***matemática*:** Una publicación del ICM – ESPOL
2011, Vol. 9, No. 1

USO DE ÁRBOLES DE DECISIÓN PARA LA ENSEÑANZA DE LA ESTADÍSTICA INFERENCIAL

[1]Valdivieso Carlos, [2]Valdivieso Oscar, [3]Valdivieso Roberto

Resumen. *El artículo muestra la utilidad de los árboles de decisión para facilitar al alumno la elección del intervalo de confianza adecuado, pudiéndose estimar la media, proporción o varianza poblacionales o la diferencia de medias, proporciones o cociente de varianzas.*
Esta elección puede dificultarse mucho, ya que depende de muchos factores, entre los cuales están:

- *El número de muestras que se comparan.*
- *La normalidad de las poblaciones de las que provienen las muestras.*
- *El tamaño de las muestras.*
- *El conocimiento de las varianzas poblacionales.*
- *La dependencia de las muestras.*
- *La igualdad de las varianzas poblacionales.*

Los autores han configurado 3 árboles de decisión tomando en cuenta estos factores y otras consideraciones teóricas y empíricas:
1. Un árbol general que ayuda a establecer el intervalo de confianza adecuado para la estimación.
2. Un árbol que se usa cuando se quiere estimar la media poblacional.
3. Un árbol útil para elegir el intervalo de confianza cuando se quiere estimar la diferencia de medias.
El artículo expone el uso de los árboles de decisión con un ejemplo para ilustrar su fácil manejo, rapidez y así realzar su utilidad en la enseñanza. Una de las principales virtudes de los árboles de decisión configurados es que presentan exhaustivamente todas las opciones disponibles para realizar la estimación. Comparándolos con otros métodos de ayuda a la elección, presentan las siguientes diferencias:

- *Son más fáciles de usar y requieren de menor tiempo y espacio que los flujogramas.*
- *Son más completos que las tablas resumen.*

Este método puede ser usado para otras áreas de la estadística. Por otro lado los árboles de decisión tienen implicaciones de uso en la informática, al sistematizar el proceso de elección, y pueden ser usados para la configuración de un software estadístico más amigable que los actuales.

Palabras claves: Estadística inferencial, Árboles de decisión, Estimación estadística, Intervalos de confianza, Estadística educacional.

Abstrat. *The article shows the usefulness of decision trees to provide students with the choice of appropriate confidence interval, being able to estimate the mean, proportion or population variance or difference of means, proportions or ratio of variances.*
This choice can be a tough lot, and it depends on many factors, among which are:

- *The number of samples to be compared.*
- *The normality of the populations from samples.*
- *The size of the samples.*
- *Knowledge of the population variances.*
- *The dependence of the samples.*
- *The equality of population variances.*

The authors have set 3 decision trees taking into account these factors and other theoretical and empirical considerations:
1. A general tree that helps to establish appropriate confidence interval for the estimate.
2. A tree that is used when we want to estimate the population mean.
3. A useful tree to choose the confidence interval when you want to estimate the mean difference.
The article describes the use of decision trees with an example to illustrate its ease of use, speed and thus enhance their usefulness in teaching. One of the main virtues of the decision trees is that they have fully configured all the options available to make the estimate. Compared with other methods of assistance to the election, show the following differences:

- *They are easier to use and require less time and space that the flow diagrams.*
- *Are more complete than the summary tables.*

This method can be used for other areas of statistics. On the other hand the decision trees have implications for use in computing, to systematize the process of choice, and can be used for configuring a statistical software more friendly than at present.

Key words. Inferential statistics, decision trees, statistical estimation, confidence intervals, Educational Statistics.

Recibido: Noviembre, 2010
Aceptado: Febrero, 2011

[1] Valdivieso Taborga Carlos Eduardo, Ing., Universidad Privada Boliviana (UPB). (e_mail: cvaldivieso@upb.edu)
[2] Valdivieso Taborga Oscar Álvaro, Ing., Universidad Privada Boliviana (UPB). (e_mail: oscarvaldivieso@lp.upb.edu)
[3] Valdivieso Castellón Roberto, Ing., Universidad Privada Boliviana (UPB). (e_mail: rvaldivieso@upb.edu)

1. LA ESTADÍSTICA INFERENCIAL
1.1. CONCEPTOS

Es la rama de la Estadística que se ocupa del uso de los conceptos de la probabilidad para afrontar el riesgo y la incertidumbre en la toma de decisiones. Los conceptos básicos de las distribuciones muestrales, sirven como base para la estadística inferencial. Esta comprende dos áreas:

- *La estimación*, que consiste en estimar los valores de los parámetros de la población bajo estudio.
- *Las pruebas de hipótesis*, que constituyen el proceso de aceptar o rechazar declaraciones o supuestos, generalmente relacionadas con características poblacionales.

