**Estrategia didáctica 3.4.4. Pruebas de hipótesis**

En la Estrategia didáctica 3.3.2.1. Distribución Normal se habló de la Regla Empírica. En ella se afirma que si una variable aleatoria es normal entonces ciertos porcentajes de los datos se cubrirán en intervalos que se forman al sumarle una dos y tres desviaciones estándar a la media poblacional. Comoes una variable aleatoria normal siempre que se cumplan las condiciones dadas por el TCL, entonces a ella se le aplicará la regla empírica. Esta regla se puede escribir de manera más concisa usando la notación de probabilidad de la siguiente manera:



En general se puede escribir como

 (1)

Donde el valor de z se elige según como se indicó en el estrategia anterior y el área que se cubrirá es la que está señalada en las tablas, por ejemplo, si se elige z = 1.5, entonces



Lo que significa que el 86.64% de las medias muestrales caerán en el intervalo , con z = 1.5.

El valor 1-α es la probabilidad asociada al intervalo, también llamada nivel de confianza. Esta probabilidad se establece a priori, es decir, confiamos en que nuestra predicción se cumpla un cierto porcentaje de veces que se realice el experimento, según el valor que hayamos elegido previamente para z. Así, por ejemplo, si el nivel de confianza es de 98%, entonces 1-α = 0.98, de donde α = 0.02. Esto significa que el área central bajo la curva normal es de 0.98 y la suma de las áreas de las dos colas es de 0.02. Por simetría, el área en cada cola es de α/2, es decir de 0.01 en cada una de ellas. Esto puede observarse en la siguiente figura. (Para comprender la figura deben hacerse las siguientes equivalencias: Z1 =  ; Z2 =  ; 1- a = 1 – α y a/2 = α/2. Además, el área bajo la curva es de 0.9772 pues Z1 = -2 y Z2 = 2. Por último, el área en cada cola es, por lo tanto, de 0.0114)

 

Asimismo como se establece la regla empírica para  al ser normal, también puede establecerse para Z. Sabemos que ambas variables están relacionadas mediante la ecuación de transformación, por lo tanto, podemos establecer la regla empírica para Z, como se observa en la figura anterior.

Todo esto significa que

 (2)

La fórmula (2) equivale a (1), pues en esta última se tiene que la probabilidad de que Z caiga en el intervalo dado por [-zα/2 , zα/2], es de 1-α. Por ejemplo, si elegimos z como 1.5, entonces 1-α valdrá 86.64%, es decir 1-α = 0.8664, de donde α = 0.1336.

Por el TCL para la media, sabemos que se cumple . Si sustituimos esta ecuación en (2) tendremos que



Multiplicando por σ/, ambas desigualdades y restando además de multiplicar por -1, se tendrá

 (3)

La expresión (3) indica algo realmente interesante. Observa que μ, el parámetro cuyo valor es de especial interés, pues si lo conocemos sabremos el estado que guarda estadísticamente la población, se encuentra en el intervalo dado por . Esto significa que si tomamos una muestra de tamaño n de la población objetivo y calculamos , entonces podemos afirmar, con una probabilidad o nivel de confianza determinado de antemano para calcular , que esta será la frecuencia con que μ quedará atrapada en el intervalo centrado en ,según el valor elegido para 1-α. Es necesario, sin embargo, que conozcamos σ. Por ejemplo, hagamos la predicción acerca de cuál será el intervalo en el que deberá caer la verdadera media poblacional de los salarios de los obreros textiles cuando se toma una muestra de 5 salarios (no debe olvidarse que se debe cumplir el TCL para que estos cálculos puedan realizarse.) En el boletín 3 se tomó una muestra de este tamaño y se obtuvo un valor de =236.11. Si decimos que se desea una probabilidad de que cuando se tome una muestra de este tamaño la probabilidad (o nivel de confianza) sea de 95% de que μ sea atrapada en el intervalo centrado en , entonces debemos elegir 1-α = 0.95 y por tanto =1.96. Como sabemos que σ = 22.3, entonces



Lo que significa que se está afirmando que la probabilidad de que μ caiga en el intervalo centrado en , es de 95%. Si realizamos cálculos



Lo cual también se explica como que (216.56,255.55) es el intervalo de confianza del 95% para μ, lo que significa que si realizamos el experimento de tomar una muestra de 5 salarios de los obreros textiles 100 veces y calculamos , entonces el intervalo que se forme contendrá a la media poblacional μ 95 veces de cada 100. Esta es la interpretación frecuentista de la probabilidad. Los valores extremos del intervalo, que en adelante llamaremos intervalo de confianza para μ, se les llama límites inferior y superior del intervalo.

