Universidad Nacional Autónoma de México

Facultad de Ciencias

# El tamaño del efecto

TESIS

que para obtener el título de

Actuaria

presenta

**Zuleima Coronel Nuñez**

Directora de tesis:
Dra. Silvia Ruiz Velasco Acosta

Ciudad Universitaria, CDMX, 2019

*Dedicado a:*
*Sofía Nicté.*

# Agradecimientos

En primer lugar quiero expresar mi gratitud a la Universidad Nacional Autónoma de México, mi alma mater, por haberme brindado la oportunidad de formar parte de ella y a quien debo mi formación.

A mi asesora de tesis, la Dra. Silvia Ruiz Velasco Acosta, por si tiempo y el apoyo brindado en la realización de esta tesis.

A mis padres por haberme brindado la oportunidad de cumplir mis metas, por su apoyo y gran esfuerzo.

A las personas que me han apoyado y acompañado en todo momento. Mi madre quien me ha dado su apoyo incondicional y de manera muy especial a Ulises por estar a mi lado en cada paso que he dado, con amor y paciencia.

A todos aquellos que de alguna forma me ayudaron a concluir este ciclo, en particular a Janet por haberme ayudado a dar el primer paso hacia la universidad.

# Introducción

En la presente tesis se estudia un método alternativo a las pruebas de hipótesis con el fin de generar conclusiones más útiles sobre nuestros datos, ya que a través de una prueba de hipótesis podría decirse que se obtienen conclusiones un tanto limitadas, las cuales impiden un mejor análisis de los datos, por ejemplo en una diferencia de medias al concluir simplemente si las medias son o no distintas y no mostrar nada sobre qué tan grande o pequeña resulta tal diferencia. Dicho método es llamado *"effect size"* o tamaño del efecto el cual tiene una serie de facetas, características y propiedades presentadas en este trabajo.

El tamaño del efecto busca saber qué tan grande es un fenómeno de estudio, o dicho de otra forma, conocer el grado en el que un fenómeno se encuentra presente en la población de estudio (Cohen, 1988), y en caso de que no exista experiencia suficiente para generar conclusiones sobre qué tan grande o significativa resulta la presencia de determinado fenómeno en la población, tomar como referencia valores establecidos para definir qué tan grande o pequeño es el efecto, en caso de que se encuentre presente en la población.

Este método puede ser representado mediante diversos índices, entre los más utilizados se encuentran el índice $d$ para diferencia de medias, el índice $f$ para análisis de varianza y el índice $f^2$ para análisis de regresión. Tales índices fueron detallados y aplicados al estudio del logro académico en el área de matemáticas al término de la educación básica en México, con el fin de buscar una mejor interpretación y análisis más completo sobre los resultados.

# Índice general

# Capítulo 1

# Inferencia estadística

La inferencia estadística es una rama de las matemáticas que a través de una serie de métodos se encarga de obtener conclusiones sobre el comportamiento de la población a través de una muestra representativa de la misma.

Los métodos paramétricos de la estadística se dividen fundamentalmente en dos: Estimación y Pruebas de Hipótesis, éstos son los puntos centrales que se tratan en este capítulo.

## 1.1. Estimación paramétrica

En la estimación paramétrica de la estadística inferencial, se supone que el comportamiento de la población sigue una distribución de probabilidad conocida, en donde la única incógnita es el valor de los parámetros. En la estadística inferencial existen dos formas de estimar parámetros: la estimación puntual y la estimación por intervalos. En la primera se busca estimar un valor único para un parámetro $\theta$, donde $\theta$ puede ser univariado o multivariado, mientras que en la segunda se busca un intervalo dentro del cual se encuentre el parámetro, con una probabilidad determinada.

Antes de mostrar cada uno de los métodos es importante señalar la diferencia entre un estimador y una estimación. El estimador $\hat{\theta}$ de un parámetro $\theta$, es una función de la muestra, mientras que una estimación es el valor que toma $\hat{\theta}$ sobre la muestra.

### 1.1.1. Estimación puntual

Existen varios métodos de estimación puntual entre los más conocidos se encuentran el método de momentos y el método de máxima verosimilitud, los cuales se presentan en la siguiente sección.

#### 1.1.1.1. Método de momentos

Sea $X_1, \ldots, X_n$ una muestra aleatoria de la población la cual tiene la misma función de probabilidad que la distribución de la población cuya función de probabilidad de la muestra es $f(X_1, \ldots, X_k; \theta) = f(X_1, \theta) \cdot f(X_2, \theta) \cdot \ldots \cdot f(X_n, \theta)$, entonces $X_1, \ldots, X_n$ forman un conjunto de variables aleatorias independientes e idénticamente distribuidas que constituyen una muestra aleatoria de la población.

El estimador por el método de momentos se obtiene resolviendo el sistema de ecuaciones que resulta de igualar los momentos muestrales, usualmente los primeros $m_1, \ldots, m_k$, con los correspondientes $k$ momentos poblacionales, $\mu'_1, \ldots, \mu'_k$. Lo cual es equivalente a

$$m_r = \mu'_r(\theta_1, \ldots, \theta_k), \quad r = 1, \ldots, k$$

con

$$m_r = \frac{\sum_{i=1}^n X_i^r}{n}, \quad \mu'_r = \mathbb{E}[X^r]$$

Donde $\mu'_r = \mu'_r(\theta_1, \ldots, \theta_k)$, para mostrar la dependencia del momento $\mu'_r$ respecto del conjunto de parámetros, obteniendo como soluciones

$$\hat{\theta}_1(X_1, \ldots, X_n), \ldots, \hat{\theta}_k(X_1, \ldots, X_n)$$

EJEMPLO **1.1** *Sea $X_1, \ldots, X_n$ una muestra aleatoria de la distribución $N(\mu, \sigma^2)$, con ambos parámetros desconocidos. Entonces k=2 y $\theta_1 = \mu, \theta_2 = \sigma^2, m_1 = \bar{X} = \sum_{i=1}^n \frac{X_i}{n}$, $m_2 = \frac{\sum_{i=1}^n X_i^2}{n}$, $\mu'_1 = \mathbb{E}[X] = \mu$, $\mu'_2 = \mathbb{E}[X^2]$ y $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ entonces $\mu'_2 = \sigma^2 + \mu^2$. Por lo que el sistema a resolver es:*

$$\begin{aligned} \mu &= \bar{X} \\ \sigma^2 + \mu^2 &= \frac{\sum_{i=1}^n X_i^2}{n} \end{aligned}$$

*Entonces,*

$$\begin{aligned} \hat{\sigma^2} &= \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 \\ &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \\ \hat{\mu} &= \bar{X} \end{aligned}$$

*La última igualdad para $\hat{\sigma^2}$ se obtiene de*

$$\begin{aligned} \sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n (X_i^2 - 2\bar{X}X_i - \bar{X}^2) \\ &= \sum_{i=1}^n X_i^2 - 2n\bar{X}\frac{\sum_{i=1}^n X_i}{n} + n\bar{X}^2 \\ &= \sum_{i=1}^n X_i^2 - 2n\bar{X}^2 + \bar{X}^2 \\ &= \sum_{i=1}^n X_i^2 - n\bar{X}^2 \end{aligned} \tag{1.1}$$

*Por lo tanto,*

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2$$

#### 1.1.1.2. Método de máxima verosimilitud

El estimador por el método de máxima verosimilitud es el que se obtiene al maximizar la función de verosimilitud, la cual se define a continuación.

**Definición 1.1** *Se define la función de verosimiltud, $L(\theta|\boldsymbol{X})$, de una muestra aleatoria, $\boldsymbol{X}$, como la función de distribución conjunta, i.e.*

$$L(\theta|\boldsymbol{X}) = L(\theta_1, \ldots, \theta_k|X_1, \ldots, X_n) = \prod_{i=1}^{n} f(x_i|\theta) \tag{1.2}$$

*Es importante notar que la función de verosimilitud es una función del parámetro $\theta$ dada una muestra fija.*

**Definición 1.2** *Se define al estimador de máxima verosimilitud, $\hat{\theta}(\boldsymbol{X})_{mv}$, como el valor que maximiza la función de verosimilitud, ie.*

$$\hat{\theta}(\boldsymbol{X})_{mv} = \max_{\theta\in\Theta} L(\theta|\boldsymbol{X}) \tag{1.3}$$

Generalmente, para el cálculo del estimador máximo verosímil se utiliza la función de *log verosimilitud*, definida como

$$l(\theta) = \ln L(\theta|\mathbf{X}) \tag{1.4}$$

la cual permite facilitar los cálculos, además de que maximizar la función de verosimilitud es equivalente a maximizar su logaritmo.

Ejemplo **1.2** *Sea $X_1, \ldots, X_n$ una muestra aleatoria con distribución $N(\theta, \sigma^2)$ con $\sigma^2$ y $\theta$ desconocidas. La función de verosimilitud está dada por*

$$\begin{aligned} L(\theta|\boldsymbol{X}) &= \prod_{i=1}^{n} \frac{\exp\left(-\frac{(X_i-\theta)^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}} \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\sum_{i=1}^{n} \frac{(X_i-\theta)^2}{2\sigma^2}\right) \end{aligned}$$

*Por lo que la función de log verosimilitud es igual a*

$$l(\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{i=1}^{n} \frac{(X_i-\theta)^2}{2\sigma^2}$$

*Maximizando la función de log verosimilitud para encontrar los estimadores máximo verosímiles, se tiene*

$$\begin{aligned} \frac{\partial l(\theta)}{\partial\theta} &= \sum_{i=1}^{n} \frac{(X_i-\theta)}{\sigma^2} \\ &= \frac{\sum_{i=1}^{n} X_i - n\theta}{\sigma^2} \\ &= 0 \end{aligned}$$

$$\begin{aligned}\frac{\partial l(\theta)}{\partial \sigma^2} &= -\frac{n}{2}\frac{2\pi}{2\pi\sigma^2} - \left(-\frac{\sum_{i=1}^{n}(X_i-\theta)^2}{(2\sigma^2)^2}\right)\\ &= \frac{\sum_{i=1}^{n}(X_i-\theta)^2}{n}\\ &= 0\end{aligned}$$

*Entonces, los estimadores están dados por*

$$\hat{\theta}(\boldsymbol{X}) = \sum_{i=1}^{n}\frac{X_i}{n} = \bar{X}$$

$$\hat{\sigma^2}(\boldsymbol{X}) = \frac{\sum_{i=1}^{n}(X_i-\hat{\theta})^2}{n}$$

*Ahora verifiquemos que* $\frac{\partial^2 l(\theta)}{\partial\theta^2} < 0$

$$\frac{\partial^2 l(\theta)}{\partial\theta^2} = -\frac{n}{\sigma^2} < 0$$

*Por lo que* $\hat{\theta}$ *es un máximo global.*

*Para comprobar que* $\sigma^2$ *también es un máximo basta calcular los eigenvalores de la matriz hessiana y verificar que la ecuación (1.5) y (1.6) se cumplen*

$$\frac{\partial^2}{\partial\theta_i^2}\boldsymbol{H}(\theta_1,\theta_2)|_{\theta_1=\hat{\theta_1},\theta_2=\hat{\theta_1}} < 0 \quad \textit{para i=\{1,2\}} \tag{1.5}$$

$$\left.\begin{vmatrix} \frac{\partial^2}{\partial\theta_1^2}\boldsymbol{H}(\theta_1,\theta_2) & \frac{\partial^2}{\partial\theta_1\partial\theta_2}\boldsymbol{H}(\theta_1,\theta_2) \\ \frac{\partial^2}{\partial\theta_1\partial\theta_2}\boldsymbol{H}(\theta_1,\theta_2) & \frac{\partial^2}{\partial\theta_2^2}\boldsymbol{H}(\theta_1,\theta_2) \end{vmatrix}\right|_{\theta_1=\hat{\theta_1},\theta_2=\hat{\theta_1}} =$$

$$\frac{\partial^2}{\partial\theta_1^2}\boldsymbol{H}(\theta_1,\theta_2)\frac{\partial^2}{\partial\theta_2^2}\boldsymbol{H}(\theta_1,\theta_2) - \left(\frac{\partial^2}{\partial\theta_1\partial\theta_2}\boldsymbol{H}(\theta_1,\theta_2)\right)^2\Big|_{\theta_1=\hat{\theta_1},\theta_2=\hat{\theta_1}} > 0 \tag{1.6}$$

*Entonces*

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\log(L(\theta,\sigma^2|\boldsymbol{x})) &= \frac{-n}{\sigma^2} < 0\\ \frac{\partial^2}{\partial(\sigma^2)^2}\log(L(\theta,\sigma^2|\boldsymbol{x})) &= \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}(x_i-\theta)^2 = \frac{n}{2\sigma^4} - \frac{n}{\sigma^4} = \\ n\left(\frac{1}{2\sigma^4}-\frac{1}{\sigma^4}\right) &= -n/2\sigma^4 < 0\\ \frac{\partial^2}{\partial\theta\partial\sigma^2}\log(L(\theta,\sigma^2|\boldsymbol{x})) &= -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i-\theta) < 0\end{aligned} \tag{1.7}$$

$$\tag{1.8}$$

$$\left| \begin{matrix} \frac{-n}{\sigma^2} & -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i-\theta) \\ -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i-\theta) & \frac{n}{2\sigma^4}-\frac{1}{\sigma^6}\sum_{i=1}^{n}(x_i-\theta)^2 \end{matrix} \right|_{\theta=\bar{x},\sigma^2=\hat{\sigma}^2} \tag{1.9}$$

$$= \frac{1}{\sigma^6}\left[\frac{-n^2}{2}+\frac{n}{\sigma^2}n\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{n}-\frac{1}{\sigma^2}\left(\sum_{i=1}^{n}(x_i-\theta)\right)^2\right]_{\theta=\bar{x},\sigma^2=\hat{\sigma}^2}$$

$$= \frac{1}{\hat{\sigma^6}}\left[\frac{-n^2}{2}+\frac{n^2}{\hat{\sigma}^2}\hat{\sigma}^2-\frac{1}{\hat{\sigma^2}}\left(\sum_{i=1}^{n}(x_i-\bar{x})\right)^2\right]$$

$$= \frac{1}{\hat{\sigma^6}}\left[\frac{-n^2}{2}+\frac{n^2}{\hat{\sigma}^2}\hat{\sigma}^2-\frac{1}{\hat{\sigma^2}}\left(\sum_{i=1}^{n}x_i-n\sum_{i=1}^{n}x_i/n\right)^2\right]$$

$$= \frac{1}{\hat{\sigma}^6}\frac{n^2}{2}>0 \tag{1.10}$$

### 1.1.2. Estimación por intervalos

**Definición 1.3** *Un intervalo de confianza para un parámetro desconocido, $\theta$, es un intervalo de la forma $(\mathbb{L}(\boldsymbol{X}),\mathbb{U}(\boldsymbol{X}))$, el cual se obtiene al estimar $\mathbb{L}$ y $\mathbb{U}$, donde las variables aleatorias $\mathbb{L}$ y $\mathbb{U}$ son funciones de la muestra tales que $\boldsymbol{P}(\mathbb{L}(\boldsymbol{X})\leq\theta\leq\mathbb{U}(\boldsymbol{X}))=1-\alpha$, en donde ésta se lee como la probabilidad de que el intervalo contenga el valor del parámetro y es igual $1-\alpha$. Donde $1-\alpha$ es el grado o coeficiente de confianza, que nos dirá que el $100*(1-\alpha)\,\%$ de los intervalos de confianza construidos a partir de la muestra contendrán al verdadero valor del parámetro y el $100*\alpha\,\%$ de los intervalos no lo contendrá.*

A continuación se presentan algunos métodos para la construcción de intervalos de confianza para un parámetro desconocido $\theta$.

#### 1.1.2.1. Método pivotal

**Definición 1.4** *Una variable aleatoria $Q(\boldsymbol{X},\theta)$ es una cantidad pivotal si su distribución no depende de $\theta$.*

La idea de este método está en encontrar una función de la muestra aleatoria y del parámetro desconocido, llamada cantidad pivotal. Cuya distribución es conocida y a partir de la cual pueda obtenerse un intervalo de confianza para $\theta$.