Las pruebas de hipótesis pueden ser a su vez:

- *Pruebas paramétricas*, si lo que se desea es probar una hipótesis acerca de un parámetro poblacional en estudio. Se debe realizar la suposición de que las muestras obtenidas por el proceso de muestreo aleatorio deben haber sido extraídas de una población normal
- *Pruebas no paramétricas*, si lo que se desea es probar una hipótesis en las cual no se requiere el uso de un parámetro poblacional y no se requiere realizar ninguna suposición acerca de la forma de la población.

A continuación se muestra en la Tabla I los principales métodos estadísticos inferenciales paramétricos y no paramétricos.

TABLA I
Uso de árboles de decisión para la enseñanza de la estadística inferencial
Métodos estadísticos inferenciales

Tipo de inferencia	Número de poblaciones		
	Una	Dos	Mayor a 2
Paramétrica	Media Varianza Proporción	Diferencia de medias Cociente de varianzas Diferencia de proporciones	ANOVA simple ANOVA de bloques
No paramétrica	Prueba de signos Prueba de bondad de ajuste	Prueba de signos Prueba U de Mann-Withney Prueba de independencia de atributos	Prueba H de Kruskal Wallis

Fuente: Elaboración propia

1.2. APLICACIÓN A LA ENSEÑANZA DE LA ESTIMACIÓN

Ahora bien, si nos ocupamos en la enseñanza de la estimación estadística, el estudiante generalmente tiene bastantes problemas para fijar un esquema mental de conocimientos que le permita analizar un problema planteado y reconocer sin ayudas adicionales (resúmenes, cuadros sinópticos, flujos de decisión, etc.) rápidamente el tipo de intervalo de confianza que deberá utilizar para la estimación, junto con el cálculo de los valores críticos que toman en cuenta distintas distribuciones de probabilidad (t de Student, distribución normal estándar z, chi-cuadrada , F de Fisher, etc.). Levin y Rubin (1996) y Berenson, Levine y Krehbiel, (2001), usan flujogramas para ayudar a la decisión de elegir el intervalo adecuado, sin embargo este método no tiene la contundencia deseada y ocupa mucho espacio para plasmar todas las alternativas. Generalmente la mayoría de los autores usan tablas resumen que no son exhaustivas (Freund y Simon, 1994; Miller, Freund y Jonson, 1992; entre otros), o ninguna ayuda adicional para la elección (Mason y Lind, 1995; Lobez y Casa, 1967; Maisel, 1973; García, 1985, entre otros). En este artículo se propone que esa ayuda requerida por el estudiante sea un árbol de decisión por la facilidad y rapidez en lograr el objetivo. Mendenhall (1990) propone el uso de un árbol de decisión, pero es muy simple y con pocas alternativas de análisis. A continuación se mostrará el uso de árboles de decisión en la estimación de parámetros poblacionales.

2. LA ESTIMACIÓN ESTADÍSTICA
2.1. INTRODUCCIÓN

En muchos experimentos, problemas o investigaciones, el estudiante quiere saber si la diferencia de medias, proporciones o varianzas entre dos poblaciones es significativa, o a veces si la media, proporción o varianza se conforman a un valor patrón; también estará interesado en conocer un intervalo en el que se espera contenga el valor de la diferencia de medias, proporciones o varianzas de las dos poblaciones; o de la media, proporción o varianza de una población; y por lo tanto, deberá usar la estimación estadística.

2.2. TIPOS DE ESTIMACIONES

Un estimador es un estadístico muestral con el cual se estima un parámetro de la población (media, proporción o varianza). Una estimación es un valor específico observado de un estadístico. En la vida real se desconocen los valores de los parámetros de las poblaciones que se estudian. Se pretende entonces estimarlos mediante información muestral. Para esto puede hacerse una:

- *Estimación puntual*, que es un valor único que pretende estimar el valor del parámetro. Dos estimadores puntuales son: la media muestral (para la media poblacional): $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, y la varianza muestral, (para la varianza poblacional): $s^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1}$.
- *Estimación por intervalo*, que es un intervalo numérico en donde se pretende encontrar el valor del parámetro bajo estudio.

2.3. CRITERIOS DE UN BUEN ESTIMADOR

- Imparcialidad, cuando, en promedio, los valores del estimador coinciden con el verdadero valor del parámetro.
- Consistencia, si al aumentar el tamaño de la muestra, se logra una seguridad casi absoluta de que el valor del estadístico se acerca mucho al valor del parámetro.
- Eficiencia, designa el tamaño del error estándar del estadístico. Cuanto menor sea, el estimador es más eficiente.
- Suficiencia, si utiliza la información contenida en la muestra, al punto que ningún otro estimador podría extraer de esta última más información referente al parámetro de la población que va a ser estimado.