Además sabemos que  no toma siempre el mismo valor para cada muestra de tamaño 5 (o de cualquier otro tamaño), por ejemplo, si se toma otra muestra de tamaño 5, se obtiene una = 237.72, y si usamos este valor para construir otro intervalo para μ, se tendrá



y al realizar los cálculos



Lo que significa que el intervalo puede cambiar, pero aún así, se sigue garantizando que μ estará contenida en el intervalo según el nivel de confianza elegido de antemano. Los dos intervalos que se calcularon antes, se pueden dibujar para que se comprenda mejor la idea del cálculo de dichos intervalos. En la figura siguiente se observa que la media poblacional está contenida en ambos intervalos que están dibujados con segmentos cuyo punto medio es la media muestral . Los intervalos aparecen desplazados con respecto a la media poblacional, porque en cada muestra de n = 5,  ha cambiado. Sin embargo, de todas formas el parámetro de interés, μ, ha sido atrapado por el intervalo construído con los datos de la muestra.



Este procedimiento puede hacerse de manera continua, es decir, en un proceso industrial, en el que se desea determinar si todos los días está bajo control, se pueden calcular intervalos como los anteriores para la media poblacional y si en cada caso μ cae en el intervalo predeterminado, se puede afirmar que el proceso está bajo control. Por ejemplo, si se desea que el llenado medio de todos los envases de refresco que se producen en un día en una embotelladora no caigan fuera de los límites de un intervalo de confianza, entonces se pueden hacer mediciones continuas del proceso, tomando muestras de tamaño n y construyendo el intervalo de confianza del 95%. Si el intervalo no contiene el valor medio de llenado con el que se espera que se produzcan los envases, entonces existe la posibilidad de que el proceso ya esté fuera de control, ya sea porque los envases están siendo llenados con exceso de líquido, o bien, porque están siendo insuficientemente llenados.

La construcción de intervalos de confianza para predecir con cierta probabilidad en qué región cae el parámetro de interés, en este caso μ, es muy útil para decidir si un fenómeno muestra tendencias fuera de control o si está en ciertos límites tolerantes. Observa que el investigador selecciona el nivel de confianza con que desea hacer las predicciones, (que comúnmente es del 95%), y también puede calcular el tamaño de la muestra que requiere según su presupuesto económico (porque el tamaño de la muestra, en muchos casos, aumenta el costo del muestreo), el error de estimación que asumirá en sus cálculos y con estos datos podrá hallar la media muestral. Sin embargo, hay una dificultad seria en el cálculo del intervalo: en cualquier problema práctico, no es razonable pensar que si se desconoce el valor de μ, se conozca el valor de σ. Es más común que ambos parámetros sean desconocidos. En este caso, ¿cómo calcular tanto el tamaño de muestra como el intervalo de confianza para estimar a μ?

Muchos investigadores afirman que cuando n>30, entonces en la fórmula (3) del intervalo de confianza, s puede sustituir a σ pues a medida que el tamaño de la muestra aumenta, s se aproxima a σ. Por lo tanto, podemos hacer estimaciones de μ usando exclusivamente la información que nos da la muestra, pues tanto , como s, se calcula a partir de los datos que la conforman. El intervalo que se construye bajo esta suposición es el siguiente: .

Una situación similar a la comentada antes surge también para estimar a π. Siguiendo el mismo razonamiento que se hizo para construir un intervalo de confianza para μ, también se puede hallar un intervalo de confianza para π. Se deja al lector que reproduzca el razonamiento hecho para μ, de manera que obtenga el intervalo de confianza para π. El resultado será el siguiente



Puede observarse que es necesario, para construir un intervalo de confianza para π, tomar una muestra de tamaño n, calculada a partir del error de estimación que se propone de antemano, y con la muestra se calcula p, la proporción muestral. El nivel de confianza predeterminado dará el valor de z con que se desea construir dicho intervalo.

