidir si tales diferencias son válidas se emplea la Teoría de la decisión estadística

CUADRO 9.1 Diferencia de 2 muestras. μ

Caso 1) La vida útil de un medicamento fabricado por el Laboratorio A es de 1.400 días, con una desviación estándar de 200 días. Por su parte, el mismo medicamento fabricado por el Laboratorio B de la competencia tiene una duración de 1.200 días con un desvío de 100 días. Se eligen 125 medicamentos de cada Laboratorio con un muestreo al azar. Calcular la probabilidad que los del Laboratorio A duren 250 días más que los del B.

$\mu_{\overline{x}_A - \overline{x}_B} = \mu_A - \mu_B = 1400 - 1200 = 200$ días

$\sigma^2_{\overline{x}_A - \overline{x}_B} = (\sigma^2_{\overline{x}_A} + \sigma^2_{\overline{x}_B}) = [(\sigma^2_A / n_A) + (\sigma^2_B / n_B)]$ Como $n_A = n_B = 125$

$\sigma^2_{\overline{x}_A - \overline{x}_B} = [(200)^2 + (100)^2] / 125 = 400$ O sea: $SE(\overline{x}_A - \overline{x}_B) = \sigma_{\overline{x}_A - \overline{x}_B} = 20$ días

La tipificación de esta diferencia se hace con: $Z_{A\text{-}B} = (250 - 200) / 20 = 2{,}5$

Entonces, la probabilidad pedida es igual al área a la derecha de este valor en la curva de Gauss. $P = (0{,}5 - 0{,}4938) = 0{,}0062$. Lo que significa que hay un 0,6% de probabilidad porcentual de que la vida útil del medicamento fabricado por A dure 250 días más que el de su competencia.

Caso 2) La Sensibilidad de una prueba clínica es del 85% si se usa la marca A, y del 80% si es la marca B. Se toman 100 muestras al azar para cada caso, y de desea saber la probabilidad que la diferencia de sensibilidades entre ambas sea del 10 % o más. Se estima $\pi_1 \approx 0{,}85$ y $\pi_2 = 0{,}8$

(a) $\mu_{A\text{-}B} = \mu_A - \mu_B = 0{,}85 - 0{,}80 = 0{,}05$ y $SE_{(A\text{-}B)} = \sigma_{A\text{-}B} = 0{,}05362$ pues:

$\sigma^2_{A\text{-}B} = [\pi_1 (1-\pi_1) / n_1] + [\pi_2 (1-\pi_2) / n_2] = (1/100) [(0{,}85 \,.\, 0{,}15) + (0{,}8 \,.\, 0{,}2)] = 0{,}002875$

Si se hace la Ho de que no hay diferencia entre ambas la variable tipificada es:

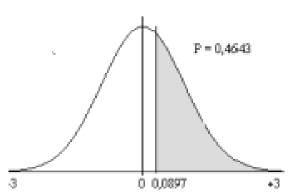
$Z_{A\text{-}B} = [(0{,}1 - 0) - (0{,}85 - 0{,}80)] / 0{,}05362 = 0{,}09325$ O sea $P(x < 10) = 0{,}5371$ y la probabilidad pedida se calcula con $P (X \geq 10) = (1 - 0{,}5371) = 0{,}4629$

(b) Otra forma de resolver este problema es hacer el supuesto de que las muestras provienen de dos poblaciones diferentes, la primera con una probabilidad de 0,85 y la otra con 0,80.
En ese caso se puede suponer que la proporción $\pi = 0{,}05$ es la diferencia real entre ambas poblaciones y que tiene un desvío $\sigma = 0{,}05362$. En n = 200 muestras el valor esperado se calcula con $\mu_r = 200 \,.\, 0{,}05 = 10$ y se pide hallar la probabilidad de que $r \geq 10$. O sea $[1 - P_{Bi} (r \leq 9)]$

Calculando el valor tipificado con la corrección de Molenaar resulta:

$z = [(4 \,.\, 9 + 3)(1 - 0{,}05)]^{1/2} - [(4 \,.\, 200 - 4 \,.\, 9 - 1)(0{,}05)]^{1/2} = 0{,}0897$

O sea $P = 0{,}5357$ y la probabilidad pedida es $(1 - 0{,}5357) = 0{,}4643$



El área a la derecha del valor tipificado es $p = 0{,}4643$ O sea, existe una probabilidad del 46,43% que la marca A tenga una sensibilidad mayor del 10% a la de la marca B.

(c) Utilizando los mismos supuestos que en el punto anterior, pero trabajando con la probabilidad exacta, la respuesta se calcula con:

$P_{Bi} (x > 10) = 1 - P_{Bi} (x \leq 9) = 1 - [P_{Bi} (x = 0) + P_{Bi} (x = 1) + \ldots + P_{Bi} (x = 9) = 0{,}4547$

Notar que ambas aproximaciones normales son bastante buenas. La idea de usar la diferencia de dos poblaciones binomiales como si fuese una sola se explicará más adelante en más detalle.

Caso 1) La vida útil de un medicamento fabricado por el Laboratorio A es de 1.400 días, con una desviación estándar de 200 días. Por su parte, el mismo medicamento fabricado por el Laboratorio B de la competencia tiene una duración de 1.200 días con un desvío de 100 días. Se eligen 125 medicamentos de cada Laboratorio con un muestreo al azar. Calcular la probabilidad que los del Laboratorio A duren 250 días más que los del B.

$\mu_{\overline{X}_A - \overline{X}_B} = \mu_A - \mu_B = 1400 - 1200 = 200$ días

$\sigma^2_{\overline{X}_A - \overline{X}_B} = (\sigma^2_{\overline{X}_A} + \sigma^2_{\overline{X}_B}) = [(\sigma^2_A / n_A) + (\sigma^2_B / n_B)]$ Como $n_A = n_B = 125$

$\sigma^2_{\overline{X}_A - \overline{X}_B} = [(200)^2 + (100)^2] / 125 = 400$ O sea: $SE(\overline{X}_A - \overline{X}_B) = \sigma_{\overline{X}_A - \overline{X}_B} = 20$ días

La tipificación de esta diferencia se hace con: $Z_{A-B} = (250 - 200) / 20 = 2{,}5$

Entonces, la probabilidad pedida es igual al área a la derecha de este valor en la curva de Gauss. $P = (0{,}5 - 0{,}4938) = 0{,}0062$. Lo que significa que hay un 0,6% de probabilidad porcentual de que la vida útil del medicamento fabricado por A dure 250 días más que el de su competencia.

Caso 2) La Sensibilidad de una prueba clínica es del 85% si se usa la marca A, y del 80% si es la marca B. Se toman 100 muestras al azar para cada caso, y de desea saber la probabilidad que la diferencia de sensibilidades entre ambas sea del 10 % o más. Se estima $\pi_1 \approx 0{,}85$ y $\pi_2 = 0{,}8$

(a) $\mu_{A-B} = \mu_A - \mu_B = 0{,}85 - 0{,}80 = 0{,}05$ y $SE_{(A-B)} = \sigma_{A-B} = 0{,}05362$ pues:

$\sigma^2_{A-B} = [\pi_1 (1-\pi_1) / n_1] + [\pi_2 (1-\pi_2) / n_2] = (1/100) [(0{,}85 \,.\, 0{,}15) + (0{,}8 \,.\, 0{,}2)] = 0{,}002875$

Si se hace la Ho de que no hay diferencia entre ambas la variable tipificada es:

$Z_{A-B} = [(0{,}1 - 0) - (0{,}85 - 0{,}80)] / 0{,}05362 = 0{,}09325$ O sea $P(x < 10) = 0{,}5371$ y la probabilidad pedida se calcula con $P (X \geq 10) = (1 - 0{,}5371) = 0{,}4629$

(b) Otra forma de resolver este problema es hacer el supuesto de que las muestras provienen de dos poblaciones diferentes, la primera con una probabilidad de 0,85 y la otra con 0,80.