Ejemplo **1.3** *Sean $X_1,\ldots,X_n$ una muestra aleatoria de la distribución $Normal(\mu,\sigma^2)$ con ambos parámetros desconocidos y se busca un intervalo de confianza para $\mu$.*
*La estadística $T=\frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t_{n-1}$, donde $s^2$ es la varianza muestral. Entonces $T$ es un pivote pues su distribución no depende de $\mu$.*

*Entonces*

$$\begin{aligned}
\boldsymbol{P}\left(-t_{\frac{\alpha}{2},n-1} \leq \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \leq t_{\frac{\alpha}{2},n-1}\right) &= \boldsymbol{P}\left(-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1} \leq \bar{X}-\mu \leq \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1}\right)\\
&= \boldsymbol{P}\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1} \leq \mu \leq \bar{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1}\right)\\
&= 1-\alpha
\end{aligned}$$

*Donde* $t_{\frac{\alpha}{2},n-1}$ *es el cuantil asociado a una distribución t que acumula el intervalo* $(\bar{X} - \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1}, \bar{X} + \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2},n-1})$*, este último es un intervalo de confianza para* $\mu$ *cuando* $\sigma^2$ *es desconocida.*

#### 1.1.2.2. Pivoteo de la función de distribución

En este método se mostrará otro pivote para la construcción del intervalo de confianza de un parámetro $\theta$, en éste se hace uso de una estadística $T$ con distribución $F_T(t|\theta)$, usualmente se toma a $T$ una estadística suficiente para $\theta$, sin embargo esto no es estrictamente necesario. $T$ puede ser continua o discreta, a continuación se presenta el caso en el que $T$ es continua.

**Definición 1.5** *Un estadístico* $T(\boldsymbol{X})$ *es* ***suficiente*** *para el parámetro* $\theta$ *si la función de probabilidad de la muestra* $\boldsymbol{X}$ *condicionada al estadístico* $T(\boldsymbol{X})$, $P((X_1 \ldots X_n)|T(X_1 \ldots X_n))$, *no depende del parámetro* $\theta$.

**Teorema 1.1** ***Pivoteo de la función de distribución, caso continuo****.*
*Sea* $T$ *una estadística con distribución* $F_T(t|\theta)$ *continua. Para cada* $t \in T$ *y* $\alpha_1 + \alpha_2 = \alpha$ *con* $0 < \alpha < 1$*,* $lasfunciones$ $\theta_L(t)$ *y* $\theta_U(t)$ *se pueden definir de la siguiente manera.*

- *Si* $F_T(t|\theta)$ *es una función decreciente de* $\theta$ *para cada t, entonces*

$$F_T(t|\theta_U(t)) = \alpha_1 \quad y \quad F_T(t|\theta_L(t)) = 1-\alpha_2 \tag{1.11}$$

- *Si* $F_T(t|\theta)$ *es una función creciente de* $\theta$ *para cada t, entonces*

$$F_T(t|\theta_U(t)) = 1-\alpha_2 \quad y \quad F_T(t|\theta_L(t)) = \alpha_1 \tag{1.12}$$

*Obtenendo el intervalo* $[\theta_L(T), \theta_U(T)]$.

*Sea la región de aceptación* $(\alpha_1 \leq F_T(t|\theta_0) \leq 1-\alpha_2)$*, con* $1-\alpha_2 \geq \alpha_1$*,* $\theta_U(t) \geq \theta_L(t)$ *siendo los valores de* $\theta_U$ *y* $\theta_L$ *únicos.*

*Si* $F_T(t|\theta)$ *es una función decreciente, entonces se cumple además, qué*

$$F_T(t|\theta) < \alpha_1 \Leftrightarrow \theta > \theta_U(t) \quad y \quad F_T(t|\theta) > 1-\alpha_2 \Leftrightarrow \theta < \theta_L(t)$$

*Las ecuaciones* (1.11) *pueden ser escritas en terminos de la función de densidad del estadístico* $T$*, de modo que:*

$$\int_{-\infty}^{t} f_T(u|\theta_U(t))du = \alpha_1 \quad y \quad \int_{t}^{\infty} f_T(u|\theta_L(t))du = \alpha_2$$

*Si $F_T(t|\theta)$ es una función creciente, entonces se cumple, qué*

$$F_T(t|\theta) < \alpha_1 \Leftrightarrow \theta < \theta_L(t) \quad y \quad F_T(t|\theta) > 1 - \alpha_2 \Leftrightarrow \theta > \theta_U(t)$$

*Las ecuaciones (1.12) pueden ser escritas en terminos de la función de densidad del estadístico $T$, de modo que:*

$$\int_{-\infty}^{t} f_T(u|\theta_L(t))du = \alpha_1 \quad y \quad \int_{t}^{\infty} f_T(u|\theta_U(t))du = \alpha_2$$

EJEMPLO **1.4** *Sea $X_1, \ldots, X_n$ una muestra aleatoria con función de densidad $f(x|\mu) = e^{-(x-\mu)}I_{[\mu,\infty)}(x)$,entonces $T$=$min(X_1, \ldots, X_n)$conocido como el primer estadístico de orden, es una estadística suficiente para $\mu$.*

*La función de densidad del primer estadístico de orden se define como $f_x(Y) = n[1 - F(y)]^{n-1}f(y)$ por lo que la función de densidad de $T$ es*

$$f_T(t|\mu) = n[1 - (1 - e^{-(t-\mu)})]^{n-1}e^{-(t-\mu)} = ne^{-n(t-\mu)}I_{[\mu,\infty)}(t) \tag{1.13}$$

*Usualmente se usa $\alpha_1 = \alpha_2 = \alpha/2$, entonces fijando $\alpha$ y definiendo a $\mu_U(t)$ y $\mu_L(t)$ para satisfacer*

$$\int_{\mu_U(t)}^{t} ne^{-n(u-\mu_U(t))}\,du = \alpha/2 \qquad (1)$$

$$\int_{t}^{\infty} ne^{-n(u-\mu_L(t))}\,du = \alpha/2 \qquad (2)$$

*De (1) aplicando integración por cambio de variable, se obtiene*

$$1 - e^{n(\mu_U(t)-t)} = \alpha/2 \tag{1.14}$$

*Se despeja a $\mu_U(t)$ de la equación (1.4), entonces*

$$1 - \alpha/2 = e^{n(\mu_U(t)-t)}$$

*Aplicando log obse obtiene*

$$log(1 - \alpha/2) = n(\mu_U(t) - t)$$

*Por lo tanto*

$$\mu_U(t) = \frac{log(1 - \alpha/2)}{n} + t$$

*De (2) aplicando integración por cambio de variable, se obtiene*

$$e^{n(\mu_L(t)-t)} = \alpha/2 \tag{1.15}$$

*Despejamos a $\mu_L(t)$ de la equación (1.5), entonces*

$$n(\mu_L(t) - t) = log(\alpha/2)$$

*Por lo tanto*

$$\mu_L(t) = \frac{log(\alpha/2)}{n} + t$$

*Por lo que el intervalo con una confianza de 1-$\alpha$ para $\mu$ es*

$$\left[\frac{log(\alpha/2)}{n} + T, \frac{log(1-\alpha/2)}{n} + T\right]$$

## 1.2. Pruebas de hipótesis

Otro método para conocer el valor del parámetro $\theta$ es mediante pruebas de hipótesis en el que dada una muestra, se busca afirmar si el valor de $\theta$ pertenece al conjunto paramétrico $\Theta_0$ o al conjunto $\Theta_0^c$, mediante una hipótesis nula, $H_0 : \theta \in \Theta_0$ y una hipótesis alternativa, $H_1 : \theta \in \Theta_0^c$.

Al realizar una prueba de hipótesis se obtiene una de las siguientes conclusiones:

- Rechazar $H_0$ a favor de $H_1$, debido a evidencia suficiente en los datos.
- No rechazar $H_0$, al no existir evidencia suficiente en los datos.

**Definición 1.6** *En cualquier partición del espacio muestral en dos conjuntos. El subconjunto del espacio muestral para el cual $H_0$ es rechazado es llamado región crítica o de rechazo y el subconjunto para el cual $H_0$ no es rechazado es llamado región de aceptación.*

**Definición 1.7** *Al realizar una prueba de hipótesis se pueden cometer dos tipos de errores:*

- ***Error tipo I***
  *Rechazar $H_0$ cuando $\theta \in \Theta_0$, es decir rechazar $H_0$ cuando ésta es verdadera.*
- ***Error tipo II***
  *No rechazar $H_0$ cuando $\theta \in \Theta_0^c$, es decir no rechazar $H_0$ cuando ésta es falsa.*

**Definición 1.8** *La función potencia de una prueba de hipótesis con región crítica RC para probar $H_0 : \theta \in \Theta_0$ frente a $H_1 : \theta \in \Theta_0^c$ es la función de $\theta$ definida por*

$$\beta(\theta) = \boldsymbol{P}_\theta(X \in RC)$$

Por lo general se pide que la función potencia no supere cierto valor, valores próximos a 0, cuando $\theta \in \Theta_0$, lo cual indica pequeña probabilidad de error tipo I, y que su valor sea lo mayor posible, valores próximos a 1, cuando $H_1 : \theta \in \Theta_0^c$, lo cual indica pequeña probabilidad de error tipo II.

**Definición 1.9** *El nivel de significancia de una prueba de hipótesis con región crítica $RC$ para probar $H_0 : \theta \in \Theta_0$ frente a $H_1 : \theta \in \Theta_0^c$, es cualquier número $\alpha \in (0,1)$ tal que $\boldsymbol{P}(RC|\theta) \leq \alpha$.*

El nivel de significancia $\alpha$ es la cota de los errores tipo I, por lo que se busca emplear valores pequeños para $\alpha$.

A continuación se mostrará el método más utilizado para la construcción de pruebas de hipótesis.

### 1.2.1. Razón de verosimilitudes

Este método está ligado con el método de la máxima verosimilitud que se ha desarrollado para la construcción de estimadores, visto en la sección 1.1.1.2, en donde la función de verosimilitud se define en la ecuación (1.1).

**Definición 1.10** *La prueba de razón de verosimilitudes para contrastar $H_0$ :$\theta \in \Theta_0$ frente a $H_1 : \theta \in \Theta_0^c$, al nivel de significancia $\alpha$ tiene como región crítica*

$$RC = \{(X_1 \dots X_n) : \lambda(\boldsymbol{X}) \leq c\}$$

*Con $c$ una constante que satisface, $0 \leq c \leq 1$ y $\lambda(\boldsymbol{X})$ es la estadística para la prueba de razón de verosimilitudes la cual es igual a*

$$\lambda(\boldsymbol{X}) = \frac{\sup_{\theta \in \Theta_0} L(\theta|\boldsymbol{X})}{\sup_{\theta \in \Theta} L(\theta|\boldsymbol{X})} = \frac{L(\hat{\theta}_0|\boldsymbol{X})}{L(\hat{\theta}|\boldsymbol{X})} \tag{1.16}$$

*donde $c$ es elegido tal que $\sup_{\theta \in \Theta_0} \boldsymbol{P}_\theta(\lambda(\boldsymbol{X}) \leq c) = \alpha$.*

$\hat{\theta}_0$ y $\hat{\theta}$ son los estimadores de máxima verosimilitud. El primero sobre el espacio definido por la hipótesis nula y el segundo sobre todo el espacio parametral. Si este cociente es pequeño (cercano a cero), entonces bajo la muestra obtenida, no existe evidencia suficiente para decir que el valor de $\theta$ se encuentre en $\Theta_0$. Por lo que $H_0$ debe ser rechazada.

Ejemplo **1.5** *Sea $X_1, \dots, X_n$ una muestra aleatoria de una población con distribución $N(\theta, \sigma^2)$, con $\sigma^2$ conocida. Se quiere realizar una prueba de hipótesis sobre el posible valor de la media poblacional, $\theta$. Realizando la prueba de razón de verosimilitudes, definimos*

$$H_0 : \theta = \theta_0 \quad versus \quad H_1 : \theta \neq \theta_0$$

*Ahora realizamos el cociente entre las dos verosimilitudes; con $\theta \in \Theta_0$, en el numerador; y $\theta$ en todo el espacio paramétrico $\Theta$ en el denominador.*

$$L(\hat{\theta_0}|\boldsymbol{X}) = (2\pi)^{-n/2}\sigma^{-n} \exp\left(-\frac{\sum_{i=1}^n (X_i - \theta_0)^2}{2\sigma^2}\right)$$

*El estimador máximo verosímil para $\theta$ encontrado en el ejemplo, 1.1.1.2, es $\hat{\theta} = \bar{X}$, así*

$$\begin{aligned} L(\hat{\theta}|\boldsymbol{X}) &= (2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\theta)^2}{2\sigma^2}\right) \\ &= (2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{2\sigma^2}\right) \end{aligned}$$

*Entonces*

$$\begin{aligned} \lambda(\boldsymbol{X}) &= \frac{(2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\theta_0)^2}{2\sigma^2}\right)}{(2\pi)^{-n/2}\sigma^{-n}\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{2\sigma^2}\right)} = \frac{\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\theta_0)^2}{2\sigma^2}\right)}{\exp\left(-\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{2\sigma^2}\right)} \\ &= \exp\left(-\frac{\sum_{i=1}^{n}(X_i-\theta_0)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{2\sigma^2}\right) \qquad (1.17) \end{aligned}$$

*obsérvese que*

$$\begin{aligned} \sum_{i}^{n}(X_i-\bar{X})^2 + n(\bar{X}-\theta_0)^2 &= \sum_{i=1}^{n}X_i^2 - 2\bar{X}\sum_{i=1}^{n}X_i + 2n\bar{X}^2 - 2n\bar{X}\theta_0 + n\theta_0^2 \\ &= \sum_{i=1}^{n}X_i^2 - 2n\bar{X}\theta_0 + n\theta_0^2 \cancelto{0}{- 2\bar{X}n\frac{\sum_{i=1}^{n}X_i}{n} + 2n\bar{X}^2} \\ &= \sum_{i=1}^{n}X_i^2 - 2n\frac{\sum_{i=1}^{n}X_i}{n}\theta_0 + n\theta_0^2 \\ &= \sum_{i}^{n}(X_i-\theta_0)^2 \end{aligned}$$

*Por lo que la ecuación (1.17) se reduce a*

$$\lambda(\boldsymbol{X}) = \exp\left(-\frac{n(\bar{X}-\theta_0)^2}{2\sigma^2}\right) \qquad (1.18)$$

*De la ecuación anterior se obtiene que la región crítica es*

$$\begin{aligned} RC &= \left\{(X_1,\ldots,X_n) : \exp\left(-\frac{n(\bar{X}-\theta_0)^2}{2\sigma^2}\right) \le c\right\} \\ &= \left\{(X_1,\ldots,X_n) : \frac{n(\bar{X}-\theta_0)^2}{n(\frac{\sigma}{\sqrt{n}})^2} \ge c_1\right\} \text{ donde } c_1 \text{ es función de } c. \\ &= \left\{(X_1,\ldots,X_n) : \left|\frac{\bar{X}-\theta_0}{\frac{\sigma}{\sqrt{n}}}\right| \ge c_2\right\} \end{aligned}$$

*El valor de $c_2$ se obtendrá de acuerdo a la ecuación (1.16), donde*

$$\alpha = \sup_{\theta=\theta_0}\boldsymbol{P}_\theta(\lambda(\boldsymbol{X}) \le c) = \sup_{\theta=\theta_0}\boldsymbol{P}_\theta(\lambda(\boldsymbol{X}) \le c) = \sup_{\theta=\theta_0}\boldsymbol{P}_\theta\left(\left|\frac{\bar{X}-\theta_0}{\frac{\sigma}{\sqrt{n}}}\right| \ge c_2\right)$$

*Ahora, sabemos que* $\frac{\bar{X}-\theta_0}{\sigma/\sqrt{n}} \sim N(0,1)$*, por lo que el valor de* $c_2$ *se obtiene al fijar un valor* $\alpha$*, así* $c_2 = z_{\alpha/2}$*. Entonces la región crítica es igual a*

$$RC = \left\{(X_1 \dots X_n) : \left|\bar{X} - \theta_0\right| \geq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right\}$$

*Entonces rechazaremos* $H_0$ *si* $\left|\bar{X} - \theta_0\right| \geq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

*Obteniendo el intervalo de confianza*

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \theta_0 < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

## 1.3. Tamaño de muestra ($n$)

El tamaño de muestra influye de manera importante en la inferencia estadística. Por ejemplo, la elección de un tamaño de muestra muy pequeño puede provocar una pérdida de potencia en las pruebas de hipótesis, además de que puede ser difícil detectar resultados significativos. Al aumentar el tamaño de muestra es más probable encontrar significancia estadística cuando ésta realmente existe, así como la obtención de un aumento en la potencia. Una muestra demasiado grande podría provocar resultados erróneos, ya que mediante un aumento en el tamaño de muestra, los efectos son cada vez menos significativos, hasta que la muestra se vuelve demasiado grande y entonces cualquier efecto puede resultar significativo (F. Hair et al., 2006), lo cual puede provocar que se generen conclusiones falsas sobre las hipótesis planteadas.

### 1.3.1. Consistencia y eficiencia

Dos propiedades importantes de mencionar son la **consistencia**, en la que un estimador es consistente cuando éste converge al valor verdadero cuando $n \to \infty$, y la eficiencia en la que un estimador se dice que es más eficiente que otro si tiene menor variabilidad, lo cual lo hará un estimador más preciso.