El intervalo, en suma, será: .

**EJERCICIOS**

1. Para el problema de los salarios de los obreros, construye con los mismos datos el intervalo de confianza para la verdadera media de los salarios de los obreros, μ, usando los siguientes niveles de confianza: 90%, 95% y 99%. En cada caso interpreta el intervalo de confianza obtenido. ¿Qué observas en la longitud de los intervalos?
2. De la hoja de datos de los salarios de los obreros textiles, toma una muestra de 40 salarios, calcula el intervalo de confianza para μ de dos formas
	* Considera que σ = 22.3
	* Considera que σ es desconocida.
3. Toma una muestra de 50 salarios y repite el problema anterior. ¿Qué observas en el tamaño de los intervalos? ¿Por qué crees que ocurre esto?
4. Repite los ejercicios 1 al 3 pero usando la población de salarios de los obreros petroleros. Revisa cuál es el valor de σ para esta población o calcúlala.
5. Deduce el intervalo de confianza para π. Las consideraciones que deben hacerse para que se pueda construir el intervalo es que se necesita que np y n(1-p) sean ambos mayores a 5. Esto es porque así se “garantiza” que la distribución de p tiende a ser normal y por tanto se puede calcular z a partir de las tablas de la distribución normal. Bajo estas condiciones p es un buen estimador puntual de π.
6. Una máquina produce piezas metálicas de forma cilíndrica. Se toma una muestra de las piezas y los diámetros son 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01 y 1.03 centímetros. Calcula el intervalo de confianza del 99% para el diámetro medio de todas las piezas. La población de los diámetros es normal. Además σ = 0.05.
7. El tiempo medio de estudio de los alumnos de Educación Básica es normal con una desviación estándar de 35 minutos. Se toma una muestra de 90 alumnos y se les pregunta el tiempo de estudio que usan para realizar su tarea, obteniéndose una media de 2 horas y 15 minutos. Calcula el intervalo de confianza de la verdadera media de tiempo de estudio del 95%. Interprétalo.
8. Una muestra de 12 graduadas de cierta escuela secretarial escriben a máquina un promedio de 79.3 palabras por minuto con una desviación estándar muestral de 7.8 palabras por minuto. Calcula el intervalo de confianza de 95% para el número promedio de palabras escritas por todas las graduadas de la escuela.
9. Calcula el error máximo de estimación de la proporción de alumnos que trabajan, cuando se toma una muestra de 200 alumnos de las Escuelas Normales del país, obteniéndose una proporción de 0.22. Hazlo para un nivel de confianza del 98%. Explica el resultado.
10. Calcula el intervalo de confianza del 98% para la verdadera proporción de los alumnos que trabajan en el CCH. Explica el significado del intervalo. Usa los datos del ejercicio anterior.
11. El 76% de los alumnos de las Escuelas Normales estudiarán una maestría en Estadística. Este dato resultó al seleccionar 120 alumnos al azar. Construye el intervalo de confianza del 95% para la verdadera proporción de alumnos de las Escuelas Normales que estudiarán una maestría en estadística.
12. La proporción de hombres obesos en una muestra de cierta ciudad fue de un 35%. La muestra fue de 200 hombres elegidos al azar. Construye el intervalo de confianza del 99% para la verdadera proporción de obesos en la ciudad.
13. De los tabulados del INEGI encontrarás el ingreso de la población en salarios mínimos según la ocupación. Diseña muestreos para construir intervalos de confianza, sin calcularlos, para la diferencia de proporciones entre los salarios mínimos, a partir de 5 o menos salarios mínimos, que reciben hombres y mujeres; diseña otro para la diferencia media entre salarios que perciben hombres y mujeres. Tampoco lo calcules.
14. En una encuesta diseñada por Gallup en E.U. en mayo del 2001, se entrevistaron 600 adultos, 42% los cuales opinaron que el aborto era moralmente aceptable. Calcula un intervalo de confianza de la verdadera proporción de adultos que opinan que el aborto es moralmente aceptable. Usa un nivel del 98%. En la misma encuesta se encontró que el suicidio era moralmente equivocado en un 78%. Calcula un intervalo de confianza del 99% para la verdadera proporción de los adultos que tienen esta opinión. Interpreta Los resultados.
* **Guardar** con el nombre **nombre-apellido.E3.4.4.Pruebas de hipotesis-grupo.doc**