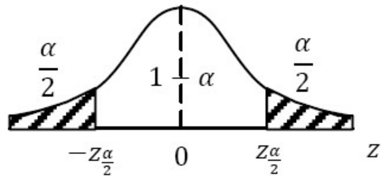
Los métodos para determinar los mejores estimadores son los de máxima verosimilitud y el de momentos (Mood/Graybill, 1976; Maisel, 1973; entre otros).

2.4. ESTIMACIÓN POR INTERVALOS

Una desventaja de los estimadores puntuales, es que no se sabe que tan cerca se encuentran del valor del parámetro que estiman. Con el fin de obtener alguna medida de la precisión en una estimación, se determina un intervalo de valores (intervalo de confianza), que incluirá el valor del parámetro con una probabilidad prefijada $(1-\alpha)$, donde "α", es la probabilidad de que el intervalo no contenga al verdadero valor del parámetro (nivel de significancia). Los intervalos de confianza pueden ser construidos utilizando el método del pivote (basado en el teorema central del límite y la ley de los grandes números). A continuación se muestra un ejemplo.

De una población normal, de la cual se quiere estimar la media, pero por alguna razón se conoce su desviación estándar, se extrae una muestra aleatoria sin importar su tamaño. La variable tipificada, para una población infinita es: $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

Sabiendo que tiene una distribución normal:



El intervalo de confianza donde se pretende hallar la media de la población está comprendido entre $-z_{\frac{\alpha}{2}}$ y $z_{\frac{\alpha}{2}}$ y $1-\alpha$ indica la probabilidad de que la media poblacional se encuentre en ese intervalo:

$$P\left(-z_{\frac{\alpha}{2}} \le z \le z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

Sustituyendo z: $P\left(-z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \le z_{\frac{\alpha}{2}}\right) = 1-\alpha$.

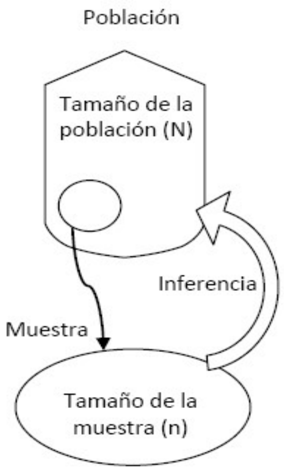
Multiplicando por el error estándar: $P\left(-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \bar{x}-\mu \le z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$. Despejando la media de la desigualdad simultánea, se tiene el intervalo deseado:

$$P\left(\bar{x}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

2.5. PROCESO DE LA ESTIMACIÓN ESTADÍSTICA

Con los conceptos anteriores, es importante que el alumno entienda el proceso que se debe seguir para realizar una estimación. Se lo presenta en el esquema 1.

ESQUEMA 1
Uso de árboles de decisión para la enseñanza de la estadística inferencial
Proceso de la estimación estadística



1. ¿Qué se quiere estimar?

$$\begin{cases} \text{Media de la población } (\mu) \\ \text{Proporción de la población (P)} \\ \text{Varianza de la población } (\sigma^2) \end{cases}$$

2. ¿Cuál es el procedimiento?

- Se determina el tamaño poblacional (N)
- Se determina el tamaño muestral (n)
- Se extrae una muestra representativa de la población:
- Se determinan los estadísticos de la muestra:

$$\begin{cases} \text{Media de la muestra } (\bar{x}) \\ \text{Proporción de la muestra } (\hat{p}) \\ \text{Varianza de la muestra } (s^2) \end{cases}$$

- A partir de los estadísticos se calcula el intervalo de confianza para estimar los parámetros poblacionales:

$$\begin{cases} \bar{x} \ para\ estimar\ \mu \\ \hat{p} \ para\ estimar\ P \\ s \ para\ estimar\ \sigma \end{cases}$$

Fuente: Elaboración propia

2.6. ÁRBOLES DE DECISIÓN

Se ha deducido un intervalo de confianza. Sin embargo, existen muchas combinaciones, ya que el intervalo depende de varias condiciones:

- ¿Cuántas poblaciones se están investigando? Pudiendo ser una o dos poblaciones.
- Si es una población, ¿qué parámetro poblacional se quiere estimar? Pudiendo existir las siguientes alternativas: Media, proporción y varianza.
- Si se quieren comparar dos poblaciones, ¿qué parámetros poblacionales se quieren estimar? Pudiendo existir las siguientes alternativas: Diferencia de medias, proporciones o cociente de varianzas.
- Si se quiere estimar la media poblacional, ¿la muestra proviene de una población normal o no?
  - Si proviene de una población normal, ¿se conoce la varianza de la población o no? Si no se conoce, ¿el tamaño de la muestra es menor a 30 o mayor que ese valor?
  - Si no proviene de una población normal, ¿se conoce la varianza de la población o no? Si se conoce, ¿el tamaño de la muestra es menor a 20 o mayor que ese valor? Si no se conoce, ¿el tamaño de la muestra es menor a 50 o mayor que ese valor?
- Si se quiere estimar la diferencia de medias poblacionales, ¿las varianzas de las poblaciones son conocidas?
- Si no son conocidas, ¿Los tamaños muestrales son menores a 30 o mayores? Si son mayores, ¿las muestras son dependientes o independientes? Si son menores, ¿las muestras son dependientes o independientes? Si son independientes, ¿las varianzas de las poblaciones son iguales o diferentes?

Con la información contenida en las preguntas anteriores, se han configurado 3 árboles de decisión que permitan sistematizar la elección del intervalo de confianza adecuado.

- Figura 1. Muestra un árbol de decisión general para elegir el intervalo de confianza adecuado al problema que se quiere resolver. Si la elección es utilizar un intervalo de confianza para la media poblacional, se debe ir a la Figura 2, y si es un intervalo de confianza para la diferencia de medias, se deberá ir a la Figura 3.
- Figura 2. Muestra un árbol de decisión para elegir el intervalo de confianza adecuado para

la estimación de la media poblacional, tanto para poblaciones normales como no normales.

- Figura 3. Muestra un árbol de decisión para elegir el intervalo de confianza adecuado para la estimación de la diferencia de medias poblacionales, para poblaciones normales.

Vale la pena analizar cuáles han sido las bases teóricas y empíricas para la configuración de la Figura 2. García (1985), Hays y Winkler (1971) y muchos otros autores, afirman que cuando el tamaño muestral es mayor a 100 se puede sustituir la curva normal por la distribución t como distribución muestral. Sin embargo también afirman que empíricamente para n>30 la media aritmética muestral presenta ya una distribución que se aproxima suficiente a la normal y que los intervalos de confianza resultantes son casi idénticos a los obtenidos utilizando la varianza verdadera. Cuando n<30 es mejor utilizar la distribución t de Student. Esto se cumple si la muestra proviene de una población normal.

Si una muestra proviene de una población que no es normal, y la varianza poblacional es conocida, el tamaño muestral que separa el uso de la distribución z o no es de 20, ya que las pruebas no paramétricas de signos de datos pareados o la U de Mann-Whitney se basan en un proceso binomial, en la que tal distribución se vuelve simétrica para una probabilidad de éxito de 0.5 y tamaño 20 (Mason y Lind, 1995; Freund y Simon, 1994; Miller, Freund y Jonson, 1992; Mendenhall, 1990, entre otros). Si la muestra es menor a 20, se deberá hacer uso del Teorema de Tchebysheff para el cálculo del valor crítico k (Hays y Winkler, 1971; Freund y Simon, 1994) (Ver pie de página de la Figura 2).

Si una muestra proviene de una población que no es normal, y la varianza poblacional es desconocida, el tamaño muestral que separa el uso de la distribución z o no es de 50. Este valor de nuevo, sale de la práctica empírica, esta vez de los autores del texto, quienes han demostrado que si se separa un estudio con tamaño muestral n en varios con tamaño 50, las conclusiones son prácticamente las mismas. Si n<50 no hay más salida que desistir de obtener un intervalo de confianza.

Cuanto mayor incertidumbre existe (es decir la población no es normal, no se conoce la varianza poblacional y el tamaño muestral es pequeño), el valor crítico k (determinado por la distribución z, t o el teorema de Tchebysheff) será más grande, y por lo tanto, el intervalo tendrá un rango mayor y poseerá menor precisión.

El estudiante podrá usar el árbol de decisión de la Figura 1 (y si es el caso remitirse a los árboles de las Figuras 2 y 3), cuando tenga que resolver cualquier problema para la estimación de un parámetro poblacional. Su uso es sencillo, sin embargo, se mostrará un ejemplo para mayor claridad.

2.7. EJEMPLO PARA EL USO DEL ÁRBOL DE DECISIÓN

Un nuevo dispositivo de filtrado se instala en una planta química. Muestras aleatorias arrojaron la siguiente información del porcentaje de impurezas, antes y después de la instalación:

Muestra/Estadísticos	Media	Varianza	Tamaño
Antes	12.5	101.17	8
Después	10.2	94.73	9

a) Establezca una forma para estimar si el porcentaje de impurezas presenta la misma variabilidad antes y después de la instalación del nuevo dispositivo de filtrado. Use $1-\alpha = 0.95$.
b) ¿El dispositivo de filtrado ha reducido el porcentaje medio de impurezas de forma significativa?
c) Si el porcentaje medio de impurezas permitido en la planta química es de 5%, ¿se llegó a cumplir la meta?