UNIDAD 2 TEORÍA DE PEQUEÑAS MUESTRAS

En este capítulo se presentan tres nuevos modelos estadísticos: el llamado *t* de Student, el modelo de la *Chi-cuadrado* (χ2) y el modelo *F* de Fisher. Los tres no requieren ya más del supuesto de un tamaño muestral grande. Ahora con dos o más mediciones se puede trabajar; por eso se usa la expresión *Teoría de pequeñas muestras* para este tema. El empleo de cualquiera de ellos es enteramente similar al visto en el capítulo anterior. Cambia la manera de calcular el estadígrafo de comparación y su respectiva tabla de valores críticos de la distribución muestral.

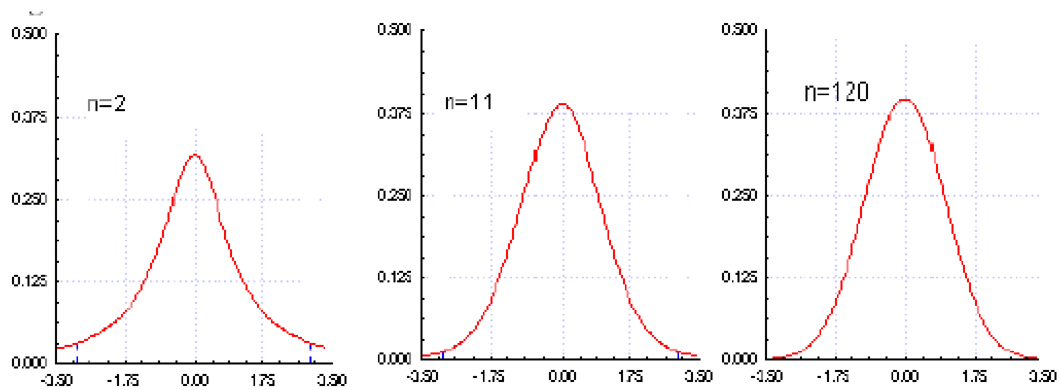
Mientras que el modelo de la *t* se aplica a medias y proporciones, los dos últimos se usan para el estudio de las desviaciones o dispersiones. También se la llama Teoría Exacta del Muestreo, pues ahora no hay que efectuar la aproximación DS2 . σ² ya que el valor muestral viene en la fórmula de cálculo del estadígrafo de comparación, en lugar del poblacional. Eso hace que no sea necesario efectuar una estimación y se tiene una mayor exactitud que con la gaussiana. Es importante destacar que los tres modelos son válidos tanto para pequeñas como para grandes muestras. Esto amplía el campo de aplicación del modelo de Gauss. Además, al no tener que hacer tantas pruebas disminuye el costo y se gana en tiempo. Todas estas ventajas tienen una contrapartida: se pierde un poco de precisión pues, como se verá, el intervalo de confianza se hace más grande para un mismo caso. Estos modelos se prefieren al de Gauss porque sus ventajas valen la pena, al precio de perder un poco de precisión. Se mostrará su empleo tanto para el caso de una sola muestra de mediciones como para la comparación de dos muestras o grupos de mediciones.

1. EL MODELO DE STUDENT

$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}}$$

Sea un estadígrafo t calculado para la media con la relación

Figura 2.1: La distribución de Student



Si de una población normal, o aproximadamente normal, se extraen muestras aleatorias e independientes y a cada una se le calcula dicho estadígrafo usando los valores muestrales de la media

y el desvío estándar, entonces se obtiene una distribución muestral *t* que viene dada por la fórmula de Student. En realidad, fue obtenida por R. A. Fisher y la bautizó *Student* en honor a W. S. Gosset, quien usaba ese seudónimo para poder publicar sus trabajos en la revista *Biometrika.* Esta función matemática tiene un parámetro que la define en forma unívoca: el número de grados de libertad υ= *n*-1 (donde *n* es el tamaño muestral). El concepto matemático de υ está relacionado con la cantidad de observaciones independientes que se hagan y se calcula con el tamaño muestral *n,* menos la cantidad *k* de parámetros poblacionales que deban ser estimados a través de ellas. O sea: υ = *n*.k. Si se observa la ecuación superior, se ve que el único parámetro poblacional

que figura es μ, por lo tanto k = 1 y así resulta υ = *n*.1. Cuando el tamaño muestral es mayor que 30 la distribución de Student se aproxima mucho a la de Gauss, en el límite ambas son iguales.

Es decir que la función Student tiende asintóticamente a la función de Gauss.

Para cada grado de libertad hay una tabla de valores que pueden obtenerse variando el nivel de significación, parecida a la de Gauss. Sería muy engorroso tener una hoja con la tabla para cada grado de libertad. Esto se soluciona de dos formas: una es usando computadoras para resolver los cálculos (programas estadísticos como Mini-Tab, SPSS, Statistica, Excel, etc.). La otra y más común, es preparar una tabla donde en cada fila se coloquen encolumnados los valores críticos más usuales para cada valor de grados de libertad. Como interesan únicamente los valores pequeños, se listan correlativamente de 1 a 30 y luego algunos como 40, 60, 120 e ∞. Este último tendrá los valores vistos para la normal. Así, en una sola hoja se presentan los valores útiles para el empleo de este modelo, como se muestran en el Tabla 5 del Anexo con las tablas estadísticas.

La distribución de Student, al igual que la de Gauss, es simétrica respecto al origen de coordenadas y se extiende desde – ∞ hasta + ∞. Pero a diferencia de la normal, puede adoptar diferentes formas dependiendo del número de grados de libertad. Por ejemplo, la que tiene un solo grado de libertad (n = 2 y υ = 1), se desvía marcadamente de la normal, como se puede ver en la Figura 13.1 anterior. Luego, a medida que los grados van aumentando, se acerca cada vez más, hasta igualarla en el infinito. Se puede ver esto en las tablas y en la Figura 2. Los valores críticos de la Tabla Student, para una confianza del 95 % y dos colas, para 1, 5, 10, 30 y ∞ grados de libertad son 12,71; 2,57; 2,23; 2,04 y 1,96 respectivamente. Estos valores críticos se denotan con sus dos parámetros así: tα ; υ = t0,05 ; ∞ = 1,96 = zα.

Los intervalos de confianza para esta distribución se arman en forma análoga a la vista para el caso de Gauss. Con la única diferencia en cómo se calcula el valor crítico tα;υ en lugar de zα.

$$\boxed{\mu \in (\mu_e \pm t_{\alpha;\upsilon}\ SE(e)) = (\mu_e \pm t_{\alpha;\upsilon}\ \sigma_e)}$$ *Intervalo de confianza con Student*

De nuevo, el par de valores (μ e ; σe) se saca de la Tabla 4, con la salvedad que ahora no se usa más la aproximación DS . σ; pues en el cálculo de *t* se emplea DS directamente. Esto hace que el modelo sea más exacto que el de Gauss. Generalmente, este modelo se aplica al caso de la media, proporciones y sus diferencias o sumas. Para una estimación con 30 o más grados de libertad, se pueden usar tanto el modelo de Gauss, como el de Student. El intervalo es casi igual, salvo que en este último el valor crítico es mayor. En efecto, si se tienen 31 muestras, t = 2,09, mientras que z = 1,96. Esto hace mayores a los intervalos obtenidos con Student que sus equivalentes gaussianos. Por eso, se dice que el modelo Student tiene menor *precisión* que el de Gauss.

La teoría de decisiones se usa en forma análoga, empleando los intervalos de confianza visto más arriba. Pero para poder aplicar este modelo se deben tener en cuenta los requisitos siguientes:

1) Las muestras fueron extraídas de una población normal o aproximadamente normal.
2) La selección de las muestras se hizo en forma aleatoria.
3) Las muestras son independientes entre sí.

Si alguno de ellos no se cumple, las conclusiones que se obtengan no son válidas. Los supuestos se pueden resumir así: *para poder usar Student, se deben tener muestras normales, aleatorias e independientes.* Notar que el error estándar de estimación es SE (*e*)= σe.