**Definición 1.11** *Una sucesión de estimadores* $\hat{\theta}_n$ *es una sucesión consistente del parámetro* $\theta$ *si, para cada* $\epsilon > 0$ *y para cada* $\theta \in \Theta$*,*

$$\lim_{n\to\infty} \boldsymbol{P}(|\hat{\theta}_n - \theta| < \epsilon) = 1$$

Ejemplo **1.6** *Consistencia de* $\bar{X}$*.*
*Sea* $X_1, \dots, X_n$ *una sucesión de variables aleatorias independientes e identicamente distribuidas con* $\mathbb{E}(X_i) = \mu$ *y* $Var(X_i) = \sigma^2 < \infty$*. Demuestre que la media muestral es un estimador consistente de la media poblacional. Entonces utilizando convergencia en probabilidad y aplicando una de las formas de la desigualdad de Chevyshev se tiene*

$$\boldsymbol{P}(|\bar{X} - \mu| \geq \epsilon) = \boldsymbol{P}(|\bar{X} - \mu| \geq k\sigma_{\bar{x}}) \leq \frac{1}{k^2} \quad \forall\epsilon > 0 \tag{1.19}$$

$\epsilon = k\sigma_{\bar{x}} \Rightarrow k = \frac{\epsilon}{\sigma_{\bar{x}}}$

*Entonces*

$$\boldsymbol{P}(|\bar{X} - \mu| \geq \frac{\epsilon}{\sigma_{\bar{x}}}\sigma_{\bar{x}}) \leq \frac{1}{(\epsilon/\sigma_{\bar{x}})^2}$$

*Sabemos que*

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{\sum_i = 1^n X_i}{n}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \mu$$

$$\sigma_{\bar{x}} = V(\bar{X}) = V\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n}V(X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \quad \boldsymbol{P}(|\bar{X} - \mu| \geq \frac{\epsilon}{\sigma_{\bar{x}}}\sigma_{\bar{x}}) \leq \frac{\sigma_{\bar{x}}^2}{\epsilon^2} = \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

*Entonces*

$$\lim_{n\to\infty} \boldsymbol{P}(|\bar{X} - \mu| \geq \epsilon) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0$$

$$\Rightarrow$$

$$\lim_{n\to\infty} \boldsymbol{P}(|\bar{X} - \mu| \geq \epsilon) = 0$$

*Por lo que $\bar{X}$ es un estimador consistente para $\mu$.*

**Definición 1.12** *Sea $\hat{\theta}_1$ y $\hat{\theta}_2$ dos estimadores insesgados de $\theta$. Entonces $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ si*

$$var(\hat{\theta}_1) < var(\hat{\theta}_2)$$

Ejemplo **1.7** *Eficiencia.*

*Sea $\hat{\theta}_1 = \sum_{i=1}^n X_i/n$ y $\hat{\theta}_2 = \sum_{i=1}^{n-1} X_i/(n-1)$ estimadores insesgados para la media de una población con distribución $N(\mu, \sigma^2)$ y $\sigma^2$ conocida. Obtenga el estimador más eficiente para $\mu$.*

$$var(\hat{\theta}_1) = var\left(\sum_{i=1}^{n} X_i/n\right) = \sigma^2/n$$

$$var(\hat{\theta}_2) = var\left(\sum_{i=1}^{n-1} X_i/(n-1)\right) = \sigma^2/(n-1)$$

*Entonces*

$$var(\hat{\theta}_1) < var(\hat{\theta}_2)$$

*Por lo que $\hat{\theta}_1$ es un estimador más eficiente que $\hat{\theta}_2$ para la media poblacional.*

### 1.3.2. Intervalos de confianza

Otra forma en la que se puede ver el efecto del tamaño de muestra es en los intervalos de confianza. En la figura 1.1 se muestran los intervalos de confianza para la media de una población normal bajo distintos tamaños de muestra, en donde se puede observar que conforme aumenta el tamaño de muestra, el intervalo se hace más pequeño acercándose más al valor real de la media.

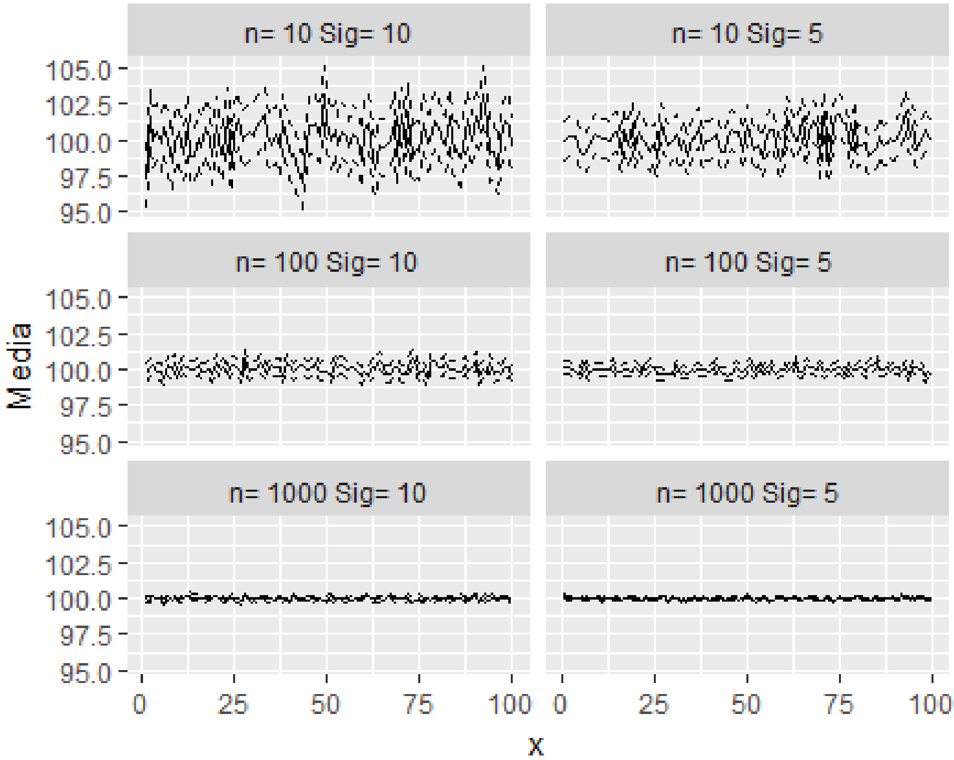


Figura 1.1: Intervalos de confianza para la media de una población normal, donde n=tamaño de muestra y Sig = $\sigma$.

## 1.4. Análisis de varianza y Regresión lineal

En la siguiente sección se presentarán brevemente las metodoligias de análisis de varianza y regresión lineal. En la primera parte se muestra el analísis de varianza (ANOVA), la cual busca comprobar si varias medias poblacionales son iguales.
En la segunda parte de la sección se muestra el método de regresión lineal cuyo principal objetivo es estudiar la dependencia de una variable respecto a otra.

### 1.4.1. Análisis de varianza

En el análisis de varianza se comparan dos o más medias para saber si las poblaciones asociadas a éstas son estadísticamente diferentes. Para probar si las medias son diferentes se compara la varianza explicada por las distintas poblaciones o tratamientos, desde un punto de vista de diseño experimental, con la varianza no explicada debida al error. Cuando la varianza explicada predomina sobre la varianza no explicada se concluye que las medias son diferentes por lo que las poblaciones difieren.

La hipótesis a probar es

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \quad vs \quad H_1 : \mu_i \neq \mu_j \quad \text{para algún } i \neq j.$$

Donde $\mu_k$ es la media del tratamiento $k$.

En donde se supone que las $k$ poblaciones tienen distribución normal, son independientes y con misma varianza.

Suponiendo además que las variables aleatorias $y_{ij}$ pueden expresarse de acuerdo al modelo

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \text{con} \quad i = 1, \ldots, k \quad \text{y} \quad j = 1, \ldots, n_i.$$

Donde $i$ son los grupos y $j$ la observación de cada tratamiento (población), con $\mu$ una media general más una desviación $\tau_i$ causada por el tratamiento $i$ y $\epsilon_{ij}$ los errores, los cuales se suponen independientes y de distribución normal con media cero y varianza constante.

La idea original del análisis de la varianza está basada en la descomposición de la ecuación

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{\bar{y}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2 + \sum_{i=1}^{k} n_i(\bar{y}_i - \bar{\bar{y}}) \tag{1.20}$$

Donde $\bar{y}_i$ es la media de la población $i$ y $\bar{\bar{y}}$ es la media total, es decir $\bar{\bar{y}} = \frac{1}{\sum_{i=1}^{k} n_i}\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}$.

El primer término del lado izquierdo en la ecuación (1.20) suele llamarse suma de cuadrados total ($SC_T$), la segunda suma de cuadrados dentro de los grupos o residual ($SC_D$) y el tercer término suma de cuadrados entre los grupos o explicada ($SC_E$), por lo que se obtiene que $SC_T = SC_D + SC_E$.

Dada la suposición de normalidad $SC_T$, $SC_D$ y $SC_E$ tienen una distribución $\chi^2$ y en particular $SC_D$ y $SC_E$ son independientes.

Para un análisis más profundo de dichos resultados consultar Casella and Berger (2002).

Los datos obtenidos en el análisis de varianza suelen presentarse en forma de tabla llamada tabla de análisis de la varianza mostrada en la tabla 1.1.

El estadístico F tiene una distribución bajo normalidad $F_{k-1,N-k}$ con $N = \sum n_i$, el cual corresponde al cociente de verosimilitudes para probar la hipótesis de igualdad de medias y para el caso de dos poblaciones la estadística de prueba corresponde a la estadística $t$ para diferencia de medias elevada al cuadrado.

### 1.4.2. Regresión lineal

En el análisis de regresión se busca poner a una variable en función de otra u otras variables y así explicar el comportamiento de esta a través de su relación con otras. El caso más

| Fuente de variación | Suma de cuadrados | $g.l.$ | Media cuadrática | Estadístico F |
|---|---|---|---|---|
| Entre grupos | $\mathrm{SC}_E = \sum_{i=1}^{k} n_i(\bar{y}_i - \bar{\bar{y}})$ | k-1 | $\mathrm{MC}_E = \frac{SC_E}{k-1}$ | $\frac{MC_E}{MC_D}$ |
| Dentro de los grupos | $\mathrm{SC}_D = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2$ | N-k | $\mathrm{MC}_D = \frac{SC_D}{N-k}$ | |
| Total | $\mathrm{SC}_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{\bar{y}})^2$ | N-1 | | |

Tabla 1.1: Tabla ANOVA para análisis de varianza de un factor. Basada en Casella and Berger (2002).

sencillo es la regresión lineal simple en la cual una variable $Y$ es explicada por sólo una variable $X$, dicho modelo se puede representar mediante la recta $Y = \beta_0 + \beta_1 X + \epsilon_i$, en donde $Y$ es la variable dependiente y $X$ la variable independiente.

El procedimiento más usual para la estimación de los coeficientes $\beta_0$ y $\beta_1$ es el de mínimos cuadrados en donde lo que se busca es encontrar la recta que minimice la suma de los cuadrados de las distancias verticales entre cada punto y la recta con el fin de minimizar la suma de los errores cuadrados

$$SC = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 \tag{1.21}$$

Derivando la ecuación (1.21) respecto $\beta_0$ y $\beta_1$

$$\begin{aligned} \frac{\partial SC}{\partial \beta_0} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \\ \frac{\partial SC}{\partial \beta_1} &= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i = 0 \end{aligned} \tag{1.22}$$

Obteniendo del resultado anterior las ecuaciones normales

$$\begin{aligned} \sum_{i=1}^{n} y_i &= n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i x_i &= \beta_0 \sum_{i=1}^{n} x + \beta_1 \sum_{i=1}^{n} x_i^2 \end{aligned} \tag{1.23}$$

Resolviendo el sistema anterior se obtiene que los estimadores para $\hat{\beta}_0$ y $\hat{\beta}_1$ son

$$\begin{aligned} \hat{\beta}_0 &= \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} = \bar{Y} - \hat{\beta}_1 \bar{X} \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^{n} x_i y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} x_i^2 - n\bar{X}^2} \end{aligned} \tag{1.24}$$

Donde $S_{xy} = \sum_{i=1}^{n} x_i y_i - n\bar{X}\bar{Y}$ y $S_{xx} = \sum_{i=1}^{n} x_i^2 - n\bar{X}^2$

Se puede comprobar que dichos estimadores son mínimos haciendo el cálculo de la matriz Hessiana y verificar que es definida positiva, dicho resultado se observa en la ecuación

(1.25). Dichos estimadores corresponden a los estimadores máximo verosímiles cuando los errores, y por tanto $Y$, tienen una distribución normal.

$$\frac{\partial^2 SC}{\partial \beta_0^2} = 2n$$
$$\frac{\partial^2 SC}{\partial \beta_1^2} = 2\sum_{i=1}^{n} x_i^2$$
$$\frac{\partial^2 SC}{\partial \beta_1 \beta_0} = \frac{\partial^2 SC}{\partial \beta_0 \beta_1} = 2\sum_{i=1}^{n} x_i$$

Entonces

$$\begin{vmatrix} 2n & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2\sum_{i=1}^{n} x_i^2 \end{vmatrix} = 4n\sum_{i=1}^{n} x_i^2 - 4\left(\sum_{i=1}^{n} x_i\right)^2 = 4n\left(\sum_{i=1}^{n} x^2 - n\bar{x}^2\right)$$
$$= 4n\sum_{i=1}^{n}(x_i - \bar{x})^2 > 0 \tag{1.25}$$

La última igualdad de la ecuación (1.25), se obtiene del resultado obtenido en la ecuación (1.1) del ejemplo 1.1.

En caso de que la variable dependiente se explique por dos o más variables predictoras entonces se habla de una **regresión lineal múltiple** en donde el análisis se basa en el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \epsilon_i \qquad \text{con } i = 1, \ldots, n \tag{1.26}$$

Donde $x_{ji}$ es el valor de la $i$-ésima observación de la covariable $j$.

Para todas las observaciones la ecuación (1.26) puede escribirse de manera matricial como $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, es decir

$$\begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & \cdots & x_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2i} & \cdots & x_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \vdots \\ \epsilon_j \\ \vdots \\ \epsilon_k \end{bmatrix}$$

Haciendo uso del método de mínimos cuadrados, para minimizar $\sum_{i=1}^{n} \epsilon_i^2$ se puede hacer uso de la transpuesta del vector de errores, es decir, minimizar $\sum_{i=1}^{n} \epsilon_i^2$ es equivalente a minimizar $\epsilon^t\epsilon$, entonces

$$SC = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon^t\epsilon = (\mathbf{Y} - \mathbf{X}\beta)^t(\mathbf{Y} - \mathbf{X}\beta) \tag{1.27}$$

Entonces

$$SC = (\mathbf{Y}^t - \beta^t\mathbf{X}^t)(\mathbf{Y} - \mathbf{X}\beta) = \mathbf{Y}^t\mathbf{Y} - 2\mathbf{Y}^t\mathbf{X}\beta + \beta^t\mathbf{X}^t\mathbf{X}\beta$$

Derivando la ecuación anterior respecto de $\beta$ se obtiene

$$\frac{\partial SC}{\partial \beta} = -2\mathbf{X}^t\mathbf{Y} + 2\mathbf{X}^t\mathbf{X}\beta = 0$$

Entonces si $det(\mathbf{X}^t\mathbf{X}) \neq 0$ y $(\mathbf{X}^t\mathbf{X})^{-1}$ existe, se obtene que

Las ecuaciones normales son $(\mathbf{X}^t\mathbf{X})\beta = \mathbf{X}^t\mathbf{Y} \Rightarrow (\mathbf{X}^t\mathbf{X})^{-1}(\mathbf{X}^t\mathbf{X})\beta = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Y}$

y además

$$\hat{\beta} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Y}$$

Para verificar que $\hat{\beta}$ es un mínimo basta comprobar que $\frac{\partial^2 SC}{\partial \beta^2} = 2\mathbf{X}^t\mathbf{X}$ es una matriz definida positiva.

Entonces para comprobar que $\mathbf{X}^t\mathbf{X}$ es definida positiva es suficiente ver que

$$\mathbf{U}^t\mathbf{X}^t\mathbf{X}\mathbf{U} > 0$$

Donde $\mathbf{U}$ es cualquier vector distinto del vector nulo.