FIGURA 1

Uso de árboles de decisión para la enseñanza de la estadística inferencial

ÁRBOL DE DECISIÓN PARA ESCOGER EL INTERVALO DE CONFIANZA ADECUADO

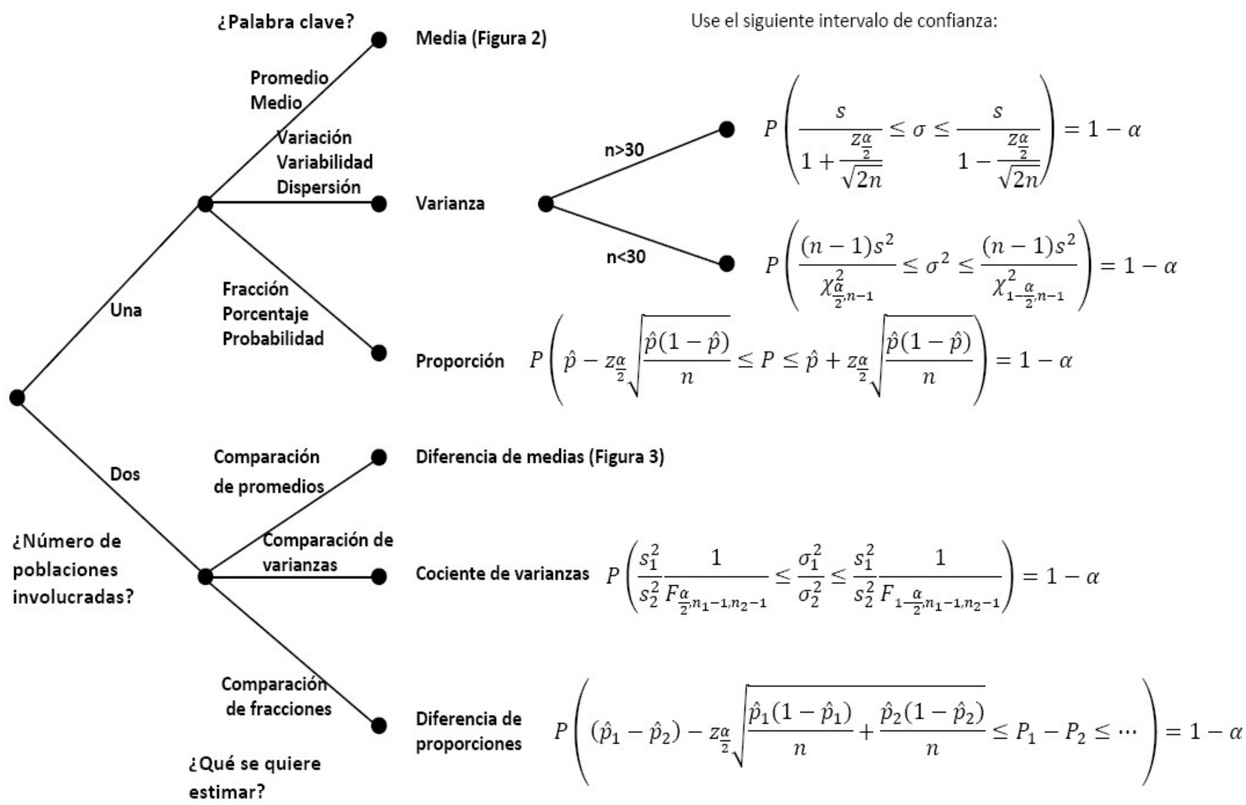
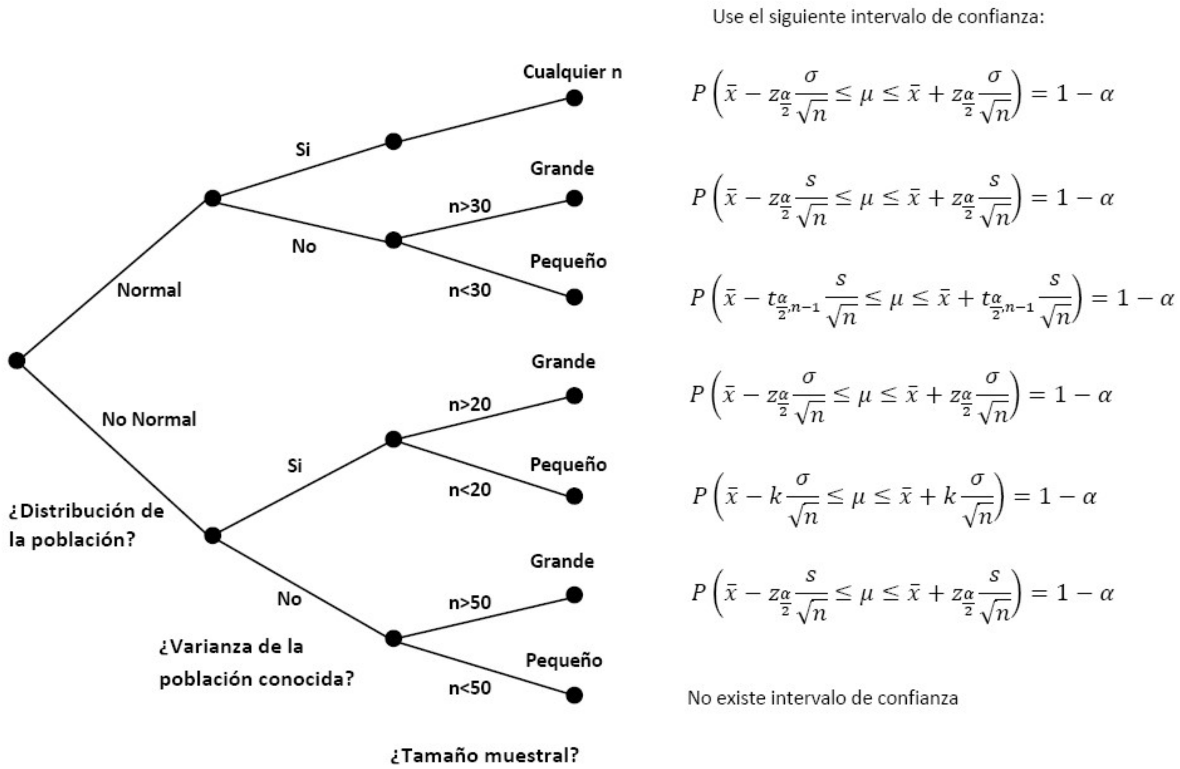
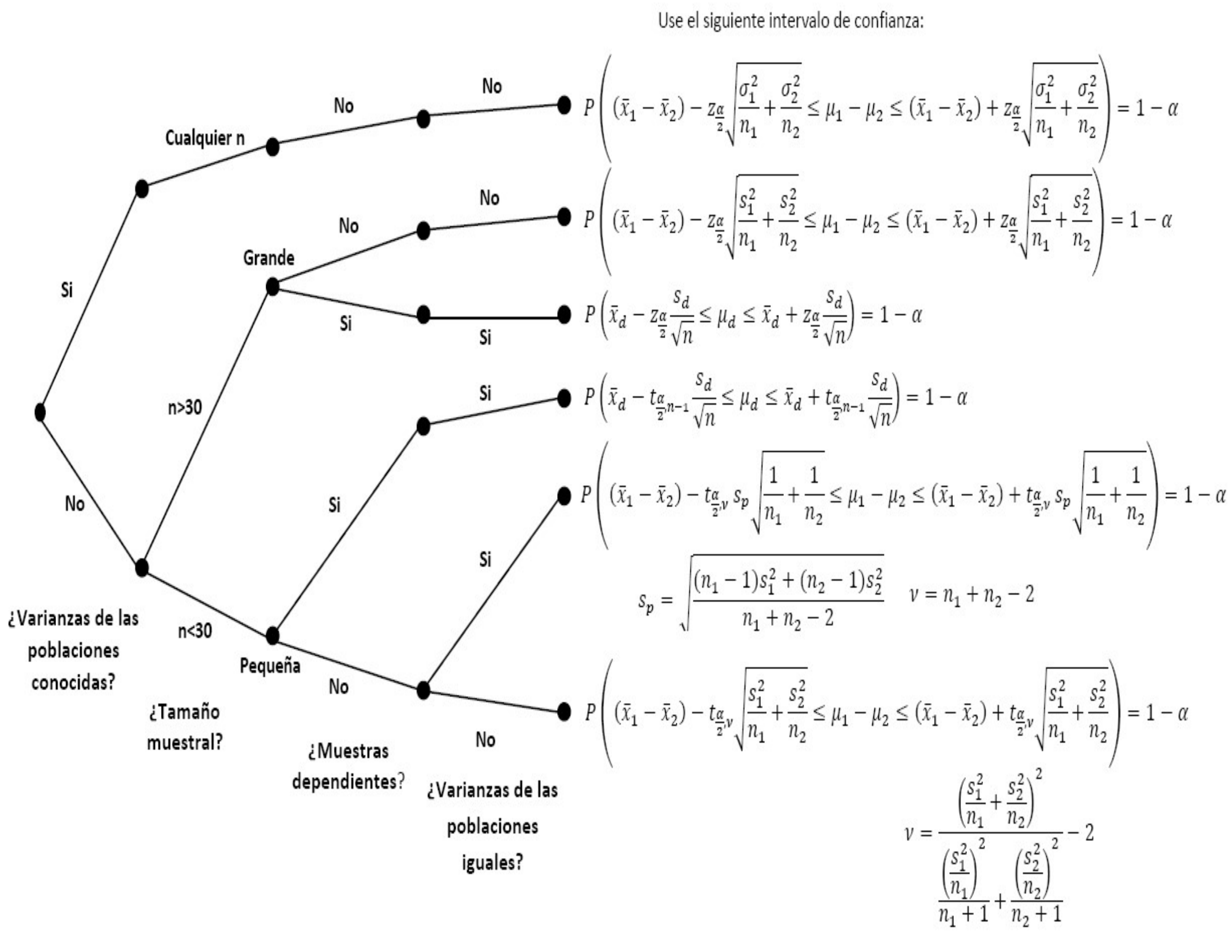