Los casos más frecuentes en la práctica son:

- **Student para medias muestrales**

En este caso *e* =.x luego: μe = μ y SE (*e*) = σe = DS / n . Por lo tanto el valor de comparación se calcula con:

$$\boxed{t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}}}$$

Ejemplo 1) Se desea saber si un instrumento de medición cualquiera está calibrado, desde el punto de vista de la exactitud. Para ello se consigue un valor patrón y se lo mide 10 veces (por ejemplo: una pesa patrón para una balanza, un suero control para un método clínico, etc.). Suponiendo que el resultado de estas mediciones arroja una media de 52,9 y un desvío de 3, usando un patrón de valor 50, se debe determinar si el instrumento está calibrado y la estimación de su error sistemático, si es que se prueba su existencia (no se usan unidades para generalizar este ejemplo).

Ho : $\mu = 50$ el instrumento está calibrado en exactitud

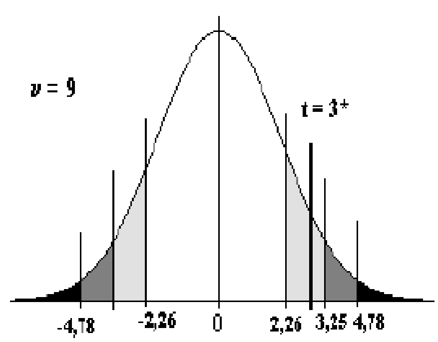
H1 : $\mu \neq 50$ no está calibrado. Hay un error sistemático

Se trata de un ensayo de dos colas donde hay $\nu = 10 - 1 = 9$ grados de libertad. De la Tabla 4 se obtienen los valores críticos para el 95% de $t_{0,05;9} = 2,262$, para el 99% de $t_{0,01;9} = 3,25$ y para un nivel del 99,9% es $t_{0,001;9} = 4,781$. Lo que permite establecer las zonas de aceptación y rechazo:

$$t = \frac{(52,9 - 50)}{3/\sqrt{10}} = 3^{*} \quad (p = 0,00135)$$

$$\mu = 50 \quad \notin \quad 95\% \ CI(50,8\ ;\ 55,1)\ ; \quad \text{o bien}$$

$$\bar{x} = 52,9 \quad \notin \quad 95\% \ CI(47,9\ ;\ 52,1)$$



Dibujando las zonas con los valores críticos, el valor de t cae en la de rechazo para el

95% y no alcanza para las otras. La conclusión es que se ha probado la existencia de un error sistemático con una confianza del 95%. Y se estima con:

$$ES \in (\bar{x} - \mu) = 2{,}90$$

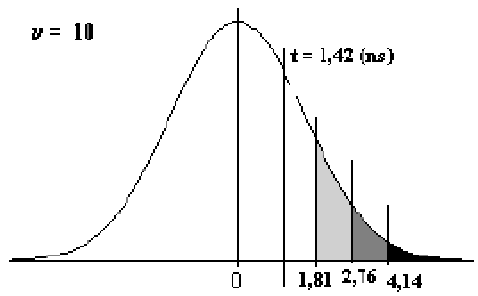
Ejemplo 2) Se midió colesterol total a 11 pacientes varones adultos escogidos al azar los resultados obtenidos arrojan una media de 235 mg/dl y un desvío estándar de 35 mg/dl. Ensayar la hipótesis de que se mantienen por debajo del valor límite de referencia (220mg/dl).

H_0 : $\mu \leq 220$ mg/dl
H_1 : $\mu > 220$ mg/dl

El valor de Student para una sola cola es:

$$t = \frac{(235 - 220)}{35 / \sqrt{11}} = 1{,}42 \text{ (ns)} \quad (p = 0{,}093)$$

Valor no significativo pues $t_{0,05;\,10} = 1{,}81$
$\mu = 220$ cae dentro de 95% CI (216 ; 254)



Para el caso de una cola, el valor de tablas para el 95% debe ser el que está en la Tabla 4 para el 90% en dos colas. La idea es que el 10% en dos colas significa el 5% en cada una, por la simetría de la curva de Student. Luego, para ν = 10, el límite para el 95% será t = 1,812 en una cola y t = 2,228 para dos colas. En la figura de más arriba se han marcado los límites del 99% y del 99,9% para una sola cola, a los efectos didácticos. La conclusión es que no puede rechazar la hipótesis nula, por lo que debe considerarse un colesterol total admisible desde el punto de vista clínico, por estar por debajo del límite de referencia.

- **Student para proporciones**

En este caso *e* = P y μ*p* = μ = π lueg $SE(e) = \sigma_\pi = \sqrt{\pi.(1-\pi)/n}$
se puede obtener el valor del estadígrafo de comparación con la relación:

$$\boxed{t = \frac{(P - \pi)}{\sqrt{\pi.(1-\pi)/n}}}$$

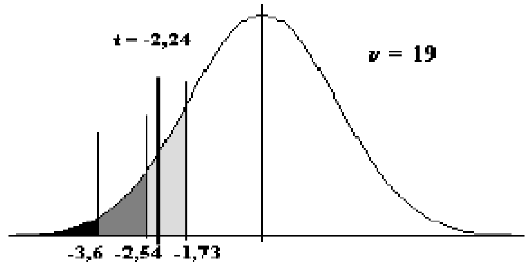
Ejemplo1) Un analgésico de plaza, afirma en su propaganda que alivia el dolor en el 90% de los casos antes de la primer hora luego de su ingesta. Para validar esa información, se hace un experimento en 20 individuos con cefalea. Se observa que fue efectivo en 15 de ellos.

H_0 : $\mu \geq 0,9$ La afirmación es correcta
H_1 : $\mu < 0,9$ La afirmación es falsa

Es un ensayo de una sola cola (a izquierda)
El porcentaje de éxitos es: P = 15 / 20 = 0,75.
Con $\mu = 0,9$ y $\sigma_p = \sqrt{0,9 \,.\, 0,1 / 20} = 0,067$

$$t = \frac{(0,75 - 0,9)}{0,067} = -2,24^* \quad (p = 0,019)$$

$\mu = 0,9$ cae fuera del CI 95% (- ∞ ; 0,78)



De tablas: $t_{0,999;\,19} = -3,579$ $t_{0,99;\,19} = -2,539$
y $t_{0,95;\,19} = -1,729$

El resultado obtenido es significativo (t = - 2,24 *). Pero la evidencia no alcanza para rechazar la hipótesis a los niveles del 99% y 99,9%. Se la rechaza al nivel de 95% únicamente. Si bien no es tan terminante, se puede afirmar que la aseveración es falsa con un 95% de confianza.

- **Student para dos muestras independientes**

El modelo de Student también se puede usar cuando se desean comparar dos muestras entre sí, para detectar si hay diferencia significativa entre ellas, debido a algún factor analizado. En primer lugar se analizará el caso de dos muestras independientes como: aplicar dos tipos de remedios a dos grupos de pacientes escogidos al azar, o las mediciones repetidas de una misma magnitud, etc. El otro caso, cuando las muestras no son independientes sino apareadas, se verá en el próximo tema. Una vez más, los supuestos para poder aplicar este modelo se resumen en: *para poder comparar con Student, las dos muestras deben ser normales, aleatorias e independientes.*

Se sacan muestras aleatorias e independientes, de dos poblaciones normales. La idea es averiguar si ambas muestras provienen de la misma población o de poblaciones diferentes. Con eso se puede ver si el efecto de los “tratamientos” aplicados a las muestras es *apreciable,* en cuyo caso las muestras parecerán provenir de diferentes poblaciones. Se usa en los casos donde se compara el efecto de una droga aplicada a un grupo de pacientes, contra otro grupo al cual se le suministra un placebo. También para comparar dos técnicas clínicas y detectar si hay diferencias, por ejemplo: dos marcas comerciales de plaza, dos instrumentos de medición, dos individuos, dos técnicas diferentes (la nueva contra la vieja), dos protocolos, etc. Con estas comparaciones se pueden realizar muchos controles internos en el laboratorio para hacer calibraciones, medir eficacia, etc. Hay una limitación: *solo se pueden comparar dos muestras entre sí a la vez y no más*. Para el caso de tener más de dos muestras, se recurre a los modelos de ANOVA.