**LECTURA**

En el libro “Bully for brontosaurus. Reflections in natural history.” Stephen Jay Gould, el gran paleontólogo y divulgador científico, escribe: “Tversky y Kahneman realizaron dos brillantes experimentos sicológicos. Estos estudios nos han proporcionado uno de los más detallados atisbos sobre el “razonamiento natural” y su curiosa desviación de la verdad lógica. Construyeron una descripción ficticia de una joven: “Linda tiene 31 años de edad, es soltera, abierta y muy brillante. Se especializó en filosofía. Cuando era estudiante, se preocupó mucho de temas de discriminación y justicia social y participó en manifestaciones antinucleares.” Con esta información, les mostraron una lista de descripciones de Linda a unos sujetos experimentales, las que ellos deberían disponer en orden de presunta verosimilitud. Las descripciones fueron:

* 1. Linda es maestra en un colegio de primaria.
	2. Linda trabaja en una librería y acude a clases de yoga.
	3. Linda es una activa feminista.
	4. Linda es asistente social especializada en siquiatría.
	5. Linda es miembro de la liga de mujeres votantes.
	6. Linda es cajera de un banco.
	7. Linda es vendedora de seguros.
	8. Linda es cajera de un banco y una activa feminista.

 En realidad sólo tres de ellas f, g y h, constituyen el verdadero experimento. (Las demás son arbitrarias porque no se puede determinar qué clase de persona es Linda con los datos enunciados). Es evidente que la afirmación h es la menos probable., puesto que cualquier conjunción ha de ser menos probable que cualquiera de sus partes considerada por separado. Pero todos los sujetos experimentales, tanto como los sujetos que han reflexionado sobre lógica y probabilidad, calificaron h como más probable que f o g.

¿Por qué cometemos continuamente este sencillo error de lógica? Tversky y Kahneman aducen que nuestra mente no está construida (por la razón que sea) para funcionar según las reglas de la probabilidad y la estadística, aunque es evidente que estas reglas gobiernan nuestro universo. Abstraemos lo que consideramos que es la esencia de una entidad y después disponemos nuestros juicios según su grado de similitud a este tipo asumido. Puesto que se nos da un “tipo” para Linda que implica feminismo, pero claramente no un trabajo para el banco, consideramos que cualquier afirmación que se ajuste al tipo es más probable que otra que sólo contiene material contrario al tipo. Esta propensión puede ayudarnos a comprender toda una gama de preferencias humanas, desde la teoría de las formas de Platón a los estereotipos modernos de raza o sexo. También podríamos comprender mejor el mundo, y liberarnos de prejuicios impropios, si comprendiéramos adecuadamente el funcionamiento de la probabilidad y de su alcance inexorable, a través de las leyes de la lógica, sobre buena parte de las pautas de la naturaleza. “Ajustar el tipo” es un error común; no comprender la pauta aleatoria en las buenas y las malas rachas es otra. Si conociéramos mejor a Doña Suerte, Las Vegas seguiría siendo seguramente una parada de carretera en el desierto.”

Es bastante difícil que en este pequeño texto pudiera explicarles la gran cantidad de errores de disponibilidad que cometemos continuamente. La gran mayoría de ellos los hacemos porque ignoramos, o en el peor de los casos no aplicamos, los conceptos estadísticos. Les recomiendo el libro “Irracionalidad: el enemigo interior” escrito por Stuart Sutherland, publicado por Alianza Editorial, #1819, como un compendio de errores que cometemos en todas las actividades humanas. Es recomendable particularmente para ustedes que en la mayoría de los casos deberán tomar varias decisiones en el futuro.