FIGURA 2
Uso de árboles de decisión para la enseñanza de la estadística inferencial
ARBOL DE DECISIÓN PARA ESCOGER EL INTERVALO DE CONFIANZA ADECUADO PARA LA MEDIA DE LA POBLACIÓN



El valor crítico k, es el valor determinado por el teorema de Tchebysheff (Freund y Simon, 1994): "Para cualquier conjunto de datos y cualquier constante k mayor que uno, el porcentaje de los datos que debe caer dentro de k desviaciones estándar de cualquier lado de la media es de por menos $1 - \frac{1}{k^2}$". Por lo tanto, uno puede estar seguro que, por ejemplo, el $1 - \frac{1}{2^2} = 75\%$ de los valores de una distribución deben caer dentro de más o menos dos desviaciones estándares. Por ejemplo, si el investigador quiere un intervalo de confianza del 95%, el valor de k debe ser de $k = \sqrt{\frac{1}{\alpha}} = \sqrt{\frac{1}{0{,}05}} = 4.47$.

FIGURA 3
Uso de árboles de decisión para la enseñanza de la estadística inferencial
Arbol de desición´para escoger el intervalo de confianza adecuado para la diferencia de medias de poblaciones normales



$\bar{x}_d$ y s_d son la media y la desviación estándar muestrales de la diferencia de datos pareados. μ_d es la diferencia de medias poblacionales de datos pareados. s_p es la desviación estándar promedio de las muestras. v representa los grados de libertad de la distribución *t* de Student.

Resolución.

a) De la Figura 1, respondemos:

- ¿Número de poblaciones? Como existe una muestra antes y otra después, éstas se extrajeron de dos poblaciones.
- ¿Palabra clave? Comparación de varianzas.

Por lo tanto usamos el intervalo para el cociente de varianzas:

$$P\left(\frac{s_1^2}{s_2^2}\frac{1}{F_{\frac{\alpha}{2},n_1-1,n_2-1}}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\frac{\alpha}{2},n_1-1,n_2-1}}\right)=1-\alpha$$

Reemplazando la información muestral y puntos críticos de la distribución de Fisher:

$$P\left(\frac{101.17}{94.73}\frac{1}{4.53}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{101.17}{94.73}\frac{1}{0.20}\right)=95\%$$

$$P\left(0{,}24\le\frac{\sigma_1^2}{\sigma_2^2}\le 5.23\right)=95\%$$

La variabilidad del porcentaje de impurezas antes y después de la instalación del nuevo dispositivo de filtrado es la misma, ya que el intervalo contiene el valor de 1.

b) De la Figura 1, respondemos:

- ¿Número de poblaciones? Como existe una muestra antes y otra después, éstas se extrajeron de dos poblaciones.
- ¿Palabra clave? Comparación de medias.
- Nos vamos a la Figura 3:

- ¿Las varianzas de las poblaciones son conocidas? "No".
- ¿Tamaños muestrales? "Menores a 30".
- ¿Muestras dependientes? "No".
- ¿Varianzas poblacionales iguales? "Si" (Resultado del inciso a)).