- **Comparación de medias**

Para estos casos, el valor de Student para validaciones de medias se calcula con:

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{DS_1^2}{n_1}\right) + \left(\frac{DS_2^2}{n_2}\right)}}$$

El cual se contrasta con tα; υ donde υ = n1 + n2 - 2 grados de libertad. Hay casos particulares como (a) las muestras son de igual tamaño y (b) son *homocedásticas* (tienen igual varianza). En ambos casos se simplifican las fórmulas de cálculo.

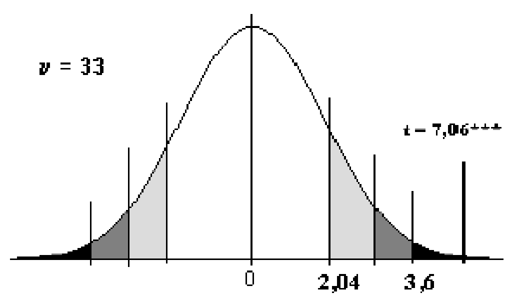
Ejemplo 1) Se aplica un medicamento a 15 pacientes que padecen cierta enfermedad, escogidos al azar, y un placebo a 20 pacientes. En el primer grupo, la desaparición del estado febril se observa a las 19 horas de tratamiento en promedio (con un desvío de 2 hs.). En el grupo control, la mejoría se observa en promedio las 25 horas con un desvío de 3 horas. Decidir si el medicamento modifica el tiempo de curación.

H_0 : $\mu_1 = \mu_2$: el medicamento es inocuo
H_1 : $\mu_1 \neq \mu_2$: el medicamento produce efecto

Es un ensayo de dos colas donde los valores críticos se buscan en la Tabla 5 interpolando entre 30 y 40 grados de libertad.

$$t = \frac{(25 - 19) - (0)}{\sqrt{\left(\frac{9}{20}\right) + \left(\frac{4}{15}\right)}} = 7{,}06^{***}$$

$\mu_1 - \mu_2 = 0$ cae fuera de CI 99,9% (3,1; 8,9).



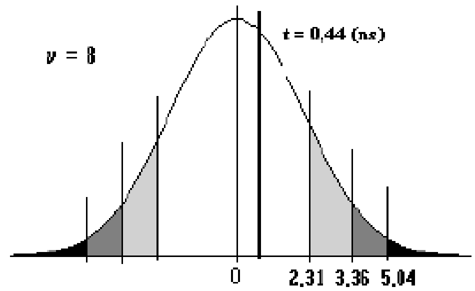
Como el valor hallado de t es mucho más grande que el valor crítico de tablas para 33 grados de libertad: tα ; υ = t 0,999; 33 = 3,44 (ensayo de dos colas y un 99,9% de confianza), la conclusión es: se obtuvieron resultados altamente significativos (t = 7,06 ***) como para rechazar la hipótesis nula. Se tiene una prueba científica del efecto del medicamento.

Ejemplo 2) Se desea verificar si hay diferencia en las mediciones a través de dos métodos clínicos diferentes. Se toma una muestra de suero lo suficientemente grande como para obtener 10 alícuotas. Se distribuyen al azar 5 alícuotas para cada método. Efectuadas las mediciones, con el primero se tuvo una media de 85 mg/dl con un desvío de 8 mg/dl. Mientras que con el segundo se tuvo una media de 83 mg/dl con un desvío de 6 mg/dl.

H_0 : $\mu_1 = \mu_2$: no hay diferencia.
H_1 : $\mu_1 \neq \mu_2$: hay diferencia

$$t = \frac{(85-83)-(0)}{\sqrt{\left(\frac{64}{5}\right)+\left(\frac{36}{5}\right)}} = 0{,}44 \text{ (ns)}$$

No se puede rechazar la Ho. Se concluye que no hay diferencia entre ambos métodos.
$\mu_1 - \mu_2 = 0$ cae dentro de 95% CI(-8,3 ; 12,3)



- **Comparación de proporciones**

Para estos casos, el valor de Student para validaciones de proporciones se calcula la misma fórmula, pero reemplazando los valores esperados con

$$\mu_{1\text{-}2} = (\pi_1 - \pi_2) \qquad \text{y} \qquad \sigma^2_{1\text{-}2} = [\pi_1(1-\pi_1)/n_1] + [\pi_2(1-\pi_2)/n_2]$$

Entonces el valor de comparación del modelo Student para este caso es:

$$\boxed{t = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{(\pi_1(1-\pi_1)/n_1) + (\pi_2(1-\pi_2)/n_2)}}}$$

Contrastando con el valor de tablas dado por tα;υ; con υ= n1 + n2 -2 grados de libertad.

Ejemplo) Se escogen al azar dos grupos formados por 20 individuos cada uno, entre los que padecen cierta alergia. Se administra una droga curativa al primer grupo y se observa una mejoría en 15 de los casos. Al segundo grupo se le administra un placebo y mejoran 13 de ellos. Ensayar la hipótesis que la droga sirve para curar ese tipo de alergia. Se emplean las hipótesis siguientes:

H0 : μ1-2 = 0 las diferencias observadas se deben al azar. H1 : μ1-2 ≠ 0 la droga produce efecto. Si se supone que ambas muestras fueron extraídas de la misma población, y por lo tanto no hay diferencias entre las muestras observadas (H0) μ1-2 = 0, eso significa que el porcentaje de curados en dicha población será π = π1 = π2 y habrá que estimarlo con los datos muestrales, calculando la proporción ponderada con: *p* = (total de curados en las muestras / total muestral) = (15+13) / 40 = 0,7. Entonces, sacando factor común en la fórmula de la varianza, esta resulta: SE2(π) = π (1.π).[ ) / 1 ( ) / 1 (2 1 *n n* +] = π (1.π) [2 / *n*] =(0,7 . 0,3) (2/20) = 0,021

Y es SE(π) = 0,145; de los datos del problema surgen P1 =15/20 = 0,75 y P2 = 13/20 = 0, 65 t = (0,75 – 0,65) / (0,021)1/2= 0,69 < t0,95 ; 38=2,02. μ1-2 = 0 cae dentro de 95% CI(-0,19 ; +0,39) Un resultado no significativo. Las diferencias observadas no se deben a la droga sino al azar.

- **Test de equivalencia biológica**

Hay ocasiones donde la Ho no busca establecer si hay o no diferencia entre dos muestras, como las del ejemplo anterior, sino que se trata de establecer si un método clínico o tratamiento nuevo es lo suficientemente bueno como para reemplaza al que se venía usando hasta entonces, el método viejo. Las ventajas de este nuevo método pueden ser: un costo menor, más rápido, menos dañino o peligroso para el paciente, etc. La cuestión básica aquí es ver si, en promedio, la diferencia entre ambos es menor que un cierto valor límite para la magnitud estudiada. Es decir que tal diferencia no implique una inferioridad del nuevo método, desde un punto de vista clínico. Para estos casos la Ho : La diferencia entre ambos promedios es mayor o igual al valor aceptable y la alternativa es H1 : Esta diferencia de medias es menor al valor crítico; en cuyo caso ambos métodos pueden ser considerados clínicamente equivalentes. La idea es que, si se rechaza la Ho se puede usar el método nuevo en lugar del viejo y aprovechar las ventajas que este posee. Pero la decisión se basa más en consideraciones médicas que estadísticas. Entonces, si se trata de magnitudes continuas, se puede usar el test de Student para comparar la diferencia de las dos medias contra el valor crítico δ o máximo aceptable desde el punto de vista clínico. El planteo se hace así: Ho : $\mu V - \mu N = . > \delta$. Donde μV es el valor poblacional que se obtiene con el método viejo y μN con el método nuevo, . es la diferencia real entre ambos métodos y δ es la diferencia máxima admisible entre ambos métodos. De esta manera, cuando Ho pueda rechazarse se tendrá evidencia suficiente como para efectuar el reemplazo, esto es cuando H1 : $\mu V - \mu N = . < \delta$.