Se tiene que $\mathbf{U}^t\mathbf{X}^t\mathbf{X}\mathbf{U} = (\mathbf{X}\mathbf{U})^t(\mathbf{X}\mathbf{U}) = ||\mathbf{X}\mathbf{U}||^2 > 0$

Los supuestos en los que se basa el análisis de regresión son:

- Linealidad entre $X$ y $Y$, es decir, se presenta una relación significativa entre la variable que se quiere predecir y las variables predictoras.
- Independencia de los residuos $e_i = y_i - \hat{y}_i$, en este supuesto se asume que los residuos no están auto-correlacionados, por lo cual son independientes $(E(e_i, e_j) = 0)$.
- Los residuos $e_i$ tienen media cero y varianza constante (homocedasticidad).
- No debe haber multicolinealidad, es decir ningún regresor $(x_i)$ debe ser una combinación lineal de otros regresores.

  Adicionalmente, cuando se desea hacer inferencia se supone que los errores y por tanto $Y$ tienen una distribución normal (con media cero y varianza constante).

Existen diversas medidas para determinar la **bondad de ajuste** del modelo que estamos aplicando y así saber qué tan bueno es el modelo, una medida, quizá la más popular, de bondad de ajuste suele ser el coeficiente de determinación $R^2$ el cual se basa en la descomposición de la ecuación (1.20)

$$R^2 = 1 - \frac{SC_D}{SC_T} \tag{1.28}$$

Donde $SC_D$ es la suma de cuadrados residual y $SC_T$ es la suma de cuadrados total.

El coeficiente $R^2$ obtiene valores entre cero y uno, tomando el valor cero cuando no existe variabilidad explicada y el valor de uno cuando el ajuste es perfecto, es decir la varianza residual es cero, por lo que con un valor de $R^2$ cercano a uno se obtendrá un mayor porcentaje de variabilidad explicada por el conjunto de variables predictoras, obteniendo así un mejor modelo.

Al aumentar el número de variables $R^2$ siempre aumentará o por lo menos se mantendrá el valor de este coeficiente sin importar si las variables agregadas aportan al modelo.

# Capítulo 2

# El tamaño del efecto

Las pruebas de significancia estadística, son quizá, el método más conocido en inferencia estadística, sin embargo éste puede llegar a tener algunas deficiencias al momento de realizar conclusiones sobre el análisis de los datos, debido a que no es suficiente saber si existe significancia o no.

Sería de gran utilidad obtener un resultado mediante el cual se pueda saber que tan grande o pequeño resulta el fenómeno planteado en términos de la pregunta de interés, con el que se pueda saber si es lo suficientemente grande para ser considerado importante o demasiado pequeño como para no mostrar relevancia en la pregunta de interés. Por lo que el uso del tamaño del efecto resulta de gran utilidad para dar una mejor interpretación de los resultados.

## 2.1. Definición

Existen muchas definiciones sobre el tamaño del efecto por ejemplo Grissom y Kim afirman que "mientras que una prueba de significación estadística se utiliza tradicionalmente para proporcionar evidencia de que una hipótesis nula es incorrecta, un tamaño del efecto (TE) mide el grado en que tal hipótesis nula es incorrecta (sólo para el caso en el que la hipótesis nula es falsa)"(Grissom and Kim, 2012). Del mismo modo Cohen (1988) define el tamaño del efecto como "el grado en que el fenómeno se presenta en la población" , o "el grado en el que la hipótesis nula es falsa". También Vacha-Haase and Thompson (2004) expresan que el término tamaño del efecto lo usarán para "referirse a cualquier estadístico que cuantifica el grado en que los resultados de la muestra divergen de la hipótesis nula".

En otras palabras se puede definir al tamaño del efecto como un valor que indica el grado en el que nuestro resultado difiere o se aproxima a lo esperado, tomando como esperado aquella pregunta de interés planteada en la hipótesis nula.

## 2.2. Facetas del tamaño del efecto

Kelley and Preacher (2012) describen tres facetas. La primera faceta del tamaño del efecto aborda el tipo de información de interés (dimensión). La segunda es sobre la ecuación mediante la cual se obtiene la magnitud del efecto (medida). Finalmente la tercera es el número que se obtiene de la medida del tamaño del efecto (valor). Dichas facetas se describen a continuación.

### 2.2.1. Dimensión del tamaño del efecto

La dimensión del tamaño del efecto es una abstracción de una cualidad cuantificable de forma generalizada que no tiene una unidad en particular. Un ejemplo de una dimensión del tamaño del efecto es la variabilidad, que puede llevarse a la práctica en unidades de varianza, desviación estándar, rango, rango intercuartil, etc. Las medidas del tamaño del efecto pueden ser multidimensionales, unidimensionales o adimensionales, las cuales afectan la forma en que un tamaño del efecto puede ser interpretado.

Cuando una dimensión no depende de otras dimensiones y además no puede ser reducida, se dice que es unidimensional. Por ejemplo, la diferencia de medias definida como: $\Delta = \mu_1 - \mu_2$ es una diferencia unidimensional ya que no depende de otras abstracciones, en donde su única dimensión es la tendencia central. Sin embargo, las dimensiones, tales como la diferencia de medias estandarizadas son multidimensionales, ya que se definen en términos de otras dimensiones, la cual se define para la población como: $\delta = \frac{\mu_1 - \mu_2}{\sigma}$ por lo tanto, la diferencia estandarizada de medias es un tamaño del efecto bidimensional, ya que se compone de dos dimensiones fundamentales: la tendencia central y la variabilidad (que son fundamentales, ya que están en el nivel más básico).

Algunas medidas del tamaño del efecto son abstracciones adimensionales, ya que el tamaño del efecto no se basa en algún componente dimensional específico (medida del tamaño del efecto =1). Las variables sin dimensiones describen lo que se puede considerar una variable natural, que es un número natural que permite describir una característica sin dimensión ni unidad de expresión explícita. Los números naturales tienen una unidad natural de 1. Por ejemplo el coeficiente de determinación múltiple definido como: $R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$, la dimensión del tamaño del efecto en el numerador es la misma que en el denominador (Suma de desviaciones al cuadrado) por lo que $R^2$ tiene dimensión igual a 1, más no una dimensión, entonces $R^2$ es adimensional.

### 2.2.2. Medida del tamaño del efecto

La medida del tamaño del efecto, es la ecuación que define la implementación de la dimensión o las dimensiones de interés del tamaño del efecto. Por ejemplo, una medida del tamaño de efecto utilizada para operacionalizar la dimensión del tamaño de efecto de la separación entre dos medias de un grupo, es la diferencia estandarizada de medias, que se define como $d = \frac{(\bar{X}_1 - \bar{X}_2)}{S}$, donde $S$ es la raíz cuadrada del estimador insesgado de la varianza dentro del grupo.

La idea de la medida del tamaño del efecto es que representa de una manera muy precisa en que los datos, estadísticas o parámetros, son utilizados para implementar una dimensión del tamaño de efecto con el fin de abordar alguna pregunta de interés.

### 2.2.3. Valor del tamaño del efecto

El número real que se obtiene de aplicar una medida del tamaño del efecto a los datos, estadísticas o parámetros es llamado valor del tamaño del efecto.

El valor del tamaño del efecto es, literalmente, la magnitud de un fenómeno. Por ejemplo, el valor obtenido para la medida del tamaño del efecto "la diferencia estandarizada de medias" que en un estudio particular puede ser $d = 0.70$. El valor de 0.70 es un número real que resulta de aplicar los datos a una medida del tamaño del efecto (es decir, a una ecuación o algoritmo).

Luego de mostrar cada una de las facetas del tamaño de efecto se tiene que; la dimensión del tamaño del efecto, es un término más general que habla del tipo de tamaño del efecto que es de interés, mientras que la medida del tamaño del efecto muestra de manera más específica la forma en que se opera la dimensión del tamaño del efecto a través de una ecuación, obteniendo de ésta un número llamado valor del tamaño del efecto.

## 2.3. Características y propiedades del tamaño del efecto

En esta sección se muestran algunas características y propiedades sobre el tamaño del efecto, las cuales son descritas en los corolarios presentados por Kelley and Preacher (2012), para dar un mejor uso e interpretación al tamaño del efecto.

### 2.3.1. Características del tamaño del efecto

En aspectos generales, una de las características más intuitivas sobre el tamaño del efecto es que éste puede representar valores teóricos, de la muestra o de la población. Estos valores del tamaño del efecto claramente pueden ser distintos, pero el valor representado debe ser especificado para evitar interpretaciones erróneas.

En cuanto a la interpretación del tamaño del efecto, ésta puede ser comparativa o absoluta; es decir, al ser absoluta, el valor del tamaño del efecto no necesita ningún parámetro de referencia para su interpretación, mientras que cuando se necesita de uno o varios valores de referencia la interpretación será comparativa. Los tamaños del efecto comparativos se utilizan cuando hay alguna comparación explícita que es importante cuantificar. Por ejemplo, el número de éxitos que se obtuvieron en determinado estudio es un valor absoluto, mientras que si lo que se obtiene es la proporción de éxitos respecto al total de eventos, entonces hablamos de una interpretación comparativa.

El valor del tamaño del efecto puede ser una cantidad estática para algunos fenómenos. Otros fenómenos pueden cambiar en función del tiempo, el contexto o las características de la población por lo que la interpretación puede cambiar, ya que en una instancia (por ejemplo, tiempo o población) puede obtenerse un valor distinto a otra instancia (es decir, en otro momento o para otra población). Por lo que el valor del tamaño del efecto puede ser dinámico y modelado a lo largo del tiempo. Por ejemplo, el tamaño del efecto que alguna vez existió (datos históricos) o que puede existir en el futuro (datos especulativos).

#### 2.3.1.1. Tamaño del efecto estandarizado, no estandarizado y parcialmente estandarizado

La idea de un tamaño de efecto estandarizado es determinar si las unidades de medida se cancelan, creando un tipo de tamaño del efecto que no necesita ninguna etiqueta de la unidad, denominado un tamaño de efecto estandarizado. Ésto resulta similar a la dimensión, sin embargo, la diferencia se encuentra en si el análisis se refiere a las dimensiones o las unidades de medida. Un tamaño de efecto estandarizado es uno en el que las unidades de medida se pueden cancelar y no necesariamente las dimensiones también, de manera que un tamaño del efecto estandarizado podría ser adimensional o de alguna otra dimensión. Por lo que las unidades de medida no son fundamentales para la interpretación del valor del tamaño del efecto. Por ejemplo, la diferencia estandarizada de medias es bidimensional y no necesita una unidad de medición específica. Por otro lado, un tamaño del efecto puede tener unidades de medición específicos (por ejemplo, un coeficiente de regresión no estandarizado, que se escala en términos de las escalas de medición particulares de las variables regresoras y las variables respuesta), a este tipo se le dice no estandarizado y su interpretación es dependiente de las unidades del instrumento de medición.

Otro tipo de escala del tamaño del efecto es la estandarización parcial. Se dice que un tamaño del efecto puede ser estandarizado parcialmente cuando al menos una componente del tamaño del efecto está estandarizado y al menos una componente es no estandarizado. En tales situaciones, la interpretación de la magnitud del efecto es al menos parcialmente sobre la base de una o más unidades de medición. Por ejemplo, la solución parcialmente estandarizada en un modelo de regresión múltiple es una en la que la estandarización de las variables independientes, pero no la variable dependiente, conduce a coeficientes de regresión que están estandarizados en parte.

#### 2.3.1.2. Tamaño del efecto independiente de la tasa-base y dependiente de la tasa-base

Un tamaño del efecto es dependiente de la tasa base cuando éste depende del tamaño de muestra o población. Por ejemplo, los elementos que cumplen cierta condición en determinado grupo (éxitos), en éste la proporción de éxitos depende del tamaño del grupo por lo que es dependiente de la tasa base, mientras que el número de éxitos es independiente del tamaño de grupo, en este último se dice que el tamaño de la muestra es independiente de la tasa base. Esto es, el tamaño del efecto puede o no depender de una tasa base (es decir, una proporción en una población o muestra).

#### 2.3.1.3. Tamaño del efecto *Ómnibus, Targeted y Semiomnibus*

Un tamaño de efecto *ómnibus* es aquel en el que se describen efectos generales. Por otro lado, un tamaño de efecto *targeted* cuantifica un efecto específico, generalmente un efecto *ómnibus* se basa en una colección de efectos *targeted*. Algunos tamaños del efecto son una combinación de un efecto *targeted* y *ómnibus*, denominados como *semitargeted o semi-ómnibus*. Por ejemplo en una regresión múltiple a través del coeficiente de correlación múltiple se puede observar la efectividad global del modelo lo que representaría un efecto *ómnibus*, y al determinar el efecto específico de cada variable explicativa sobre la variable dependiente $Y$, se estaría hablando de un efecto *targeted*, y por último, si se considera el aumento en el coeficiente de correlación múltiple al cuadrado al agregar un $x_3$ y un $x_4$ a un modelo en donde ya se encuentra $x_1$ y $x_2$ modelando a $Y$ será un efecto *semi-ómnibus* en donde sólo se conocerá el cambio agregado pero no la contribución exacta de cada componente.

### 2.3.2. Significancia sustantiva del tamaño del efecto

En la significancia estadística se obtiene un resultado preciso a través de una probabilidad, en donde lo único que se concluye es si el resultado obtenido es estadísticamente significativo o no, mientras que en la significancia sustantiva se puede hablar sobre la importancia del resultado obtenido o lo que representa dicho valor, dependiendo del contexto y del fenómeno que se está estudiando, siendo la significancia sustantiva de gran utilidad al momento de crear las conclusiones. Sin embargo, al no tener un significado formal y bien definido ésta puede tener ciertas inconsistencias ya que la interpretación suele ser en algunos casos subjetiva.

En cuanto a la interpretación, sería un error intentar asignar etiquetas para clasificar de manera general los tamaños del efecto por su valor. Glass et al. (1981) argumentan que “no hay sabiduría alguna al intentar asociar regiones de la métrica del tamaño del efecto con adjetivos descriptivos como “pequeño”, “moderado”, “grande” y “similares” . En otras palabras, se puede decir que no se puede hacer una clasificación para el valor del tamaño del efecto, como “pequeño, mediano o grande” de manera general, para cualquier valor del tamaño del efecto, ya que éste puede variar de un área a otra, puesto que un valor pequeño que se considera marginal para alguna área, en otra un valor igual de pequeño podría ser de gran importancia.

### 2.3.3. Propiedades del tamaño del efecto

Kelley and Preacher (2012) proporcionan varias propiedades deseables que un buen tamaño del efecto debería cumplir, en donde los tamaños del efecto que mejor satisfagan dichas propiedades serán preferibles. Sin embargo, siempre pueden ser presentados más de un tamaño del efecto con el fin de obtener mejores conclusiones sobre la pregunta de interés. Las propiedades de un buen tamaño del efecto son:

1. Los valores del tamaño del efecto deben escalarse de manera apropiada, dada la medición y la pregunta de interés.

2. Los valores de tamaño del efecto deben ir acompañados de intervalos de confianza.

3. La estimación puntual del valor del tamaño del efecto poblacional debe ser independiente del tamaño de la muestra.

4. Las estimaciones de los valores del tamaño del efecto deben tener propiedades de estimación deseables; es decir, deben ser imparciales (sus valores esperados deben ser iguales a los valores poblacionales correspondientes), consistentes (deben converger al valor poblacional correspondiente a medida que aumenta el tamaño de la muestra) y eficientes (deben tener una variación mínima entre las medidas competidoras).

# Capítulo 3

# Cálculo del tamaño del efecto

En este capítulo se muestra el cálculo del tamaño del efecto mediante tres índices, los cuales son útiles cuando las mediciones no tienen un significado intrínseco, por lo que no se puede hacer una comparación directa, ya sea porque se han utilizado escalas distintas o porque el tamaño del efecto es estudiado en términos de la variabilidad de la población de estudio.

El primer índice mostrado es el índice $d$ para diferencia de medias estandarizadas visto como un porcentaje de no superposición, así como en términos de correlación y proporciones. El índice $f$ para análisis de varianza a través de una medida de rango estandarizado y en términos de correlación y proporción de varianza. Por último se muestra el índice $f^2$ para análisis de regresión a través de la correlación parcial y semiparcial múltiple cuadrada. Se muestran además al final de cada sección los valores de referencia para un tamaño de efecto pequeño mediano y grande para cada uno de los índices, para el caso en el que no se cuente con un valor propio de referencia.

## 3.1. Índice $d$ para la prueba $t$ de diferencia de medias

El índice $d$ de Cohen es una medida estandarizada en la que el tamaño del efecto, se calcula como:

$$d = \frac{\mu_A - \mu_B}{\sigma} \tag{3.1}$$

Donde $\mu_A$ y $\mu_B$ son medidas asociadas a dos muestras independientes y $\sigma$ es la desviación estándar de ambos grupos definida de manera general como

$$\sigma = \sqrt{\frac{(n_a - 1)\sigma_a^2 + (n_b - 1)\sigma_b^2}{n_a + n_b - 2}}. \tag{3.2}$$

En cuanto a su interpretación, se puede decir que $d$ representa el número de desviaciones típicas que separan a dos grupos. Más adelante se mostrarán los valores de referencia dados por Cohen para interpretar la magnitud de los tamaños del efecto.