En otro libro titulado “The Panda´s Thumb” también escrito por S. J. Gould, hay unos capítulos acerca del lugar que ocupaban las mujeres en el pensamiento de algunos científicos del siglo XIX. Algunas “disciplinas“ que sirvieron para tal uso eran la craneometría o medición del cráneo y la antropometría o medición del cuerpo humano. En realidad se trataba de probar una idea preconcebida que consistía en probar la inferioridad de las mujeres. Paul Broca, profesor de la Universidad de París, intentó realizar la prueba. Era un medidor cuidadoso y reunió multitud de datos que usó para apoyar su hipótesis. El afirmaba que era el apóstol de la objetividad y que ante los hechos había que dejar de lado las supersticiones y sentimentalismos. Según sus datos, las mujeres tenían cerebros más pequeños que los de los hombres y por lo tanto no podían ser iguales en inteligencia. Un discípulo de Broca, que no concordaba con los resultados de aquel, decía: “ ...los teólogos se han preguntado si las mujeres tenían alma. Varios siglos después, algunos científicos están dispuestos a negarles una inteligencia humana.”

Broca realizó autopsias en hospitales de donde extrajo sus datos. Su argumento se basaba en el mayor tamaño del cerebro de los hombres y en un supuesto incremento de la superioridad del hombre al paso del tiempo. Con 292 cerebros de hombre calculó un peso medio de 1325 gramos; de 140 cerebros de mujer halló que la media era de 1144 gramos, lo que daba 181 gramos de diferencia de pesos medios de los cerebros. Anticipó una objeción a sus resultados: la diferencia de estaturas en el hombre y la mujer. Decía: “Podemos preguntarnos si el pequeño tamaño del cerebro de la mujer depende exclusivamente del tamaño de su cuerpo. Tiedemann ha propuesto esta explicación. Pero no debemos olvidar que las mujeres son, por regla general, un poco menos inteligentes que los hombres. Una diferencia que no debemos exagerar, pero que es real.” Broca también midió la capacidad de cráneos prehistóricos (7 de varón y 6 de hembra) y halló una diferencia de 99.5 centímetros cúbicos entre los de hombre y mujer, mientras que cuando midió cráneos modernos, ellos mostraban una diferencia mayor: de 129 a 220.7 centímetros cúbicos. La explicación de la diferencia de capacidades era debida, según un discípulo de Broca, a que “el hombre que combate por dos o más en la lucha por la supervivencia, que carga con todas las responsabilidades y preocupaciones del día de mañana, que está continuamente en activo, combatiendo contra su medio ambiente y contra sus rivales humanos, necesita más cerebro que la mujer a la que debe proteger y alimentar, la mujer sedentaria, carente de vida interior alguna, cuyo papel es criar hijos, amar y ser pasiva.”

Un científico social, Gustave Le Bon, También en el siglo XIX, escribía ataques tan furibundos contra las mujeres como el siguiente: “En algunas razas superiores, el tamaño del cerebro de las mujeres está más próximo al de los gorilas que de los varones. Todos los sicólogos que han estudiado la inteligencia reconocen que ellas representan las formas más inferiores de la evolución humana y que están más próximas a los niños y a los salvajes que al adulto civilizado. Son insuperables en su veleidad, su inconstancia, en su carencia de ideas y en su incapacidad para razonar.” Le Bon, sin embargo, aceptaba que existían mujeres más inteligentes que muchos hombres, pero, la explicación de ello era que “resultan tan excepcionales como el nacimiento de una monstruosidad, como un gorila de dos cabezas, que podemos olvidarlas por completo.”