Se trata de un ensayo de una sola cola. Pero cuando se trate de ver si en valor absoluto la diferencia entre ambos métodos no supere a un cierto valor δ, porque aquí no interesa tanto que sea menor, sino que también interesa que no sea mayor (dependiendo de la magnitud clínica analizada); entonces la Ho será : $\mu V - \mu N = . = \delta$ y el ensayo será de dos colas. Análogo al visto en el punto anterior. Para ilustrar este procedimiento se usará un ejemplo tomado de la obra de Armitage

Ejemplo) Sea el índice cardíaco CI (respuesta cardiaca normalizada para la superficie del cuerpo) el cual se mide con un procedimiento invasivo como es el colocar un catéter en el corazón del paciente llamado Termo-dilución (el método viejo) y la unidad de medición son litros por minuto tomado por m2 de superficie del cuerpo humano. Se ha propuesto una nueva manera de medir esa magnitud con una técnica no invasiva, llamada el método de la Bioimpedancia, en la cual se le adosa un instrumento al cuerpo de paciente en forma externa, y mide en forma eléctrica el valor del CI usando una escala adecuada (el método nuevo). El criterio clínico de aceptación es: el nuevo método se considerará equivalente al viejo cuando, en promedio, el valor obtenido difiera en un 20% respecto al promedio aceptado de 2,75 l / min. / m2 para el método del catéter. Esto significa que el 20% de tal valor es δ = 0,55. Luego el planteo se hace así:

Ho : $.\mu V - \mu N . = ... > \delta = 0{,}55$ o lo que es lo mismo $(\mu V - \mu N) = \delta = 0{,}55$

H1 : .μV – μN . = ... < δ = 0,55 cuyo equivalente es (μV – μN) = δ ≠ 0,55

Se toma una muestra de N = 96 individuos a los cuales se le aplica el método nuevo, los valores encontrados fueron un promedio de 2,68 l / min. / m2, y un desvío estándar de 0,26 l / min. / m2 luego será:

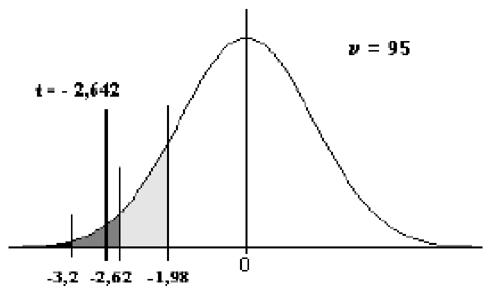
$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}} = (2{,}68 - 2{,}75) / (0{,}26 / \sqrt{96}) = -\,0{,}07 / 0{,}0265 = -\,2{,}642^{**} > t_{0,99\,;\,95} = -\,2{,}62$$

Lo que indica que hay evidencia significativa como para rechazar la Ho.

Otra manera de ver lo mismo es cuando se usa el valor límite de confianza, para la diferencia en valor absoluto, la cual es: 0,07. Para un 95% de confianza el error de estimación se calcula con:

$$t_{0,95\,;\,95} \,.\, (DS / \sqrt{N}) = 0{,}053$$

Luego 0,07 + 0,053 = 0,123 < 0,55 con lo que se puede rechazar la Ho.



La conclusión final es que se puede usar el método nuevo en lugar del viejo, con una gran ventaja para el paciente, pues ahora ya no tendrá que ser cateterizado para efectuarle su medición del índice cardíaco. A este procedimiento estadístico aparecido en los últimos años en Medicina se lo conoce también con el nombre de *test de equivalencias médicas o biológicas*.

- **Student para dos muestras apareadas**

El modelo de Student se puede usar para el caso especial de muestras *apareadas*, esto es, cuando se le efectúan dos tratamientos a la misma muestra; por ejemplo, del tipo *antes–después* donde al mismo individuo se lo mide dos veces para ver el efecto del tratamiento realizado, o el caso de método nuevo contra el método viejo, donde al mismo grupo de pacientes se le hacen dos mediciones a cada uno, la del método de rutina habitual y una extra con el nuevo método a probar para decidirse entre ambos. La idea básica es como sigue: se sacan n muestras aleatorias e independientes de una población normal. A cada muestra se le aplican dos “tratamientos” A y B diferentes y lo que interesa detectar es si producen algún efecto apreciable. Este caso es muy diferente al anterior si bien las muestras son independientes entre sí, los tratamientos no lo son, porque a un mismo individuo se le aplican ambos tratamientos. Entonces, la misma persona aparecerá dos veces en los resultados: uno en el grupo A y el otro en el grupo B. El truco para resolver este problema de la independencia es trabajar con la diferencia de los resultados de cada par de mediciones efectuadas: d = xA - xB. Luego se tendrán n diferencias d1;

d2; d3...dn, que son independientes entre sí, puesto que cada valor di corresponde a un solo individuo. Luego, se le aplica el modelo Student para una sola muestra, ensayando la hipótesis de que no hay diferencias entre ambos grupos. O sea, efectuando la hipótesis: H0 : □d □□0 resultará:

$$t = \frac{\bar{d} - 0}{DS_d / \sqrt{n}}$$

La hipótesis alternativa implica un efecto diferente para cada grupo H1 : μd ≠ 0. Si se prueba que el valor esperado del promedio de las diferencias es diferente de cero, entonces el tratamiento aplicado produce un efecto demostrable. Para aclarar estas ideas se presenta el siguiente caso:

Ejemplo) Se escogen 5 pacientes al azar, del grupo que concurre diariamente al Laboratorio de Análisis Clínicos a efectuarse una determinación de Uremia. Las muestras extraídas se miden con el procedimiento habitual y además con una nueva técnica clínica que se desea probar. Ver si hay diferencia entre ambas técnicas. Los resultados expresados en g/l fueron:

Pacientes	Vieja	Nueva	Diferencias
1	0,38	0,33	0,05
2	0,54	0,45	0,09
3	0,22	0,15	0,07
4	0,11	0,09	0,02
5	0,23	0,22	0,01

Promedio = 0,048

Desvío estándar = 0,033

Con los valores de las diferencias se calculan $\bar{d}$ = 0,048 g/l y DS_d = 0,033 g/l. Luego:

$$t = \frac{0{,}048}{0{,}033 / \sqrt{5}} = 3{,}25^* > t_{0,95\,;\,4} = 2{,}776.$$ O bien, el valor 0 cae fuera del CI 95% (0,01; 0,09)

Se obtuvo evidencia significativa t = 3,25 * que hay diferencia entre ambas técnicas.

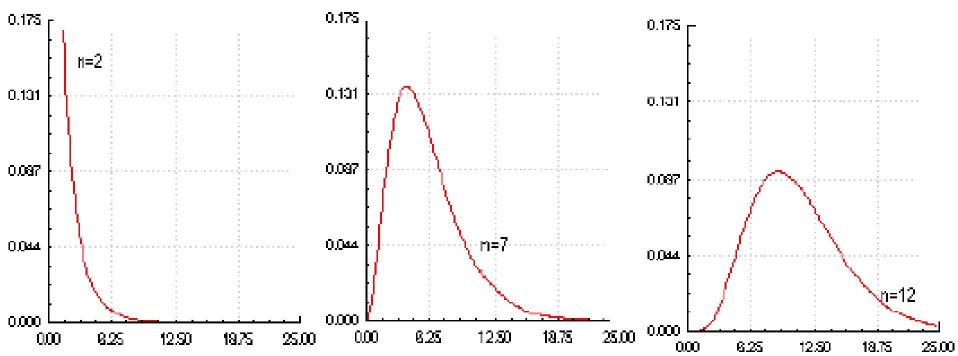
3. EL MODELO DE CHI-CUADRADO

La función Chi-cuadrado es igual a la función normal elevada al cuadrado. Esto es, el producto de dos distribuciones de Gauss es una distribución de Chi-cuadrado. Si de una población normal, o aproximadamente normal, se extraen muestras aleatorias e independientes, y se le calcula el estadígrafo χ^2 usando el valor muestral de la varianza y el poblacional con:

$$\boxed{\chi^2 = (n-1)\ DS^2 / \sigma^2}$$

La distribución muestral de χ2 viene dada por la fórmula de K.R. Pearson. Esta función matemática está caracterizada por el valor del número de grados de libertad υ = n - 1 (donde n es el tamaño muestral). Al igual que la Student, el valor total del área bajo la curva es igual a la unidad, pero la diferencia principal es que esta no es simétrica respecto al origen, sino que se extiende desde 0 hasta + ∞ porque no puede ser negativa. A medida que los grados de libertad aumentan, la curva cambia de forma y sus valores se han tabulado en el anexo de tablas estadísticas (Tabla 6), donde se muestran los valores del área bajo la curva, para los principales valores de χ2 , a la derecha de éste. O sea, se muestra la zona de rechazo para diferentes niveles de significación y de grados de libertad, lo cuales varían entre 1 y 100. Más allá, conviene usar directamente la función de Gauss.