Por ejemplo si se tiene una población A con una media de 280 y la población B con una media de 270, la diferencia es de 10 puntos, pero qué representan esos diez puntos, qué tan grande es cada punto; entonces al utilizar la $d$ de Cohen es posible responder a estas preguntas expresando estas distancias que son representadas por puntos en unidades de variabilidad, de este modo si la desviación estándar común en los grupos anteriores es $\sigma = 100$, entonces se obtendrá una $d = 0.1$, con la cual concluye que las medias difieren en una décima parte de una desviación estándar. En otro caso si se tuviera $\sigma = 5$ entonces $d$ sería igual a 2, lo que diría que las medias difieren en dos desviaciones estándar y esto es mucho más grande que $d = 0.1$. Pero, qué tan grande es el segundo respecto al primero. Cohen(1988) muestra varias formas en las que se puede entender el valor de $d$, dichas formas se exponen a continuación.

### 3.1.1. $d$ como un porcentaje de no superposición: las medidas $U$

Bajo la suposición de que las poblaciones son normales, con igual varianza y del mismo tamaño, Cohen (1988) define a las medias de no superposición $U_1, U_2$ y $U_3$, las cuales se representan a partir de los percentiles de la distribución normal estándar ($P_d$), en donde $d$ será la desviación en la curva de la normal, entonces

$$\begin{aligned} U_3 &= P_d \\ U_2 &= P_{d/2} \\ U_1 &= \frac{2U_2 - 1}{U_2} \end{aligned}$$

Por ejemplo en el caso en el que $d = 0$ entonces, las medidas de no superposición serían:

$U_1 = 0$; lo que dice que hay un 0 % de área de no superposición, es decir, la superposición de la población A es absoluta con la población B.
$U_2 = 0.5$; entonces el 50 % de la población B excede el 50 % de la población A (considerando que $\mu_B > \mu_A$).
$U_3 = 0.5$; esto indica que el 50 % de la población A es superada por la media de la población B.

Ahora si $d = 3$ entonces las medidas de no superposición serían

$U_1 = 92.8\,\%$; lo que dice que prácticamente toda el área de la población A y la población B no se superponen al haber un 92.8 % de no superposición.
$U_2 = 93.3\,\%$; entonces se dice que el 93.3 % de la población B excede el 93.3 % de la población A (considerando que $\mu_B > \mu_A$).
$U_3 = 99.9\,\%$; esta última medida indica que el 99.9 % de la población A es superada por la media de la población B.

### 3.1.2. $d$ en términos de correlación y proporciones

Una forma más general de representar el tamaño del efecto sería considerar la pertenencia en la población A o en la población B a través de una dicotomía, en la que para medir el grado de asociación entre una variable dicotómica $X$ y una variable continua $Y$, se realiza mediante el coeficiente de correlación de Pearson ($r$) mostrado en términos de $d$. Cohen (1988) propone una fórmula para representar a $r$ en términos de $d$, presentada en la ecuación (3.3):

$$r = \frac{d}{\sqrt{d^2 + 1/pq}} \tag{3.3}$$

Donde $p$ y $q$ son las proporciones de los grupos A y B respectivamente, siendo $p$ igual al número de individuos en el grupo A ($n_A$) dividido por el número total de individuos en el estudio ($n_A + n_B = n$), por lo que $p = n_A/n$ y $q = 1 - p = n_B/n$. Note que cuando $n_A = n_B$ las proporciones son iguales y el término $(1/pq) = 4$ reduciendo la ecuación (3.3) a $r = \frac{d}{\sqrt{d^2+4}}$.

Tomando una $d$ fija igual a 0.8, en la figura 3.1, se puede observar que conforme la diferencia de las proporciones $p$ y $q$ disminuye, $r$ aumenta; ya que a medida que $p$ y $q$ se separan el denominador se hace mayor, por lo que el valor del denominador aumenta y por consiguiente la correlación disminuye.

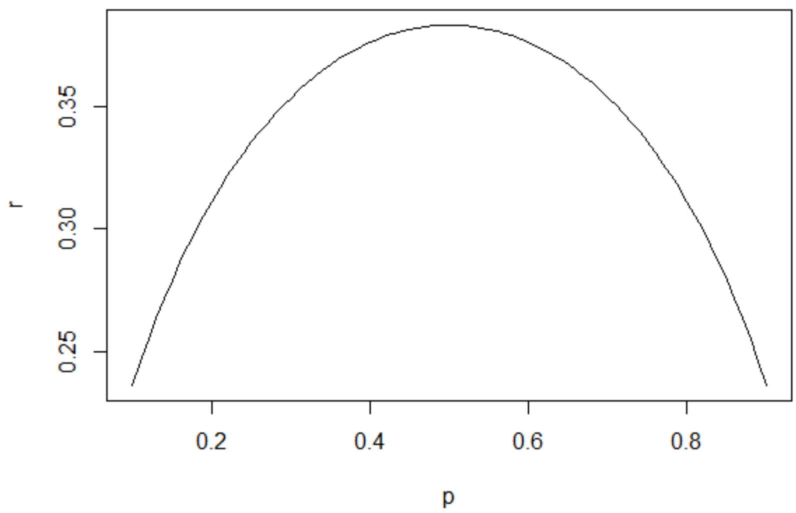


Figura 3.1: comportamiento de $r$ conforme las proporciones se acercan.

Considerando $p = q$, en la figura 3.2 se puede observar que conforme $d$ aumenta el valor de $r$ crece de manera asintótica a 1.

La estimación del tamaño del efecto para medir la relación entre dos variables mediante la conversión a $r$ siempre dará un número entre 0 y 1, lo cual puede ser una ventaja a diferencia de la estimación de $d$ mostrada en la ecuación (3.1) pues su valor no se encuentra acotado. Entonces el uso de $r$ nos puede ayudar a una interpretación más fácil, además $r$

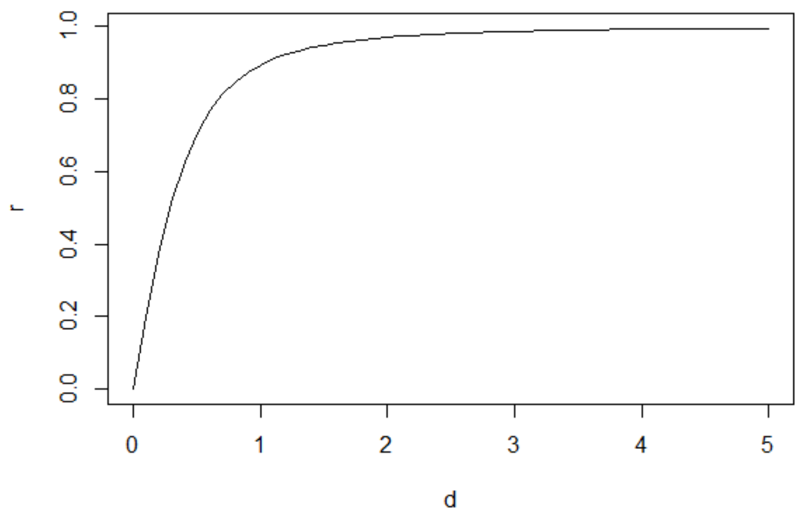


Figura 3.2: comportamiento de r conforme $d$ aumenta.

también puede ser utilizada con su cuadrado, la cual representa el porcentaje de variación total observada en la variable dependiente. El coeficiente de determinación $r^2$ suele ayudar en la obtención de otros resultados.

### 3.1.3. Valores pequeños, medianos y grandes de $d$

Los términos pequeño, mediano y grande son relativos pues éstos dependen del área y tipo de investigación que se ésta realizando. Por lo que anteriormente se mencionó que no se puede dar valores de referencia generales para el tamaño del efecto, sin embargo Cohen (1988) propone tres valores de referencia para el tamaño de efecto $d$, para cuando no se dispone de una mejor base para la interpretación del índice lo cual nos da un marco de referencia general para este índice.

*Tamaño del efecto pequeño* ***d=0.2****.*
Este valor para $d$, visto como una diferencia pequeña de medias, en términos de las medidas de no superposición; $U_1$ representa un 14.7 % de las áreas de las poblaciones A y B no coincide. $U_2$ indica que considerando a B como la población con la media más grande y A con la media más pequeña ($\mu_B > \mu_A$), el 54 % de la población B excede el 54 % de la población A; y en términos de $U_3$ indica que el 57.9 % de la población A es superada por la media de la población B. Todo esto corresponde a poblaciones con distribución normal, misma variabilidad y mismo tamaño.

En cuanto a la $d$ en términos de correlación para poblaciones del mismo tamaño, con $d = 0.2$ se obtiene una $r = 0.1$ y una $r^2 = 0.01$ lo que indica que la diferencia de las medias de la población representan el 1 % de la variación total.

*Tamaño del efecto mediano* ***d=0.5****.*

El valor de $d = 0.5$ se considera "mediano" para una diferencia de medias, el cual en términos de las medidas de no superposición indican que, $U_1 = 33\,\%$ es decir el $33\,\%$ es el área de no superposición. Considerando a $\mu_B > \mu_A$, $U_2 = 59.9\,\%$ de la población B excede el $59.9\,\%$ de la población A y que con $U_3 = 69.1\,\%$ de la población A es superada por la media de la población B. Todo esto corresponde a poblaciones con distribución normal, misma variabilidad y mismo tamaño.

En cuanto a la correlación para poblaciones del mismo tamaño, d=0.5 da una $r = 0.243$ y un $r^2 = 0.059$ , lo que indica que el $5.9\,\%$ de la variación total es explicada por la diferencia de las medias de la población.

*Tamaño del efecto grande* ***d=0.8***.
Cuando la diferencia de medias es igual a 0.8 esta diferencia es considerada como grande y en cuanto a las medidas de no superposición esto representa que $U_1 = 47.7\,\%$ lo que indica que prácticamente la mitad de sus áreas no se superponen, un $U_2 = 65.5\,\%$, considerando a $\mu_B > \mu_A$ el $65.5\,\%$ de la población B excede el $65.5\,\%$ de la población A y un $U_3 = 78.8\,\%$ indicando que la media de la población B excede el $78.8\,\%$ de la población A.

En este valor la $r$ es igual a 0.371 y nuestra $r^2$ por tanto es igual a 0.138 , teniendo entonces un $13.76\,\%$ de la variación total explicada por la diferencia de las medias de la población y un $r$ que en algunos métodos no es considerada como lo suficientemente grande como para resultar significativa.

## 3.2. Índice $f$ para análisis de varianza

El índice $f$ se puede traducir como la generalización del índice $d$, pues a diferencia de este último, el índice $f$ busca contrastar la igualdad de $k$ medias, donde $k$ puede ser mayor o igual a dos. Cohen (1988) define el índice $f$ en términos de la suma de los cuadrados de la diferencia media de cada grupo con la media general, dividido por la desviación estándar de los $k$ grupos misma que es utilizada en el cálculo de la $d$ de Cohen, por lo que el índice es definido como.

$$f = \sigma_m/\sigma \tag{3.4}$$

Considerando muestras de igual tamaño, $\sigma_m = \sqrt{\sum_{i=1}^{k}(\mu_i - \mu)^2/k}$, con $\mu$ igual a la media de las $\mu_i$, es decir $\mu_i = \sum_{i=1}^{k} \mu_i/k$ .

Cuando todas las medias son iguales, $f$ toma el valor de cero y en caso de que al menos una media sea distinta, el índice $f$ puede tomar valores infinitamente grandes conforme $\sigma_m$ aumente respecto a $\sigma$. Cohen (1988) muestra dos formas de representar el índice $f$ los cuales se muestran a continuación.

### 3.2.1. $f$ como una medida de rango estandarizado

Una manera sencilla de representar el índice $f$, es a través de una diferencia de medias estandarizada, la cual en este caso se toma como la distancia entre la media más grande

($\mu_{max}$) y la media más pequeña ($\mu_{min}$) de las $k$ medias, lo cual se escribe como:

$$d = \frac{\mu_{max} - \mu_{min}}{\sigma} \tag{3.5}$$

Bajo el supuesto de muestras de igual tamaño, se puede observar que cuando $k = 2$ la ecuación (3.5) es igual a la ecuación (3.1) del índice $d$ para la diferencia de dos medias y como se observa en el resultado de la ecuación (3.6) la desviación estándar de dos valores, en este caso dos medias, resulta ser simplemente la mitad de su diferencia. Por lo que la relación de $f$ con $d$ es $f = d/2$, entonces $d = 2f$:

$$\begin{aligned}
\sigma_m &= \sqrt{\frac{(\mu_1 - \mu)^2 + (\mu_2 - \mu)^2}{2}} \\
&= \sqrt{\frac{(\mu_1 - \frac{\mu_1+\mu_2}{2})^2 + (\mu_2 - \frac{\mu_1+\mu_2}{2})^2}{2}} \\
&= \sqrt{\frac{(-\mu_1\mu_2 + (\frac{\mu_1+\mu_2}{2})^2) + (-\mu_2\mu_1 + (\frac{\mu_1+\mu_2}{2})^2)}{2}} \\
&= \sqrt{\frac{-2\mu_1\mu_2 + \frac{(\mu_1+\mu_2)^2}{2}}{2}} \\
&= \sqrt{\frac{-2\mu_1\mu_2 + \frac{\mu_1^2+2\mu_1\mu_2+\mu_2^2}{2}}{2}} \\
&= \sqrt{\frac{\frac{\mu_1^2-2\mu_1\mu_2+\mu_2^2}{2}}{2}} \\
&= \frac{\sqrt{\mu_1^2 - 2\mu_1\mu_2 + \mu_2^2}}{2} \\
&= \frac{\sqrt{(\mu_2 - \mu_1)^2}}{2} = \frac{\mu_2 - \mu_1}{2}
\end{aligned} \tag{3.6}$$

Entonces

$$f = \frac{\frac{\mu_2-\mu_1}{2}}{\sigma} = \frac{\mu_2 - \mu_1}{2\sigma} \Rightarrow 2f = \frac{\mu_2 - \mu_1}{\sigma} = d$$

Representar a $f$ en términos de $d$ mediante el rango estandarizado de medias para el caso de dos medias facilita su uso y comprensión; sin embargo a medida que el número de medias aumenta la relación entre su desviación estándar ($f$) y su rango ($d$) depende de cómo se dispersan las medias en su rango. Por lo que al tener $k$ medias, $d$ define a la media de cada extremo, pero la distribución de las $k - 2$ medias restantes puede ser de diversas formas a lo largo del intervalo. Cohen (1988) define tres patrones de distribución para estas medias mostrando la relación en cada caso de $f$ con $d$. Dichos patrones son:

- Patrón $1(f_1)$. Define el caso en el que la variabilidad es mínima, la cual se produce cuando en cada extremo del rango $d$ hay una media y las $k - 2$ medias restantes se concentran en el punto medio:

$$f_1 = d\sqrt{1/2k} \Rightarrow d_1 = f\sqrt{2k}$$

- Patrón 2 ($f_2$). Define una variabilidad media la cual se da cuando las $k$ medias se encuentran equitativamente espaciadas en intervalos de $d/(k-1)$ a lo largo del rango $d$.:

$$f_2 = \frac{d}{2}\sqrt{\frac{k+1}{3(k-1)}} \Rightarrow d_2 = 2f\sqrt{\frac{3(k-1)}{k+1}}$$

- Patrón 3 ($f_3$). Define la variabilidad máxima la cual se produce cuando las $k$ medias caen en los extremos del rango $d$, teniendo $k/2$ en cada extremo cuando $k$ es par y cuando $k$ es impar en un extremo de $d$ se tendrán $(k+1)/2$ medias y en el otro $(k-1)/2$:

  Para $k$ par,

$$f_3 = d/2 \Rightarrow d_3 = 2f$$

  Para $k$ impar,

$$f_3 = d\frac{\sqrt{k^2-1}}{2k} \Rightarrow d_3 = f\frac{2k}{\sqrt{k^2-1}}$$

EJEMPLO **3.1** *Obtenga $f_1$, $f_2$ y $f_3$ para cinco medias con $d$ especificado y considerando que los valores se encuentran estandarizados.*

*Entonces:*

*$f_1 = d\sqrt{1/10} = 0.3162d$, es decir, la desviación estándar es 0.31, casi un tercio del rango. La dispersión de las medias bajo el criterio de este patrón sería en -d/2 la $\mu_5$, en 0 $\mu_4, \mu_3, \mu_2$ y en d/2 se encontraría $\mu_1$.*

*$f_2 = \frac{d}{2}\sqrt{6/3(4)} = 0.3535d$, como se puede observar la desviación estándar bajo $f_2$ es mayor que en $f_1$. En cuanto a la dispersión bajo este patrón sería en intervalos de $d/4$, es decir $\mu_5$ en $-d/2$, $\mu_4$ en $-d/4$, $\mu_3$ en cero, $\mu_2$ en $d/4$ y $\mu_5$ en $d/2$.*

*Para el caso de $f_3$, como 5 es impar, entonces $f_3 = d\sqrt{24}/10 = 0.4898d$, es decir la desviación estándar es casi la mitad del rango además $f_3$ es mayor que $f_2$ y que $f_1$. En cuanto a la dispersión se tiene que en un extremo del rango se tendrán 3 medias y en el otro extremo 2 medias.*

### 3.2.2. $f$ en términos de correlación y medidas de proporción de varianza

En esta sección se describen otras dos formas para analizar $f$. En términos de correlación entre los miembros de la población y la variable dependiente, y en términos de la proporción de la varianza total de las $k$ poblaciones combinadas que es explicada por los miembros de la población.