S. J. Gould examinó estadísticamente los datos que legó Broca. Además del peso del cerebro, Broca registró la edad y la estatura. Actualmente se sabe que el tamaño del cerebro disminuye con la edad y que el peso del cerebro aumenta con la estatura. Gould descubrió que los cráneos de mujer que usó Broca eran en término medio más viejos que los de los hombres. Usó la regresión múltiple, para evaluar la influencia de la edad y la estatura del individuo sobre el tamaño del cerebro. En un análisis de los datos sobre las mujeres descubrió que, con la estatura y edad media de los varones, el cerebro de una mujer pesaría 1212 gramos, de manera que la corrección de edad y estatura era menor que la obtenida por Broca. Existía, sin embargo, una diferencia de pesos de 113 gramos. Gould escribe: “No sé que conclusiones sacar de esta diferencia, porque existen otros factores que tienen una importante influencia sobre el tamaño del cerebro. Las enfermedades degenerativas a menudo llevan consigo una disminución sustancial del tamaño del cerebro, independientemente del decrecimiento debido a la edad. El tipo de muerte también influye sobre el peso del cerebro, de manera que existe un aumento de unos 60 gramos si hay muerte por accidente. Actualmente, no existe un procedimiento para eliminar la influencia de la estatura o el peso del individuo en el tamaño del cerebro. Otros autores intentaron relacionar la masa muscular con el tamaño del cerebro. Manouvrier realizó esto último y halló que incluso en hombres y mujeres de la misma estatura había marcada diferencia del tamaño del cerebro en los hombres. Pero cuando efectuó algunas correcciones de lo que llamó “masa sexual”, resultó que las mujeres iban por delante en el tamaño cerebral. Así pues, la diferencia entre tamaños del cerebro, al realizar correcciones, está probablemente en cero y puede perfectamente ser favorable a las mujeres. Además, la diferencia de 113 gramos de peso, es la diferencia de tamaño del cerebro entre un hombre de 1.62 y uno de 1.92, según los datos de Broca. En pocas palabras, ¿qué hacer con los datos de Broca? Desde luego, no permiten afirmar confiadamente que los hombres tienen el cerebro más grande que las mujeres.”

Hay que anotar que en el siglo XIX (Y hoy, sin duda, en muchos países), las mujeres, los negros, los indios y los pobres, sufrían de alta discriminación. Especialmente, las mujeres eran denigradas porque se disponían de datos de fácil acceso nunca apoyados por teorías que relacionaran la inteligencia con el tamaño del cerebro. Una explicación alternativa al hecho del dominio de los hombres a lo largo de la Historia, la dio María Montessori quien decía que las mujeres eran intelectualmente superiores, pero los hombres habían prevalecido hasta ese momento en virtud de su fuerza física y, puesto que la tecnología había abolido la fuerza como instrumento de poder, la era de la mujer había llegado. A pesar de esta opinión tan valiosa, Montessori también quería establecer certezas a partir de la medición del tamaño del cráneo, pues había usado una técnica muy parecida a la de Broca para probar una correlación entre tamaño del cerebro y la inteligencia: medía la circunferencia de las cabezas de los niños e infería que los niños prometedores tenían cabeza más grande.

En realidad el concepto de inteligencia es muy escurridizo. Se han inventado varias definiciones que pretenden dar un procedimiento objetivo para medirla, como los tests de inteligencia. Probablemente el entorno social influya en la inteligencia, pero es difícil dar una medida precisa de él. Algunos animales muestran indicios de inteligencia al resolver problemas de alimentación pero no es fácil decidir si han hecho un acto inteligente o reflexivo. Broca fue víctima de los prejuicios de su época e intentó probar una idea que dominaba a la mayoría de sus contemporáneos. Nunca escuchó la voz del filósofo francés Raymond Lecastelier que escribía sobre el escaso entendimiento que había entre el hombre y la mujer: “Un error inacabable, entre los infinitos errores de los hombres, ha sido el de mantener a la mujer enclaustrada en un estado de insensata domesticidad, donde aparece como un ser de ralas e insípidas ideas, y en el de no consentir en su integración plena a la vida y a las actividades humanas. El hombre, al minimizar la capacidad intelectual de la mujer, se desligó de ella y la volvió una criatura lejana e incomprensible. La idealización nos ha llevado a negar de una manera absoluta a la mujer, a considerarla una criatura de naturaleza inferior con la que el hombre sólo puede compartir los hechos mínimos e intrascendentes, es decir, los meramente instintivos y exteriores, es decir todos aquellos en los que no intervienen o apenas si intervienen, las funciones de la inteligencia y del espíritu.”

Lecastelier murió en París en 1899. Fue contemporáneo de Broca (1824-1880), pero sus libros fueron escasamente conocidos por los estudiosos franceses, a diferencia de las opiniones de Broca. Gould escribió una dura crítica a Broca en el libro “La falsa medida del hombre”, donde usa una gran variedad de conceptos estadísticos y evalúa la facilidad con que estos pueden manipularse para obtener conclusiones falsas o prejuiciadas, como las que Broca adujo.