Figura 3.1: La distribución de Chi-cuadrado



Para cada grado de libertad hay una tabla de valores que pueden obtenerse variando el nivel de significación, parecida a la de Gauss. El problema de calcular los valores críticos, para un nivel de confianza dado, se resuelve de dos maneras: usando computadoras para resolver los cálculos, y la otra más común, usando tablas resumidas en una sola hoja como la que se muestra en la Tabla 6, en forma análoga a la vista para el modelo de Student. La distribución de χ2 se usa principalmente para analizar dispersiones. Se compara la dispersión muestral expresada a través de sus cuadrados medios (MS) contra la dispersión poblacional cuantificada a través de la varianza (σ2). El valor MS = (n – 1) DS2 = v DS2 es otra forma de mostrar la precisión del sistema de medición. El uso más difundido en Bioquímica es para controlar la dispersión de la técnica de Análisis Clínicos empleada en el Laboratorio; se compara la obtenida en forma experimental, contra un valor considerado como aceptable en los libros de texto denominado la *dispersión máxima admisible* (σmáx). Usualmente, se toma el CV%, como se

Mostró en el ejemplo de varianzas visto en el capítulo anterior. Existen otros criterios, como el de Thonks, que usa un error relativo admisible máximo, y se calcula como un cuarto del rango de los valores normales de referencia, dividido por el valor medio de dicho intervalo (referido a la magnitud clínica en cuestión y expresado en porcentajes). Todo esto se puede ver con mayor detalle en la Tabla 24.1 del Capítulo 24. También se emplea a este modelo para realizar la llamada *prueba de chi-cuadrado* en las comparaciones de frecuencias observadas contra las frecuencias esperadas, con datos de recuento. Más adelante se desarrolla mejor este tema, lo mismo que su so para testear la independencia de dos o más factores en una Tabla de Contingencia.

En la industria farmacéutica se la usa para analizar la dispersión de los componentes de los productos terminados. Todo remedio fabricado debe cumplir estrictas normas de calidad, generalmente referidas al contenido en peso de sus principales componentes. Se usan dos límites: el superior e inferior, dentro de los cuales se los debe mantener controlados. Este rango de valores define la dispersión máxima admisible y lo ideal es que la dispersión de los productos terminados sea bastante inferior a dicho rango. Ese control de la dispersión es muy similar al explicado más arriba, para los bioquímicos. Para ilustrar estas ideas se presenta un ejemplo en Control de Calidad de equipos de Laboratorio. Pero antes se debe remarcar que para poder aplicar este modelo, se deben tener en cuenta los requisitos siguientes:

1. Las muestras fueron extraídas de una población normal o aproximadamente normal.

2. La selección de las muestras se hizo en forma aleatoria.

3. Las muestras son independientes entre sí.

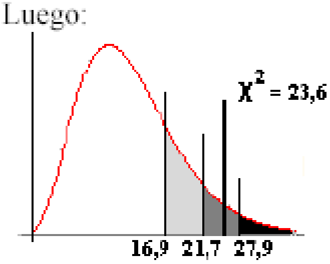
Ejemplo 1) Un bioquímico sospecha que su micro-centrífuga no mantiene constante su velocidad mientras trabaja, lo cual le da una variabilidad indeseada en sus determinaciones. Para controlarla, consigue un tacómetro regulado y mide cada minuto la velocidad durante 10 minutos. Los resultados fueron: una velocidad promedio en las 10 mediciones de 3098 rpm con un desvío de

100,4 rpm. Testear para un error relativo máximo del 2% o menos, si la centrífuga es estable.

El desvío estándar aceptable es: $\sigma_{máx}$ = 2% de 3098 rpm = 62 rpm. Luego:

$H_0 : \sigma_{máx} \leq$ 62 rpm: la micro centrífuga es estable
$H_1 : \sigma_{máx} >$ 62 rpm: la micro centrífuga no es estable

$$\chi^2 = (n-1)\ DS^2 / \sigma^2$$
$$\chi^2 = (10-1)\,(100{,}4)^2 / (62)^2 = 23{,}6^{**}$$



De la Tabla de valores críticos surge: χ2 0,99 ; 9 = 21,666 y χ2 0,991 ; 9 = 27,877. Por lo tanto, el bioquímico ha encontrado una muy fuerte evidencia que la velocidad del equipo oscila en forma indeseada, tal como sospechaba. Y deberá ajustarlo si desea disminuir la variabilidad de sus mediciones. Los resultados fueron muy significativos χ2 = 23,6**.

Ejemplo 2) Un farmacéutico Jefe del Dpto. Control de Calidad en una industria alimenticia, descubre que en su proceso de producción el contenido de ciclamato en su línea de mermeladas dietéticas varía en forma indeseada. Sospechando que se trata de una falla en el dosificador, decide tomar 10 muestras seguidas del mismo. Encuentra un promedio de 20 gramos con un desvío de 8 gramos. Si en su protocolo de fabricación la variación máxima permitida es del 3%, determinar si el dosificador debe ser corregido.

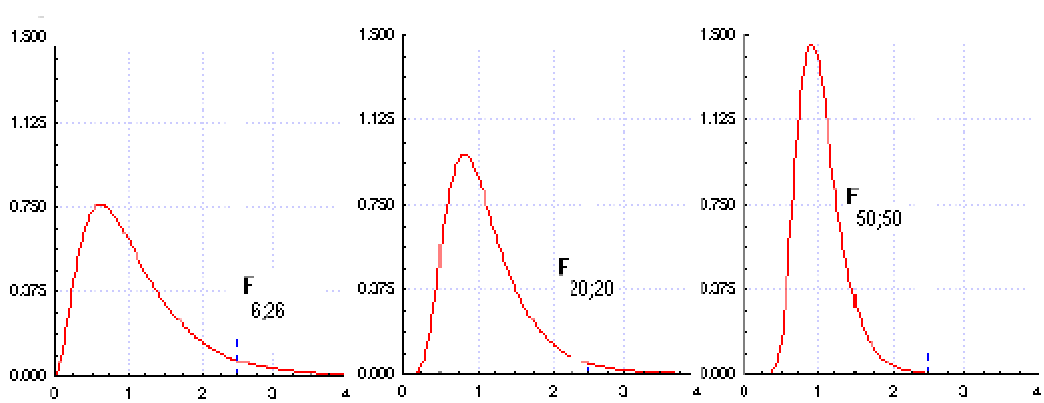
El desvío estándar aceptable es: □máx = 3% de 20 g = 6 g. Luego: H0 : □máx ≤□6 g.: el dosificador funciona correctamente H1 : □máx > 6 g.: el dosificador debe ser cambiado

□2 = (n – 1) DS2 / □2 = (10 – 1) . (8)2 / □□□2 = 16 (no significativo) De la Tabla de valores críticos surge: □2 0,95 ; 9 = 16,9. Por lo tanto, el farmacéutico no ha encontrado evidencia que respalde sus sospechas. Sin embargo, el valor hallado es muy cercano al crítico, por lo que le convendría hacer más pruebas.

4. EL MODELO DE FISHER

Si de dos poblaciones normales, o aproximadamente normales, se extraen dos muestras aleatorias e independientes, y a cada una se le calcula su respectiva varianza, el cociente de ambos valores F = DS2 1 / DS2 2 (con F > 1, esto es, siempre se coloca el más grande como numerador) tendrá una distribución de Fisher, cuyos valores críticos fueron obtenidos por W. Snedecor y se muestran en la Tabla 7 del anexo. Esta tabla se caracteriza por tener dos grados de libertad: el correspondiente al numerador υ1 = n1 - 1 y el del denominador υ2 = n2 - 1. Programas de computación permiten calcular los valores críticos respectivos. En otra forma se puede usar una tabla de doble entrada como la Tabla 7; su forma se puede ver a continuación:

Figura 4.1: La distribución de Fisher.