Por lo tanto se usa el siguiente intervalo para la diferencia de medias:

$$P\left((\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2},v}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1-\mu_2 \le (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2},v}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \right)=1-\alpha$$

$$\text{con } s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$

Reemplazando la información muestral y puntos críticos de la distribución *t*:

$$s_p = \sqrt{\frac{(8-1)101.17+(9-1)94.73}{8+9-2}} = 9.89$$

$$P\left((12.5-10.2)-(2.13)9.89\sqrt{\frac{1}{8}+\frac{1}{9}} \le \mu_1-\mu_2 \le (12.5-10.2)+(2.13)9.89\sqrt{\frac{1}{8}+\frac{1}{9}} \right) = 95\%$$

$$P\left(-7.94 \le \mu_1-\mu_2 \le 12.54\right) = 95\%$$

El dispositivo de filtrado no ha sido efectivo, porque el porcentaje medio de impurezas antes y después es el mismo (el intervalo contiene el valor de 0).

c) De la Figura 1, respondemos:

- ¿Número de poblaciones? Como solo se quiere ver si con el nuevo dispositivo se llegó a la meta: Una.
- ¿Palabra clave? Media.
- Nos vamos a la Figura 2:
  - ¿Distribución de la población? Vamos a suponer que la muestra proviene de una población normal.
  - ¿Varianza de la población conocida? "No".
  - ¿Tamaño muestral? Pequeño.

Por lo tanto el intervalo para la media es:

$$P\left(\bar{x}-t_{\frac{\alpha}{z},n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{x}+t_{\frac{\alpha}{z},n-1}\frac{s}{\sqrt{n}} \right) = 1-\alpha$$

Reemplazando la información muestral y puntos críticos de la distribución t:

$$P\left(10.2-2.31\frac{9.73}{\sqrt{9}} \le \mu \le 10.2+2.31\frac{9.73}{\sqrt{9}} \right) = 95\%$$

$$P\left(2.71 \le \mu \le 17.69\right) = 95\%$$

Se llegó a cumplir la meta del porcentaje de impurezas permitido, ya que el intervalo contiene el valor de 5%.

Aclaración: Los valores críticos F, t o z para los distintos intervalos de confianza con una probabilidad de certeza, se determinan mediante las tablas estadísticas respectivas que se pueden encontrar en cualquier libro de Estadística.

3. CONCLUSIONES E IMPLICACIONES

Se han configurado árboles de decisión para elegir el intervalo de confianza adecuado en la resolución de problemas de estimación estadística, a partir de bases teóricas y empíricas generalmente aceptadas.

Este método se constituye en una gran ayuda didáctica para el estudiante que le permitirá aprender con mayor facilidad, ya que es fácil de comprender y usar, y brinda la solución de elección de forma más rápida y exhaustiva que los métodos de flujogramas o tablas resumen. Por otro lado, se presta para el uso en la informática, por su forma sistematizada de elección, y la configuración de un software estadístico más amigable que los actuales.

Por otro lado, el método de árboles de decisión puede ser usado de manera similar a la de los intervalos de confianza en otros temas de la estadística inferencial, como ser: pruebas de hipótesis, tanto paramétricas como no paramétricas, modelos probabilísticos, modelos de regresión; o en temas de la estadística descriptiva, control estadístico de calidad o el diseño y análisis de experimentos (los autores han desarrollado el método en todas estas áreas).

REFERENCIAS BIBLIOGRÁFICAS Y ELECTRÓNICAS

[1]. **BERENSON, LEVINE Y KREHBIEL,** (2001), *"Estadística para Administración"*, Pearson Educación, Segunda Edición, México.

[2]. **LEVIN Y RUBIN,** (1996), **"**Estadística para *Administradores"*, Prentice Hall Hispanoamericana S.A., Sexta Edición, México.

[3]. **MASON Y LIND,** (1995), *"Estadística para Administración y Economía"*, Alfaomega, Séptima Edición, México.

[4]. **FREUND Y SIMON,** (1994), *"Estadística Elemental"*, Prentice Hall, Octava Edición, México.

[5]. **MILLER, FREUND Y JONSON,** (1992), *"Probabilidad y Estadística para Ingenieros"*, Prentice Hall Hispanoamericana S.A., Cuarta Edición, México.

[6]. **MENDENHALL W.,** (1990), *"Estadística para Administradores"*, Grupo Editorial Iberoamérica, Segunda Edición, México.

[7]. **GARCÍA M.,** (1985), *"Socioestadística"*, Alianza Editorial, Madrid-España.

[8]. **MOOD/GRAYBILL,** (1976), *"Introducción a la Teoría Estadística"*, Editorial Aguilar, Cuarta Edición, Madrid-España.

[9]. **MAISEL L.,** (1973), *"Probabilidad y Estadística"*, Fondo Educativo Interamericano, Colombia.

[10]. **HAYS Y WINKLER,** (1971), *"Statistics: Probability, Inference and Decision"*, Holt, Rinehart and Winston Inc., United States of America.

[11]. **LOBEZ y CASA,** (1967), *"Estadística Intermedia"*, Editorial Vicens-Vives, Primera Edición, España.