**Eta ($\eta$) y Eta cuadrada ($\eta^2$).**

El índice $\eta$ puede definirse como una generalización de la $r$ de Pearson y $\eta^2$ por consiguiente como una generalización de $r^2$ vistas en la sección 3.1.2 en donde se tiene sólo a la población A y a la población B en la cual la pertenencia a una población queda restringido a sólo dos posibilidades; sin embargo cuando se quiere representar la pertenencia a las posibles $k$ poblaciones se usa el estadístico $\eta^2$.

Ahora partiendo de la forma de $f$ definida en la ecuación (3.4) y considerando que todas nuestras poblaciones tienen la misma varianza, pero cada una con su propia media, $m_i$. Entonces elevando al cuadrado la ecuación, se tiene

$$f^2 = \sigma_m^2/\sigma^2 \tag{3.7}$$

donde $\sigma_m^2$ será una cantidad distinta de cero cuando al menos una de las $k$ medias sea distinta de las demás.
Cohen(1988) define a $\eta^2$ como

$$\eta^2 = \sigma_m^2/\sigma_t^2 \tag{3.8}$$

Donde $\sigma_t^2$ es definida como la varianza total la cual es igual a, $\sigma_t^2 = \sigma^2 + \sigma_m^2$.
Haciendo uso de la ecuación (3.7) se tiene que

$$\begin{aligned}\eta^2 &= \frac{f^2\sigma^2}{\sigma^2 + f^2\sigma^2}\\ &= \frac{f^2\sigma^2/\sigma^2}{\sigma^2/\sigma^2 + f^2\sigma^2/\sigma^2} = \frac{f^2}{1+f^2}\end{aligned} \tag{3.9}$$

Cuando todas las medias resultan ser iguales, entonces $\sigma_m^2$ será cero y $f^2$ será cero, por lo tanto $\eta^2 = 0$ lo cual indica que ninguna proporción de la varianza total es explicada por los miembros de la población. También se puede observar de la ecuación (3.8) que cuando todos los valores dentro de cada población sean iguales se obtendrá una $\sigma^2 = 0$ por lo que $\eta^2 = 1$, es decir, la varianza total es producida por la varianza de las medias.

### 3.2.3. Valores pequeños, medianos y grandes de $f$

Cohen (1988) define tres valores específicos para considerar un tamaño del efecto como pequeño, mediano o grande, los cuales son útiles como una referencia general para el caso en el que no se cuente con experiencia suficiente en el área de estudio, como para fijar dichos valores, pues como ya se mencionó anteriormente, un valor puede considerarse insignificante en determinada área pero no necesariamente en todas.

*Tamaño del efecto pequeño* ***f=0.10***.
Un $f$ igual a 0.1 indica que la desviación estándar de las $k$ medias es 0.10 de la desviación estándar de las observaciones dentro de las poblaciones. En cuanto al rango mínimo utilizando un $f$ de 0.1 es de 0.20 con $k$ igual a 2. Independientemente del patrón de distribución utilizado pues con $k$=2 los tres patrones coinciden con la relación $d = 2f$.
Por otro lado cuando se expresa a $f$ en términos de correlación y varianza total un $f = 0.1$ es equivalente a un $\eta^2 = 0.0099$ obtenido de sustituir el valor de $f = 0.1$ en la ecuación (3.9), con esto se concluye que cerca del 1 % de la varianza total es representada por la diferencia entre los miembros de la población.

*Tamaño del efecto mediano* ***f=0.25***.
Con un $f$ igual a 0.25 se dice que la desviación estándar de las $k$ medias representan un cuarto de la desviación estándar de las observaciones dentro de las poblaciones. En cuanto al rango obtenido con un $f = 0.25$ éste comienza desde $d = 0.50$ cuando $k$ es igual a dos. Por otro lado cuando se expresa a $f$ en términos de $\eta$ y $\eta^2$ un $f = 0.25$ se tiene que $\eta^2 = 0.0588$, (varianza total) y $\eta = 0.243$ (correlación), por lo que cerca del 6 % de la varianza total es representada por la diferencia entre los miembros de la población.

*Tamaño del efecto grande* ***f=0.40***.
Un $f$ igual a 0.4 indica que la desviación estándar de las $k$ medias representan el 0.40 de la desviación estándar de las observaciones dentro de las poblaciones. Se puede notar que el rango mínimo comienza en 0.8 con $k$ igual a 2 el cual coincide con el tamaño del efecto grande establecido para $d$ mostrado en la sección (3.1.3), dicha coincidencia también aplica tanto para el efecto mediano, como para el pequeño.
En términos de $\eta$ y $\eta^2$ una $f = 0.40$ indica que $\eta^2 = 0.1379$ y $\eta = 0.371$ por lo que cerca del 10 % de la varianza total es representada por la diferencia entre los miembros de la población.

## 3.3. Índice $f^2$ para análisis de regresión

El análisis de regresión múltiple y correlación es un procedimiento de análisis de datos a través del cual se puede estudiar la relación de una variable dependiente ($Y$) con múltiples factores que afectan al fenómeno a explicar, donde cada factor es un conjunto compuesto por una o más variables independientes.

Cohen (1988) muestra una forma de ver al estadístico F para el análisis de regresión y correlación en el que se busca obtener la proporción de la varianza en $Y$ explicada, por alguna fuente (conjunto de las variables predictoras). Tal índice es escrito como,

$$f^2 = \frac{PV_f}{PV_e}$$

Donde $PV_f$ es la proporción de la varianza de $Y$ explicada por la fuente ($f$) en la muestra; $PV_e$ es la proporción del error ($e$) o varianza residual, considerando a este cociente como la medida del tamaño del efecto en la muestra, el cual indica la proporción de la variabilidad de $Y$ explicada por la fuente en proporción con el error.

Cohen (1988) menciona 3 casos que se pueden presentar para definir la proporción de varianza explicada a través del índice $f^2$.

**Caso 0**. Un único conjunto B, compuesto de $u$ variables, está relacionado con $Y$, por lo que $PV_f = R^2_{Y\cdot B}$; su complemento, $PV_e = 1 - R^2_{Y\cdot B}$, es la proporción de varianza del error. La hipótesis nula es que el valor de la población de $R^2_{Y\cdot B}$ es cero, es decir, la proporción de la varianza de $Y$ explicada por B es cero; así entonces

$$f^2 = \frac{R^2_{Y\cdot B}}{1 - R^2_{Y\cdot B}} \tag{3.10}$$

**Caso 1**. En este caso dos conjuntos (A y B) se encuentran relacionados con $Y$. Se busca determinar la proporción de la varianza de $Y$ explicada sólo por el conjunto B, en donde $R^2_{Y\cdot A,B}$ es la varianza explicada por el conjunto A y B, por lo que $PV_f = R^2_{Y\cdot A,B} - R^2_{Y\cdot A} = B\cdot A$ y $PV_e = 1 - R^2_{Y\cdot A,B}$. La hipótesis nula en este caso es que la proporción de la varianza explicada sólo por B es cero. Por lo que

$$f^2 = \frac{R^2_{Y\cdot A,B} - R^2_{Y\cdot A}}{1 - R^2_{Y\cdot A,B}} \tag{3.11}$$

**Caso 2**. Análogo al caso 1, con $PV_f = R^2_{Y\cdot A,B} - R^2_{Y\cdot A}$, pero el ruido se reduce aún más con la inclusión de un conjunto C, por lo que $PV_e = 1 - R^2_{Y\cdot A,B,C}$, haciendo que la proporción de la varianza de $Y$ explicada por B.A no sea menor que en el caso 1. Entonces

$$f^2 = \frac{R^2_{Y\cdot A,B} - R^2_{Y\cdot A}}{1 - R^2_{Y\cdot A,B,C}} \tag{3.12}$$

### 3.3.1. $R^2$ parcial y $R^2$ semiparcial

La proporción de la varianza explicada para los casos uno y dos, puede interpetarse como la proporción de varianza de $Y$ explicada por $B \cdot A$, lo cual puede ser escrito como $R^2_{Y\cdot(B\cdot A)}$, lo cual es llamado como **correlación semiparcial múltiple cuadrada** ya que el conjunto B y A comparten variabilidad explicada, es decir no son fuentes independientes de variabilidad y lo que se busca es obtener la variabilidad que aporta sólo el conjunto B quitando la variabilidad que explica junto con el conjunto A, así como la variabilidad explicada sólo por el conjunto A en otras palabras lo que el conjunto B añade al conjunto A.

Por otro lado, cuando se expresa la variabiliad de $Y$ explicada por $B \cdot A$ como una proporción de la parte de la varianza de $Y$ no explicada por el conjunto A, en lugar de como una proporción de la varianza total, como es el caso de la correlación semiparcial, el coeficiente resultante es entonces de **correlación parcial múltiple cuadrada**, cuya fórmula quedaría como

$$R^2_{YB\cdot A} = \frac{R^2_{Y\cdot A,B} - R^2_{Y\cdot A}}{1 - R^2_{Y\cdot A}} = \frac{R^2_{Y\cdot(B\cdot A)}}{1 - R^2_{Y\cdot A}}$$

El cálculo de $f^2$ al basarse en la estimación de $R^2$ y la contribución de cada conjunto a la proporción de la varianza explicada se puede obtener de la ecuación (3.10) que $R^2_{Y\cdot B} = \frac{f^2}{1+f^2}$ la cual es igual a la ecuación(3.9) dada para $\eta^2$ en la sección anterior, por lo que $\eta^2$ resulta ser un caso especial de $R^2$. Por otro lado al sustituir $R^2_{Y\cdot B}$ por $R^2_{YB\cdot A}$ se puede obtener que $R^2_{YB\cdot A} = \frac{f^2}{1+f^2}$ con lo que se puede notar qué $\eta^2$ parcial es un caso especial de $R^2$ parcial como lo es $\eta^2$ de $R^2$.

### 3.3.2. Valores pequeños, medianos y grandes de $f^2$

A continuación se muestran valores fijos para $f^2$ presentados por Cohen (1988) los cuales pueden ser tomados como referencia para la interpretación del tamaño del efecto del índice $f^2$, sólo para el caso en el que el investigador no cuente con información suficiente para asignar un propio valor de acuerdo al área y tipo de investigación en el que se está trabajando.

*Tamaño del efecto pequeño* $\boldsymbol{f^2 = 0.02}$.
Dicho valor para $f^2$ en términos de $R^2$ o $R^2$ parcial da un valor de $0.02/(1+0.02) = 0.019$ es decir un valor aproximado del 2 % de varianza explicada de $Y$ puede ser representado como un efecto pequeño.

*Tamaño del efecto mediano* $\boldsymbol{f^2 = 0.15}$.
Entonces con un $f^2 = 0.15$ en términos de $R^2$ o $R^2$ parcial se obtiene un valor de $0.15/(1+0.15) = 0.13$, es decir, un 13 % de varianza explicada de $Y$ puede ser considerado como un efecto medio.

*Tamaño del efecto grande* $\boldsymbol{f^2 = 0.35}$.
Con un valor de $f^2 = 0.35$ se obtiene en términos de $R^2$ o $R^2$ parcial un valor de $0.35/(1+0.35) = 0.26$ entonces un 26 % de varianza explicada de $Y$ puede ser considerado como un efecto grande aunque para algunos tipos investigaciones esto pudiese parecer demasiado pequeño.

# Capítulo 4

# Aplicación

En este capítulo se presenta una aplicación sobre el tamaño del efecto, en el contraste de ciertas hipótesis mediante los índices vistos en el capítulo anterior, haciendo un análisis sobre logro educativo en el área de matemáticas de estudiantes de tercer grado de secundaria en México a través de los Exámenes de la Calidad y el Logro Educativo (Excale-09) aplicados a alumnos de tercer grado de Secundaria para la evaluación de los años 2008 y 2012. Información obtenida del Instituto Nacional para la Evaluación de la Educación México (INEE).

Las pruebas (Excale-09) fueron aplicadas a una muestra representativa de los alumnos del país mediante conglomerados, los cuales fueron diseñados para obtener resultados a nivel nacional por modalidad educativa (estatal, general, técnica, telesecundaria, privada) y entidad federativa. Sin embargo debido a que no es el alcance de este trabajo para los cálculos siguientes se asume que toda la muestra fue obtenida mediante un muestreo aleatorio simple.

Las puntuaciones obtenidas sobre el logro educativo fueron sometidas a una transformación mediante un modelo de Rasch bayesiano, obteniendo una distribución a posteriori con una nueva escala con valores entre 200 y 800 puntos con media nacional 500 puntos y desviación estándar poblacional de 100 puntos. Ver el Informe sobre resultados de Excale-09, INEE (Andrés Sánchez Moguel, 2009).

En la base del (Excale-09), utilizada para esta aplicación, se presentan 5 valores plausibles para el logro educativo de cada alumno los cuales fueron generados con la distribución a posteriori, de los cuales se ha tomado la media como variable de interés.

## 4.1. Cálculo del tamaño del efecto ($d$)

Mediante el índice $d$ se analiza la diferencia de medias sobre el logro educativo en el área de matemáticas medido a través de los EXCALE-09, aplicado en los años 2008 y 2012. Con el fin de conocer el cambio que se dio en el logro educativo promedio de los alumnos de nivel secundaria al terminar la educación básica.

La hipótesis a probar fue, $H_0 : \mu_{2008} = \mu_{2012}$ vs $H_1 : \mu_{2008} \neq \mu_{2012}$, donde:
$\mu_{2008}$ : media del logro educativo en matemáticas de alumnos al término de la educación

básica en el año 2008.
$\mu_{2012}$ : media del logro educativo en matemáticas de alumnos al término de la educación básica en el año 2012.

De la prueba $t$ con un $p-value$ de 0.1928, un intervalo de confianza de $(-0.557, 2.765)$ y con un nivel de significancia de 0.05, se concluye que no existe evidencia suficiente para rechazar $H_0$ por lo que se dice que las medias son iguales, es decir, el logro educativo fue igual para el año 2008 y 2012.

En la tabla 4.1 se muestran los resultados de las medias para el año 2008 y 2012 respecto al logro educativo en matemáticas, de donde se puede observar una diferencia de 1.1039 puntos y un tamaño del efecto de 0.01, el cual muestra que hay una diferencia muy pequeña en donde las medias difieren en una centésima parte de una desviación estándar por lo que se considera como un tamaño del efecto pequeño según los valores de referencia vistos en el capítulo anterior, lo cual concuerda con los valores de referencia dados por el INEE mostrados en la tabla 4.2, en donde se puede ver que una diferencia de 1.10 puntos no es de gran relevancia respecto al logro educativo.

Otra observación importante es que la media mayor no es en 2012 como se esperaría, lo cual puede deberse a un ligero retroceso en el logro educativo de los estudiantes o también podría ser debido a que las pruebas aplicadas en 2012 fueron diferentes a las utilizadas en 2008 debido a cambios curriculares programados por la SEP, a pesar de que estas fueron elaboradas intentando mantener tanto el mismo grado de dificultad como los mismos temas.

A pesar de que la diferencia resultó ser menor para el año 2012 respecto a 2008 y aunque esta diferencia fue pequeña, los resultados obtenidos no resultan satisfactorios pues el logro educativo nacional se encuentra apenas en el nivel básico.

| Año | n | media |
|---|---|---|
| 2008 | 23967 | 501.0196 |
| 2012 | 23967 | 499.9156 |
| índice d | 0.0118 | |

Tabla 4.1: Tamaño del efecto de la diferencia de medias del logro educativo para 2008 y 2012.

Otra manera de visualizar el tamaño del efecto es mediante la equivalencia del índice $d$ con el coeficiente de correlación de Pearson $r$. La ejemplificación de éste se ha realizado para probar si existe diferencia en el logro educativo dependiendo del sexo y en su caso saber qué tan grande es la diferencia. La comparación del logro educativo promedio por sexo fue realizada para los años 2008 y 2012.