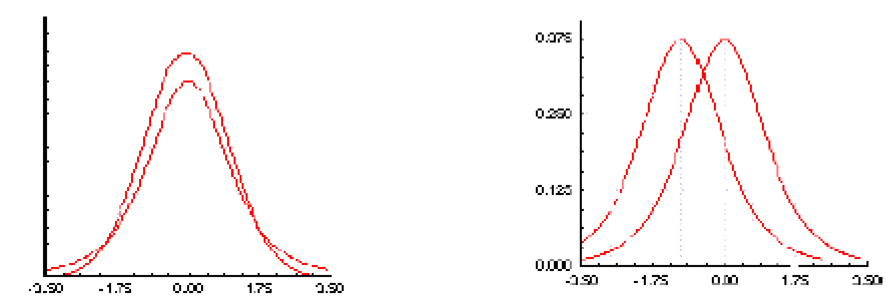


En las Tabla 7 se presenta una hoja para cada nivel de confianza, se eligen los más apropiados como: 95% ; 97,5% ; 99% ; 99,5% y 99,9%. Como siempre, el área total bajo la curva es la unidad y se extiende desde 0 a + ∞. La forma es muy parecida a la Chi-cuadrado. se muestran tres casos, con diferentes grados de libertad, y se marca el valor de F = 2,5 con una ,línea punteada vertical.

El principal uso de esta función es el Análisis de Varianza, que se verá más adelante, y es para cuando se necesita comparar más de dos medias muestrales a la vez. En estos casos la idea es detectar si el efecto de uno o más tratamientos afecta a las muestras testeadas. En cambio, cuando se tiene el caso de dos muestras, la idea es testear si hay homoscedasticidad en las dos poblaciones en estudio. Una vez verificado este supuesto, se puede avanzar más verificando si hay diferencia entre las medias muestrales, y así verificar si ambas muestras tienen igual media y varianza, porque eso significa que en realidad provienen de la misma población normal. Eso probaría que no hay efecto de un tratamiento si se lo compara con un placebo, o que dos técnicas de laboratorio son equivalentes. Si el experimento no verifica esto, entonces se deberá elegir el caso que presente menor varianza, para tener menor variabilidad en las mediciones. En Genética se puede verificar si una generación de crías es más variable en un carácter que la de sus padres. En Sistemática se puede testear si dos poblaciones locales tienen la misma variabilidad. En Bioquímica y Farmacia el uso más frecuente es comparar el error casual de mediciones de laboratorio, al introducir algún efecto o cambiar el método de medición. En el caso de testear si dos técnicas de laboratorio tienen igual dispersión, o bien, para elegir aquella con mayor precisión, conviene pensar el problema como la incidencia de un factor en estudio en lugar de dos técnicas totalmente diferentes entre sí. Por ejemplo, se trata de una misma práctica, pero se usan dos espectrofotómetros diferentes, y se trata de determinar si la modificación de la varianza se debe al uso de un aparato diferente. El factor acá sería: tipo de espectros.

También se puede estudiar la incidencia del factor humano, realizando las mismas mediciones a dos personas diferentes. De esa forma se puede imaginar que las dos muestras provienen de diferentes poblaciones, o que el efecto del factor analizado no es despreciable cuando se rechaza la hipótesis nula. En la Figura 13.4 se muestra el caso de dos poblaciones. En el caso (a) ambas poblaciones tienen la misma media, pero por efecto del error casual sus varianzas son diferentes. Si esta diferencia es significativa, resulta evidenciada por el Modelo de Fisher que permite la comparación de ambas.

Figura 4.2: Dos poblaciones con



(a) Igual media, distinta varianza

(b) Distinta media, igual varianza

En el caso (b) hay un error sistemático que desplaza la media, pero sus varianzas permanecen iguales. Es lo mismo que sumar una constante a todos los valores; ocurre un desplazamiento hacia la derecha. Student se usa para detectar esto cuando se hace el test de comparación de dos medias independientes. Como se verá más adelante, se puede construir todo un bagaje de métodos para efectuar un Control de Calidad interno en un laboratorio de medición clínica. Por ahora, basta decir que se puede controlar la exactitud con los modelos de Student y la precisión con los de Chi-cuadrado y Fisher.

Con esto se pueden comenzar a controlar y calibrar los sistemas de medición. Las limitaciones de todo esto son dos: la primera es que se puede estudiar el efecto del factor analizado en solo dos muestras y no en más de dos. La segunda es que si la calidad se entiende como exactitud y precisión, solo se pueden emplear estos modelos para magnitudes de tipo cuantitativas como las de la Química Clínica, pero no en magnitudes cualitativas como las usuales en Microbiología, Bacteriología, Micología, etc. En magnitudes cuantitativas, por “calidad “se entiende precisión y exactitud, en lugar de la capacidad de una prueba clínica para diagnosticar. Sin embargo, a pesar de estas limitaciones sigue siendo una herramienta sencilla y poderosa de control.

Para poder aplicar este modelo se deben tener en cuenta los requisitos siguientes:

1. Las muestras fueron extraídas de una población normal o aproximadamente normal.

2. La selección de las muestras se hizo en forma aleatoria.

3. Las muestras son independientes entre sí.

Ejemplo) El jefe de un laboratorio se encuentra con una técnica de medición fuera del control estadístico.

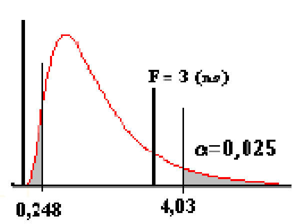
Para investigar las causas decide investigar si el factor humano tiene incidencia, y toma una muestra de suero cualquiera la divide en 20 alícuotas. Luego elige 10 de ellas al azar y se las entrega al laboratorista 1 para que haga las determinaciones; las restantes las encomienda al laboratorista 2 para que las mida. Los resultados obtenidos son: DS2 1 = 2,4 es la varianza obtenida

por el laborista, 1 y DS2 2 = 0,8 para el otro. Decidir si hay diferencia en dispersión entre ambos.

$H_0 : \sigma^2_1 = \sigma^2_2$: no hay diferencia y el factor humano no incide
$H_1 : \sigma^2_1 \neq \sigma^2_2$: hay diferencia entre ambas personas

Se calcula el estadígrafo de comparación de Fisher

$F = DS^2_1 / DS^2_2 = 2,4 / 0,8 = 3$



Como se trata de un ensayo de dos colas, para un nivel del 95% de confianza, se busca en las tablas para: $\upsilon 1 = \upsilon 2 = n1 - 1 = 9$ grados de libertad, mientras que $\alpha = 0{,}025$ para el límite inferior

y $\alpha = 0{,}975$ para el superior. Estos valores son:

$F0{,}975; (9{,}9) = 4{,}03$.

Luego, para calcular el valor no tabulado $\alpha = 0{,}025$ se aprovecha una propiedad que tiene la función F usando la inversa como sigue:

$F0{,}025; (9{,}9) = 1 / F0{,}975; (9{,}9) = 1/ 4{,}03 = 0{,}248$ Como el valor hallado $F = 3$ cae dentro de la zona de aceptación, no hay evidencia significativa como para decir que el factor humano tiene incidencia en la dispersión de las mediciones.