Las hipótesis a probar fueron, $H_0 : \mu_H = \mu_M$ vs $H_1 : \mu_H \neq \mu_M$, para los años 2008 y 2012, donde:

$\mu_H$ : media del logro educativo en matemáticas de hombres al término de la educación básica.

[1]Basado en Informe sobre los resultados del Excale 09-2008, INEE.

| Nivel de logro | Características |
|---|---|
| Avanzado (735-800) | Dominio muy avanzado de conocimientos, habilidades y destrezas. Aprovechamiento máximo de lo previsto en el currículo. |
| Medio (583-734) | Dominio sustancial de conocimientos, habilidades y destrezas escolares. Buen aprovechamiento de lo previsto en el currículo. |
| Básico (499-582) | Dominio imprescindible (suficiente, mínimo) de conocimientos, habilidades y destrezas necesarios para poder seguir progresando satisfactoriamente en la asignatura. |
| Por debajo de lo básico (menos de 499) | Carencias importantes en el dominio curricular de los conocimientos, limitación para poder seguir progresando satisfactoriamente en la asignatura |

Tabla 4.2: Niveles de logro educativo en el área de matemáticas.[1]

$\mu_M$ : media del logro educativo en matemáticas de mujeres al término de la educación básica.

Para el año 2008, de la prueba $t$ con un $p - value$ de 0.2097, un intervalo de confianza de $(-0.8962, 4.0837)$ y con un nivel de significancia de 0.05 se concluye que no existe evidencia suficiente para rechazar $H_0$ por lo que se dice que las medias son iguales, lo que indica que no hay diferencia en el logro educativo según el sexo.

En cuanto al año 2012 los resultados obtenidos por la prueba $t$ con un $p-value < 2.2e-16$, y un intervalo de confianza de $(8.4681, 12.7289)$ con un nivel de significancia de 0.05 se concluye que se rechaza $H_0$ por lo que las medias no son iguales, indicando qué a diferencia del año 2008, en los Excale-09 aplicados en 2012, se tiene que sí existe diferencia en el logro educativo según el sexo.

En la tabla 4.3 se encuentran los resultados obtenidos para las medias, el índice $d$ y $r$ por sexo para los años 2008 y 2012.

En cuanto al año 2008 se puede observar una diferencia de 1.59 la cual puede ser considerada como una diferencia poco significativa según los valores del logro educativo mostrados en la tabla 4.2 y con un tamaño del efecto de 0.008 el cual coincide con la conclusión anterior, pues este valor es considerado como un tamaño del efecto pequeño según los valores de referencia mostrados en el capítulo anterior. En cuanto al coeficiente de correlación de Pearson $r$ se puede concluir que el grado de correlación entre el sexo y el logro educativo muy pequeño igual a 0.008 y con un 0.006 % de variabilidad del logro educativo asociada al sexo.

Para el año 2012 la diferencia en las medias es de 10.6725 la cual es un poco más relevante en comparación con la obtenida en 2008; sin embargo, también es considerada como pequeña ya que el tamaño del efecto sigue siendo muy pequeño, en cuanto al coeficiente de correlación se dice que la correlación entre el sexo y el logro educativo es del .05 y el porcentaje de variabilidad del logro educativo asociada al sexo es del .25 %, la cual es un poco mayor a la obtenida para el año 2008.

| Año | sexo | n | media | d | r |
|---|---|---|---|---|---|
| 2008 | Hombre | 11516 | 502.1302 | 0.0118 | 0.008 |
| | Mujeres | 11516 | 500.5365 | | |
| 2012 | Hombre | 13358 | 505.2066 | 0.1201 | 0.0599 |
| | Mujeres | 13358 | 494.6080 | | |

Tabla 4.3: Tamaño del efecto de la diferencia de medias e índice $r$ del logro educativo para 2008 y 2012 por sexo.

Los resultados obtenidos para los dos años muestran una diferencia a favor de los hombres lo cual puede reflejar un mejor desempeño en comparación con las mujeres, sin embargo esta diferencia no es tan relevante pues para los dos años se considera como pequeña además de que para ambos sexos los resultados no son satisfactorios pues se ubican en los niveles de logro educativo básico y por debajo de lo básico.

Se comprobaron los supuestos de normalidad, utilizando las pruebas de Lilliefors y Anderson-Darling para la comprobación del supuesto de normalidad obteniendo un $p - value < 2.2e - 16$ para ambas pruebas y ambos años por lo que no se cumple con el supuesto de normalidad.

Debido a que los datos no cumplieron con el supuesto de normalidad, se realizó una transformación Box Cox de donde se obtuvo un valor de $\lambda = -0.22$ con un intervalo de $(-0.27, -0.16)$ para 2008 y para el año 2012 un valor de $\lambda = 0.06$ con un intervalo de $(0.01, 0.12)$, sin embargo para ambos años aún con la transformación no hubo mejora en los datos, debido a que existe una ligera falta de simetría, lo cual puede observarse en la figura 4.1, en donde para ambos años se presenta una pequeña asimetría positiva la cual podría deberse principalmente a los resultados poco satisfactorios obtenidos en el desempeño de los alumnos generando una media por debajo de lo esperado, por otro lado también se puede observar una ligera mejora en el año 2012 con una forma más simétrica, sin embargo tampoco cumplen el supuesto de normalidad.

En cuanto al supuesto de igualdad de varianzas se utilizó la prueba de Fligner-Killeen de donde se obtuvo un $p - value < 2.2e - 16$ para ambos años por lo que tampoco se cumple el supuesto.

Con base en lo anterior se decidió trabajar con los datos sin la transformación.

Las medidas de no superposición no fueron aplicadas ya que los datos no cumplen con el supuesto de igualdad de varianzas.

## 4.2. Cálculo del tamaño del efecto $(f)$

Se ha estudiado el logro educativo en el área de matemáticas para el año 2008 y 2012 con el fin de conocer si éste difiere dependiendo la modalidad educativa dentro de la que se encuentran; ya sea educación general, telesecundaria, privada, técnica, a través del índice $f$ mediante la comparación de las medias de las respectivas modalidades educativas.

La hipótesis a probar fue, $H_0 : \mu_{gen} = \mu_{telesec} = \mu_{tec} = \mu_{priv} = \mu \ vs \ H_1 : \mu_i \neq \mu, \ i =$

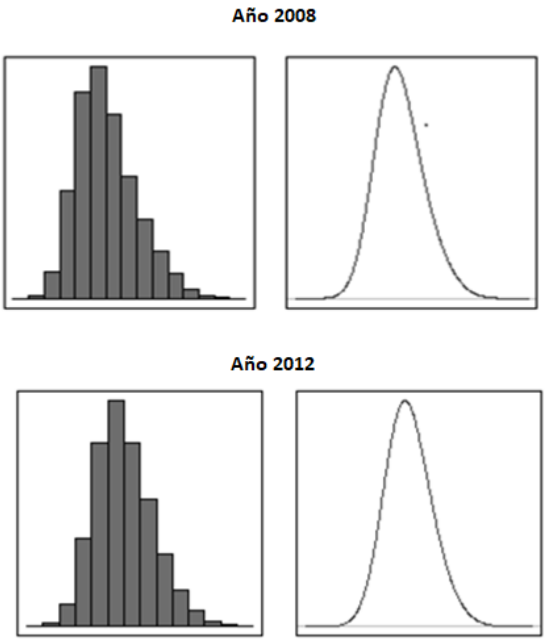


Figura 4.1: Análisis exploratorio de los datos sobre el logro educativo para los años 2008 y 2012.

*gen*, *telesec*, *tec*, *priv*, para los años 2008 y 2012, donde:

$\mu_{gen}$: media del logro educativo en el área de matemáticas de alumnos al término de la educación básica en escuelas generales.

$\mu_{telesec}$: media del logro educativo en el área de matemáticas de alumnos al término de la educación básica en telesecundarias.

$\mu_{tec}$: media del logro educativo en el área de matemáticas de alumnos al término de la educación básica en escuelas técnicas.

$\mu_{priv}$: media del logro educativo en el área de matemáticas de alumnos al término de la educación básica en escuelas privadas.

Para el año 2008 de la prueba $f$ se obtuvo un $p - value$ de $<$2e-16, con un nivel de significancia del 0.05 se concluye que existe evidencia suficiente para rechazar $H_0$ por tanto se dice que las medias son distintas, es decir, existe diferencia en el logro educativo dependiendo de la modalidad educativa. En cuanto al año 2012 se obtuvo un $p - value$ de $<$2e-16, con un nivel de signifiancia del 0.05 se concluye de igual forma así como en el año 2008 que las medias del logro educativo por modalidad educativa son distintas.

En la tabla 4.4 se encuentran los resultados obtenidos de las medias, el índice $f$ y $\eta^2$ para el logro educativo por modalidad educativa para los años 2008 y 2012. Se puede observar que tanto para 2008 como para 2012 la media más baja es representada por las telesecundarias con una media general de 477.09 y 484.48 respectivamente, lo cual indica una deficiencia importante de conocimientos ubicándose en un nivel de logro educativo por debajo de lo básico según los valores de referencia de la tabla 4.2. Mientras que la media más alta se encuentra en la modalidad privada con una media de 569.2 para el año 2008 y 570.6 para el año 2012, lo cual muestra que de manera general los alumnos en esta modalidad logran un dominio mínimo de conocimientos para progresar de manera satisfactoria en el área de matemáticas, con un nivel de logro educativo apenas básico según las referencias de la tabla 4.2.

Para los dos años de referencia la diferencia de la media más alta respecto de la media más baja es 92.11 puntos para el año 2008 y 86.18 puntos para el año 2012, lo cual indica una diferencia bastante significativa, lo cual se confirma con el tamaño del efecto obtenido mediante el índice $f$ pues con un valor de 0.42 para 2008 y 0.46 para el año 2012, se tiene un tamaño del efecto grande para los dos casos, según los valores de referencia mostrados en el capítulo anterior.

En cuanto a la correlación ($\eta$) entre la modalidad educativa y el logro educativo para el año 2008 es del 38.83 % con un 15.08 % de variabilidad del logro educativo explicado por la modalidad educativa y para 2012 la correlación es del 42.7 % con un 17.73 % de variabilidad de logro educativo explicado por la modalidad educativa.

| Año | Modalidad educativa | n | media | f | $\eta^2$ |
|---|---|---|---|---|---|
| 2008 | General | 3606 | 492.4832 | 0.4215 | 0.1508 |
| | Técnica | 3606 | 489.272 | | |
| | Telesecundaria | 3606 | 477.0906 | | |
| | Privada | 3606 | 569.2024 | | |
| 2012 | General | 3289 | 491.4081 | 0.4642 | 0.1773 |
| | Técnica | 3289 | 486.3129 | | |
| | Telesecundaria | 3289 | 484.4871 | | |
| | Privada | 3289 | 570.6073 | | |

Tabla 4.4: Tamaño del efecto y medias del logro educativo en matemáticas del año 2008 y 2012 por modalidad educativa.

De los resultados obtenidos, sobre el logro educativo en el área de matemáticas de las distintas modalidades educativas, se puede concluir que, en términos generales los resultados no son muy satisfactorios pues el mejor desempeño se encuentra en las escuelas privadas con un nivel básico, seguido de las escuelas generales y técnicas las cuales muestran resultados similares y casi en el nivel básico, encontrándose en el nivel más bajo las telesecundarias con un logro educativo por debajo de lo básico.

Estos resultados podrían deberse no sólo a la calidad de la enseñanza sino a las condiciones en las que se desenvuelven los estudiantes, pues para las escuelas telesecundarias las cuales presentan los peores resultados, éstas suelen encontrarse en poblaciones muy pequeñas con un alto grado de marginación, en donde por lo general no se cuenta con muchos servicios como internet, telefonía, computadoras, etc. A diferencia de las escuelas privadas las cuales se encuentran generalmente en zonas urbanas o con un mayor número de habitantes en donde los servicios y las condiciones socioculturales son mejores.

Al igual que para el índice $d$, el índice $f$ se basa en los supuestos de normalidad y de igualdad de varianzas y como se explicó en la sección anterior los datos no cumplieron con el supuesto de normalidad, utilizando las pruebas de Lilliefors y Anderson-Darling para normalidad obteniendo un $p-value < 2.2e-16$ y la prueba de Fligner-Killeen para el supuesto de igualdad de varianzas de donde se obtuvo un $p-value < 2.2e-16$ para ambos años, por lo que se realizó una transformación Box Cox de donde se obtuvo un valor de $\lambda = -0.22$ con un intervalo de $(-0.27, -0.16)$ para 2008 y para el año 2012 un valor de $\lambda = 0.06$ con un intervalo de $(0.01, 0.12)$, sin embargo para ambos años aún con la transformación no hubo mejora en los datos por ello se dicidió trabajar con los datos sin la transformación. La falta de normalidad puede ser explicada por una ligera asimetría positiva presente en ambos años la cual puede visualizarse en las figuras 4.2 y 4.3.

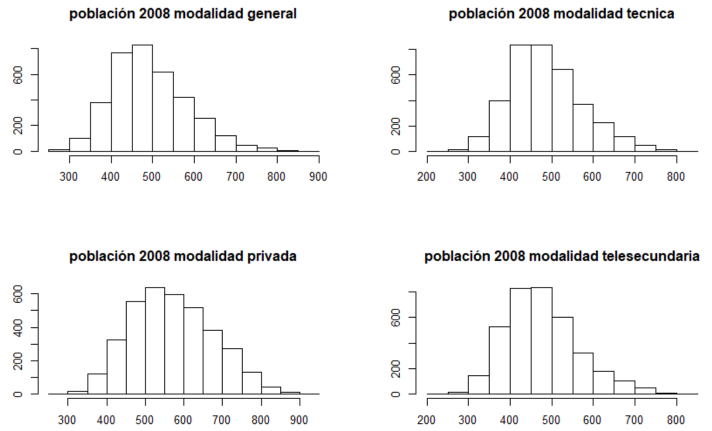


Figura 4.2: poblaciones por modalidad educativa 2008.

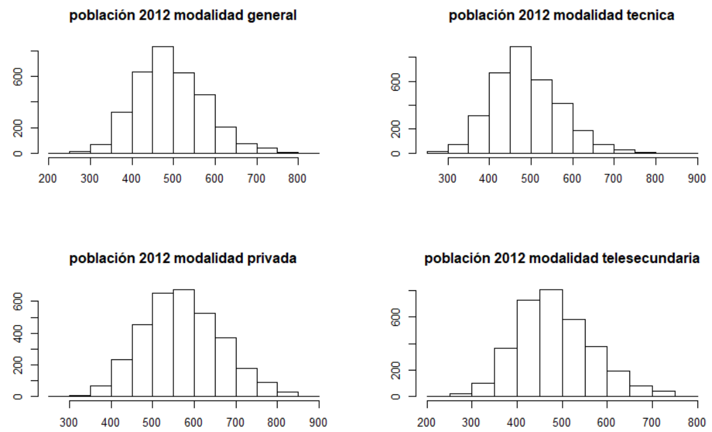


Figura 4.3: poblaciones por modalidad educativa 2012.

Para el análisis de residuales, en ambos años, utilizando las pruebas de Lilliefors y Anderson-Darling para la comprobación del supuesto de normalidad se obtuvo un $p - value < 2.2e - 16$ por lo que no se cumple dicho supuesto, sin embargo esto puede deberse a la pequeña asimetría positiva presente en los residuales de ambos años lo cual puede observarse en la figura 4.4; el supuesto de media cero se cumple, en un intervalo de $(-10.42669, 10.42669)$, con un valor de $1.650352e - 15$ para 2008 y en un intervalo de $(-9.576627, 9.576627)$, con un valor de $8.132202e - 15$ para 2012. En cuanto a la homocedasticidad esta no se cumple bajo la prueba de Levene con un $p - value < 2.2e - 16$, sin embargo en la figura 4.5 se puede observar que no existe ninguna tendencia en la distribución de los residuos lo cual indica que no existen problemas significativos de heterocedasticidad, por ultimo el supuesto de independencia de los errores se comprobo satisfactoriamente mediante la prueba de Durbin Watson con un valor de 1.8918 y 1.9383 para 2008 y 2012 respectivamente.

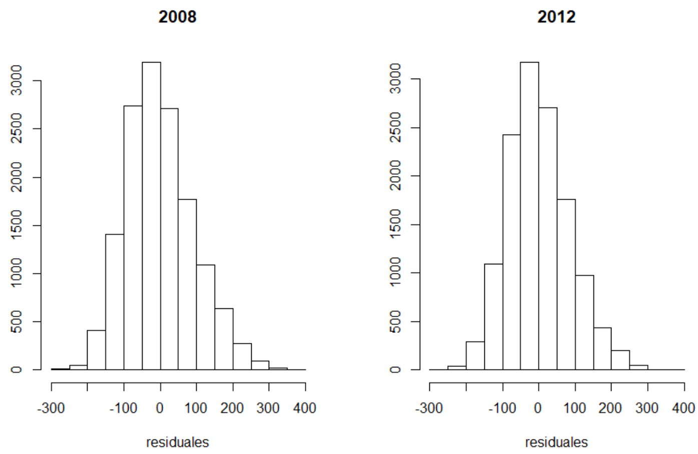


Figura 4.4: residuales

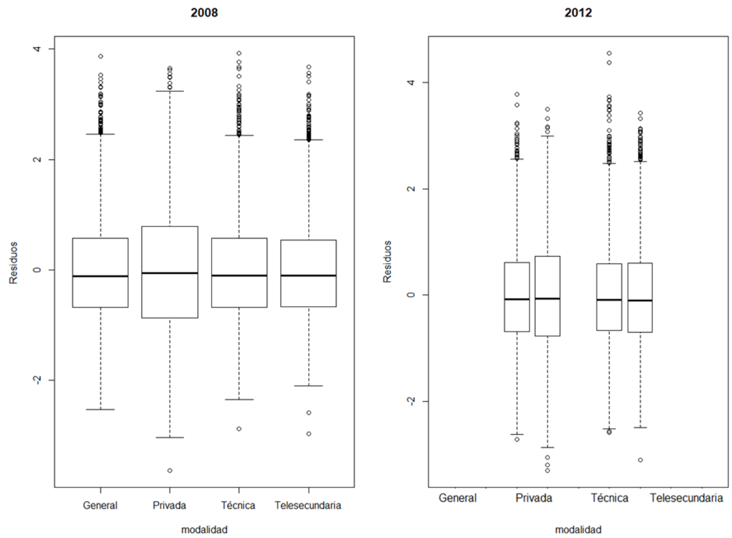


Figura 4.5: boxplot de residuos por modalidad educativa

## 4.3. Cálculo del tamaño del efecto ($f^2$)

Mediante un análisis de regresión se ha estudiado el logro educativo y su relación con el conjunto de variables predictoras; modalidad educativa, sexo, edad y entidad.

La variable modalidad educativa está conformada por cuatro categorías (general, técnica, telesecundaria, privada), en donde la categoría de referencia fue la modalidad general. La variable sexo está conformada por dos categorías (hombre, mujer), tomando como categoría de referencia a hombre. La variable entidad conformada por cada uno de los estados del país, en donde el estado de Aguascalientes fue la categoría de referencia. La variable edad conformada por dos categorías (edad normativa y extra-edad), tomando como categoría de referencia a la edad normativa. La edad normativa se refiere a la edad en la que se espera que los estudiantes terminen la secundaria la cual se encuentra entre los 14 y 15 años de edad, la extra-edad es considerada a partir de los 16 años (INEE).

El análisis fue realizado para el año 2008 y 2012, en donde la hipótesis a probar fue,

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta \ vs \ H_1 : \beta_i \neq \beta, \ i = 1, \ 2, \ 3, \ 4$, donde:

$\beta_1$ mide el cambio en el logro educativo por cada cambio en la modalidad eduucativa la cual corresponde corresponde a tres coeficientes (técnica,telesecundaria, privada), con respecto a la categoría de referencia (modalidad general) . Manteniendo el sexo, la entidad y la edad constantes.

$\beta_2$ mide el cambio en el logro educativo por cada cambio en el sexo, es decir, cambiar de hombre a mujer. Manteniendo la modalidad educativa, la entidad y la edad constantes.

$\beta_3$ mide el cambio en el logro educativo por cada cambio unitario en la edad. Manteniendo el sexo, la entidad y la modalidad educativa constantes.

$\beta_4$ mide el cambio en el logro educativo por cada cambio unitario en la entidad federativa con respecto al estado de Aguascalientes. Manteniendo constante el sexo, la edad y la modalidad educativa.

En la tabla 4.6 se encuentran los valores de los coeficientes $\beta_i$ obtenidos para el modelo en el año 2008 y 2012.

En la tabla 4.5 se encuentran los resultados obtenidos para el coeficiente de determinación y el índice $f^2$ para los años 2008 y 2012. El índice $f^2$ basado en la formula (3.10), es estimado a partir del coeficiente de determinación ($R^2$) el cual estima la proporción de varianza del logro educativo explicada por el conjunto de variables predictoras.

| Año | n | $R^2$ | $f^2$ |
|---|---|---|---|
| 2008 | 23563 | 0.144 | 0.1682 |
| 2012 | 26791 | 0.1535 | 0.1813 |

Tabla 4.5: Tamaño del efecto y coeficiente de determinación del logro educativo en matemáticas del año 2008 y 2012.

| $\beta_i$ | 2008 | 2012 |
|---|---|---|
| $\beta_0$ | 514.35 | 514.32 |
| $\beta_{mod\ priv}$ | 73.68 | 77.76 |
| $\beta_{mod\ tec}$ | -5.83 | -5.26 |
| $\beta_{mod\ telesec}$ | -13.48 | -2.42 |
| $\beta_{mujer}$ | -4.45 | -13.35 |
| $\beta_{extra\ edad}$ | -49.67 | -47.14 |
| $\beta_{BajaCalif}$ | -29.11 | -19.38 |
| $\beta_{BajaCalifS}$ | -16.08 | -17.89 |
| $\beta_{Campeche}$ | -14.90 | -23.62 |
| $\beta_{Chiapas}$ | -15.30 | -28.30 |
| $\beta_{Chihuahua}$ | -15.59 | -8.95 |
| $\beta_{Coahuila}$ | -16.90 | -19.97 |
| $\beta_{Colima}$ | -4.29 | -7.89 |
| $\beta_{DF}$ | 12.12 | 9.36 |
| $\beta_{Durango}$ | -6.14 | -8.00 |
| $\beta_{Guanajuato}$ | -2.29 | -8.29 |
| $\beta_{Guerrero}$ | -29.04 | – |
| $\beta_{Hidalgo}$ | -9.44 | -10.57 |
| $\beta_{Jalisco}$ | -11.35 | -3.95 |
| $\beta_{Mexico}$ | -0.74 | -7.96 |
| $\beta_{Michoacn}$ | -34.22 | -41.25 |
| $\beta_{Morelos}$ | -2.86 | -13.44 |
| $\beta_{Nayarit}$ | -15.45 | -8.11 |
| $\beta_{NuevoLeon}$ | -13.95 | 2.31 |
| $\beta_{Oaxaca}$ | 6.39 | – |
| $\beta_{Puebla}$ | -1.02 | -1.69 |
| $\beta_{Queretaro}$ | -0.28 | 11.30 |
| $\beta_{QRoo}$ | -6.70 | -8.36 |
| $\beta_{SnLuisP}$ | -2.87 | -12.72 |
| $\beta_{Sinaloa}$ | -2.31 | -9.79 |
| $\beta_{Sonora}$ | -7.14 | -12.71 |
| $\beta_{Tabasco}$ | -28.83 | -36.77 |
| $\beta_{Tamaulipas}$ | -19.14 | -25.86 |
| $\beta_{Tlaxcala}$ | 0.67 | -13.79 |
| $\beta_{Veracruz}$ | -10.24 | -5.30 |
| $\beta_{Yucatn}$ | -14.00 | 2.34 |
| $\beta_{Zacatecas}$ | -9.99 | -9.71 |

Tabla 4.6: Coeficientes parciales de regresión para los años 2008 y 2012.

El $p - value$ asociado al contraste tanto para el año 2008 como para el año 2012 resultó de <2.2e-16 por lo que con un nivel de significancia de 0.05 se rechaza la hipótesis nula y entonces al menos una de las variables predictoras influye en el logro educativo, por tanto el modelo en conjunto resulta significativo.

En cuanto al año 2008 se tiene que las variables predictoras del modelo explican el 14.4 % de la varianza total de la variable logro educativo el cual resulta muy bajo en cuanto a la proporción de varianza explicada, sin embargo el valor de $f^2$ indica un efecto medio de las variables predictoras en el modelo.

Para el año 2012 las variables predictoras explican un 15.35 % de variabilidad del logro educativo el cual al igual que en el año 2008 resulta bajo. El índice $f^2$ de 0.18, igual representa un efecto medio de las variables predictoras en el modelo.

En la tabla 4.7 se encuentran los resultados para el índice $f^2$ con base en la formula 3.11 para obtener la correlación semiparcial múltiple dividiendo al conjunto de predictores en dos conjuntos (A y B) en donde el conjunto B está formado por cada variable predictora y en el conjunto A el resto de las variables predictoras con el fin de obtener la proporción de la varianza explicada por de cada variable. Se observa que la variable sexo es la que menos aporta sobre lo aportado por el conjunto A, mientras que la modalidad educativa es la variable que más aporta explicación sobre el conjunto A.

| | 2008 | | 2012 | |
|---|---|---|---|---|
| conjunto B | $R^2_{Y \cdot A}$ | $f^2$ para $B \cdot A$ | $R^2_{Y \cdot A}$ | $f^2$ para $B \cdot A$ |
| Modalidad educativa | 0.0615 | 0.0963 | 0.0612 | 0.1090 |
| Sexo | 0.1434 | 0.0007 | 0.1469 | 0.007 |
| Edad | 0.108 | 0.0420 | 0.1229 | 0.0361 |
| Entidad | 0.1328 | 0.0130 | 0.1229 | 0.0181 |

Tabla 4.7: Correlación semiparcial múltiple cuadrada para los años 2008 y 2012.

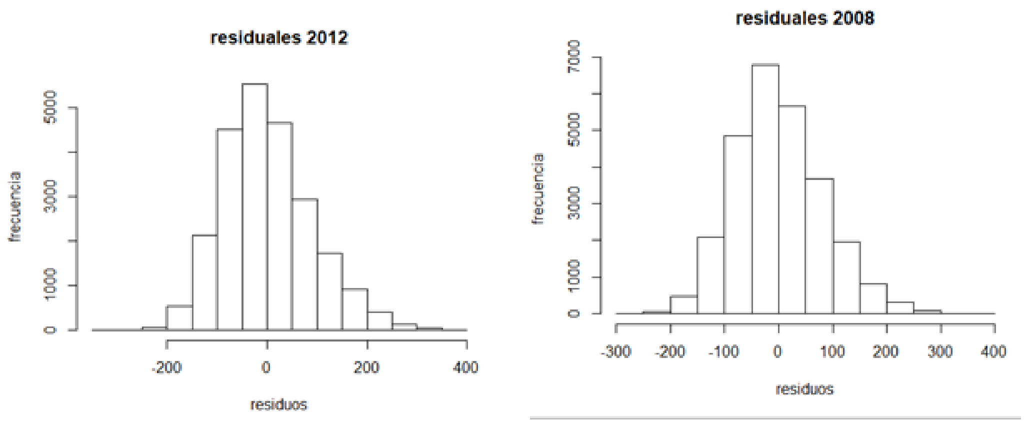


Figura 4.6: residuales para los años 2008 y 2012.

Para la verificación de los supuestos de los residuos se comprobó su independencia mediante la prueba Durbin Watson con un valor de 1.6 y 1.7 para 2008 y 2012 respectivamente. Para el supuesto de homocedasticidad este no fue superado mediante la prueba de Breusch-Pagan con un $p - value < 2.2e - 16$, sin embargo, se puede observar en la figura 4.7, que el error es constante a lo largo de las observaciones ya que no se observa ninguna tendencia por lo que se puede considerar que no existen problemas significativos de heterocedasticidad.

La media de los residuos es cero en un intervalo de $(-9.112671, 9.112671)$ con un valor de $2.709835e-15$ para 2008 y $-7.76816e-16$ para 2012. Para el supuesto de no multicolinealidad este se comprobó mediante el análisis de inflación de varianza con valores obtenidos entre 1.32 y 1.74 para cada una de las variables en ambos años, por lo que no se encuentran problemas de multicolinealidad.

La falta de normalidad en los residuos comprobada mediante la prueba de Lilliefors y Anderson-Darling con un $pvalue < 2.2e16$ puede explicarse por la presencia de una ligera asimetría, la cual puede observarse en la figura 4.6.

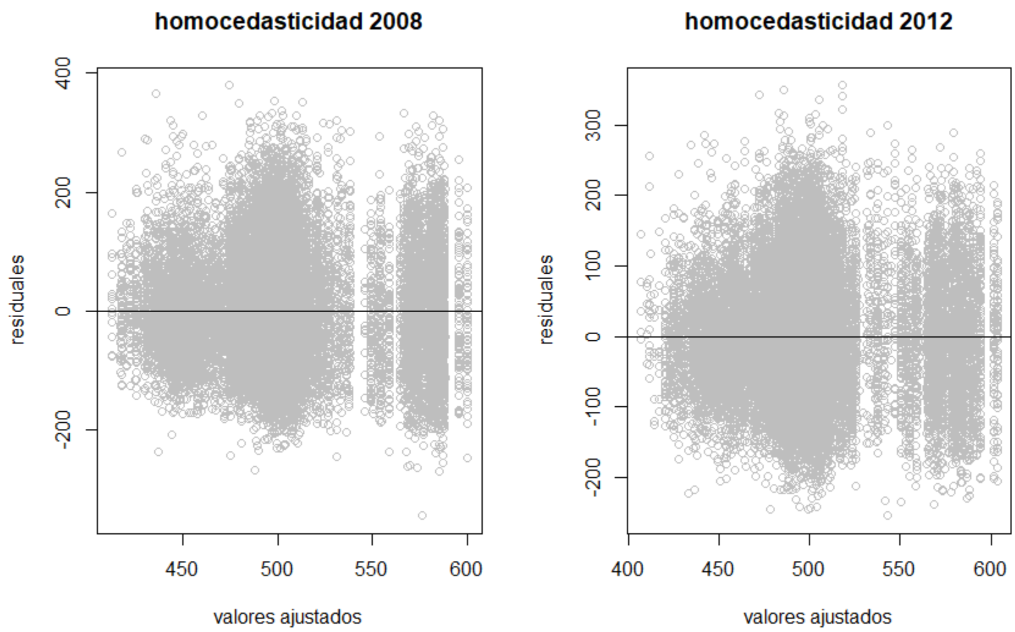


Figura 4.7: residuales para los años 2008 y 2012.

# Conclusión

Con lo expuesto anteriormente se puede concluir que el tamaño del efecto resulta de gran utilidad para mostrar la importancia de un fenómeno presente en la población de estudio, logrando disminuir las deficiencias que suelen presentarse en las investigaciones.

Como se vio en el capítulo 4 el análisis del tamaño del efecto permite generar conclusiones más completas sobre el logro educativo del área de matemáticas en el nivel básico, ya que además de la información proporcionada por las pruebas de hipótesis, se logró identificar mediante el índice $d$, que la diferencia obtenida en el logro educativo resultaba poco relevante en cuanto a una mejora en el de desempeño académico en la comparación de los resultados obtenidos para el año 2008 respecto al 2012 y en la comparación de los hombres contra las mujeres. En cuanto al resultado obtenido en el índice $f$ muestra que la modalidad educativa a la que pertenezcan los alumnos tiene un efecto grande o bastante significativo en el logro académico obtenido para ambos años. Por último, el valor obtenido para el índice $f^2$ muestra un efecto medio o medianamente significativo de las variables predictoras en conjunto, elegidas para predecir el logro académico para los años 2008 y 2012.

Lo anterior permite concluir además, que a través del análisis del tamaño del efecto se pueden generar conclusiones que aporten más información útil a la investigación ayudando a conocer la importancia que tienen los resultados obtenidos. También se pudo mostrar que el tamaño del efecto siempre dependerá del contexto en el que se encuentre, por lo que para saber si un efecto es relevante o no será necesario conocer el fenómeno de estudio, ya que aunque existan los valores de referencia estos no siempre pueden concordar con el contexto y el área de estudio, pudiendo sesgar en cierta medida las conclusiones generadas.

# Bibliografía


Andrés Sánchez Moguel, E. A. M. (2009). El aprendizaje en tercero de secundaria en México; español, matemáticas, biología y formación cívica y ética. Technical report, Instituto Nacional para la Evaluación de la Educación México (INEE).

Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury Press, USA, second edition.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum, USA, second edition.

F. Hair, J., C. Black, W., Babin, B., E. Anderson, R., and Tatham (2006). *Multivariate Data Analysis*. Pearson, sixth edition.

Glass, G., McGaw, B., and Smith, M. (1981). Meta-analysis in social research. *Sage Publications*.

Grissom, R. and Kim, J. (2005). *Effect Sizes for Research: A Broad Practical Approach*. Lawrence Erlbaum, Mahwah.

Grissom, R. and Kim, J. (2012). *Effect sizes for research: Univariate and multivariate applications*. Routledge, New York, second edition.

Kelley, K. and Preacher, K. (2012). 0n effect size. *Psychological Methods*, Vol. 17:137–152.

Vacha-Haase, T. and Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, 51:473–481.