- *Cuadro resumen*
- *Test de hipótesis para las medias muestrales.*

Se puede siempre usar el modelo Student. El de Gauss cuando la muestra sea grande ($n > 30$)

Muestra vs. Población

- σ^2 conocido $\quad Z = \dfrac{(\bar{x} - \mu)}{\sigma/\sqrt{n}} \quad$ vs. $\quad Z\alpha$
- σ^2 desconocido $\quad t = \dfrac{(\bar{x} - \mu)}{DS/\sqrt{n}} \quad$ vs. $\quad t_{\alpha;\upsilon}$ con $\upsilon = n - 1$

Muestra vs. Muestra:

<u>Caso 1</u>: Muestras independientes con σ_1^2 y σ_2^2 son desconocidos

Si $\sigma_1^2 \neq \sigma_2^2$ es

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\dfrac{DS_1^2}{n_1}\right) + \left(\dfrac{DS_2^2}{n_2}\right)}} \quad \text{vs.} \quad t_{\alpha;\upsilon} \quad \text{con } \upsilon = \frac{\left[(DS_1^2/n_1) + (DS_2^2/n_2)\right]^2}{\dfrac{(DS_1^2/n_1)^2}{n_1 - 1} + \dfrac{(DS_2^2/n_2)^2}{n_2 - 1}}$$

Si $\sigma_1^2 = \sigma_2^2$ es

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}\sqrt{\dfrac{(n_1 - 1)DS_1^2 + (n_2 - 1)DS_2^2}{n_1 + n_2 - 2}}} \quad \text{vs.} \quad t_{\alpha;\upsilon} \quad \text{con } \upsilon = n_1 + n_2 - 2$$

<u>Caso 2</u>: Muestras independientes con σ_1^2 y σ_2^2 son conocidos

Si $\sigma_1^2 \neq \sigma_2^2$ es

$$Z = [\,(\bar{x}_1 - \bar{x}_2) - (\mu_{\bar{x}_1 - \bar{x}_2})\,] \,/\, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{vs.} \quad Z\alpha$$

Si $\sigma^2 = \sigma_1^2 = \sigma_2^2$ es

$$Z = [\,(\bar{x}_1 - \bar{x}_2) - (\mu_{\bar{x}_1 - \bar{x}_2})\,] \,/\, \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad \text{vs.} \quad Z\alpha$$

<u>Caso 3</u>: Las muestras son apareadas y se suponen de una misma población con $DS_d \approx \sigma^2$

Se toma $d_i = X_1i - X_2i$ entonces $t = \dfrac{\bar{d} - 0}{DS_d / \sqrt{n}}$ vs. $t_{\alpha;\upsilon}$ con $\upsilon = n - 1$

Test de hipótesis para proporciones

Se puede siempre usar el modelo Student. El de Gauss cuando la muestra sea grande (n > 25)

Muestra versus población: π conocido o estimado con $p = r / n$

Gauss: $Z = \dfrac{p - \pi}{\sqrt{\pi.(1 - \pi) / n}}$ o $Z = \dfrac{r - n.\pi}{\sqrt{n.\pi.(1 - \pi)}}$ vs. $Z\alpha$

Student: $t = \dfrac{(P - \pi)}{\sqrt{\pi.(1 - \pi) / n}}$ o $t = \dfrac{(r - n.\pi)}{\sqrt{n.\pi.(1 - \pi)}}$ vs. $t_{\alpha;\upsilon}$ con $\upsilon = n - 1$

Muestra vs. Muestra: π_1 y π_2 conocidos

$$t = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{(\pi_1(1 - \pi_1) / n_1) + (\pi_2(1 - \pi_2) / n_2)}} \quad \text{vs.} \quad t_{\alpha;\upsilon}; \text{ con } \upsilon = n_1 + n_2 - 2$$

Ambas muestras provienen de la misma población: $\pi_1 = \pi_2 = \pi$

Se estima $\pi = [(p_1 n_1) + (p_2 n_2)] / (n_1 + n_2)$ y se reemplaza en

$$t = \frac{(P - \pi)}{\sqrt{\pi.(1 - \pi) / n}} \quad \text{o} \quad t = \frac{(r - n.\pi)}{\sqrt{n.\pi.(1 - \pi)}} \quad \text{vs.} \quad t_{\alpha;\upsilon} \text{ con } \upsilon = n - 1$$

Test de hipótesis para varianzas

Muestra vs. Población: $\chi^2 = (n - 1)\ DS^2 / \sigma^2$ vs. $\chi^2_{\alpha;\nu}$ con $\upsilon = n_1 - 1$

Muestra vs. Muestra: $F = DS^2_1 / DS^2_2$ vs. $F_{\alpha;\,\upsilon 1;\,\upsilon 2}$

con $DS^2 > DS^2_2$ y $\upsilon_1 = n_1 - 1$; $\upsilon_2 = n_2 - 1$

CONCLUSION

Las razones para efectuar un muestreo a una población, en lugar de estudiarla directamente, pueden ser varias como se puntualiza a continuación:

. El tamaño de la población es infinito.

. El muestreo es de tipo destructivo.

. La población es finita, pero demasiado grande.

. Sería muy caro estudiar a toda la población y basta con deducciones aproximadas.

. Tomaría demasiado tiempo analizar la población total.

Cuando las muestras son lo suficientemente grandes, se pueden hacer inferencias analíticas bastante extensas, con pocos y simples recursos, en comparación con técnicas más refinadas de la Estadística. Esto es conveniente desde un punto de vista didáctico. La Teoría del muestreo es el estudio de las relaciones entre una población y las muestras que se extraen de ella. Del análisis de las muestras se pueden *estimar o inferir* datos de la población como su media (μ), varianza (σ2), etc., llamados *parámetros poblacionales*, denotados usualmente con letras griegas, a partir de los valores obtenidos de la muestra, tales como la media muestral x , la varianza muestral

DS2, etc. Por ejemplo, en Bioquímica el verdadero valor de glucosa de un paciente μ es siempre desconocido y además variable con el tiempo, por ello se le extrae una muestra de sangre para poder estimarlo a través de mediciones. Así, el valor x obtenido (ya sea con una o más mediciones) es el que se pone en el informe de la determinación clínica. El problema consiste en extraer muestras lo más representativas posibles, de la población desconocida, para que tengan sentido las estimaciones realizadas a través de ellas. Cuando en Farmacia las mediciones se hacen a través de encuestas; se miden los porcentajes de las respuestas obtenidas. Esto se emplea usualmente en investigaciones de mercado, técnicas de propaganda, estudios poblacionales, etc. Por ejemplo, en una encuesta sobre el uso de determinado producto cada respuesta favorable se puede considerar un éxito, y la proporción de éxitos p en el total de las encuestas realizadas se puede usar para estimar la verdadera proporción en la población tomada como marco de referencia del estudio.

Cuando la población sea finita y de un tamaño manejable en tiempo y costo, los valores poblacionales se calculan directamente, sin necesidad del muestreo. Por ejemplo, si se trata de revisar diamantes, a nadie se le ocurriría tomar muestras sino que se controlarían uno por uno. Ahora bien, cuando se efectúan mediciones de magnitudes clínicas de tipo cuantitativo, la idea teórica es que se pueden efectuar infinitas mediciones y, en tal caso, el tamaño de la población será infinito. Es el caso de mediciones repetidas de un mismo objeto como pesar un cuerpo, o medirle su longitud, etc., se puede efectuar todas las mediciones que se deseen. Esto no ocurre cuando el ensayo es destructivo para la muestra. Por ejemplo, en mediciones clínicas de calcio, colesterol, hierro, etc., porque al suero se le adicionan los reactivos químicos y sirve para una sola vez. O bien, cuando se controla la cantidad de los componentes activos en un remedio, no hay otra forma que destruirlo para poder revisarlo. Entonces, el número total de determinaciones posibles dependerá de la cantidad de material disponible, que es finito. En resumen, en estos casos se acostumbra a considerar al valor verdadero de la magnitud medida como desconocido. El único medio para estimarlo es la toma de muestras.

La Alumna

BIBLIOGRAFIA

- Fidalgo. A. (1.998) 1Var 25. Oviedo

- J.reskog, K. G. y S.rbom, D. (1.993). *PRELIS 2 User«s reference guide.* Chicago

- Mu.iz, J. (1997) *Introducci.n a la Teor.a de Respuesta a los .tems.* Ediciones Pir.mide.Madrid

- Samejima, F. (1.969) Estimation of latent ability using a response pattern of graded scores. *Psychometric Monographs,* num 17

- Van der Linden, W.J. y Hambleton, R.K. (eds.)(1.997) *Handbook of modern item response theory*, Springer-Velac, Nueva York

- Spigel, Munrray. "Estadisticas Administrativas". Editorial McGraw Hill. Décima edición. 1.998.

ANEXOS

Teoría de las Muestras de Trabajo