PROYECTO DE GRADO

# *DESIGUALDAD EN EL BILINGÜISMO.*

Por:

**Gerardo Esteban Chaparro.**
**Orlando Andrés Moreno.**

Presentado a:

**Paola Casasbuenas.**

**Julio Cesar Alonso.**

UNIVERSIDAD ICESI
FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONOMICAS
PROGRAMA DE ECONOMIA Y NEGOCIOS INTERNACIONALES
SANTIAGO DE CALI, 2013

## Medidas de Desigualdad.

*RESUMEN*

La desigualdad ha sido interés de estudio por investigadores económicos a lo largo de la historia. Los diferentes autores plantean reflexiones sobre este tipo de asuntos, y además, ayudan a calcular y a explicar el porqué de esta desigualdad. Por lo tanto, los indicadores miden el nivel de desigualdad que existe en una distribución de ingresos, para así, encontrar índices útiles acerca del bienestar de una población, que pueden ser de gran importancia, ya que pueden servir como pautas para diseñar nuevas políticas económicas. Este documento muestra de manera sencilla 9 de los indicadores más empleados en estudios económicos de desigualdad. Posteriormente, calculamos los niveles de desigualdad en el bilingüismo, donde aplicamos los 9 indicadores vistos anteriormente en los resultados del puntaje de inglés de las pruebas SaberPro del año 2011 de 33 departamentos del país, para un total de 249.244 datos. Vemos una tendencia en los resultados de los cálculos donde Bogotá, San Andrés y Antioquia, se repiten en la mayoría de indicadores como los departamentos con las mayores diferencias en los puntajes de inglés obtenidos en las pruebas SaberPro.

**Palabras clave:** *Desigualdad, indicadores, distribución, cualidades deseables, desviación, bilingüismo, transferencia*

***INTRODUCCION***

La desigualdad ha sido interés de estudio por investigadores económicos a lo largo de la historia. En 1754 el filósofo francés Jean Jacques Rousseau escribió su *Discurso sobre el origen y los fundamentos de la desigualdad entre los hombres y otros escritos* (1755), donde afirma que la desigualdad social no es un estado que se da naturalmente, sino que ésta es causada por la propiedad privada y por los abusos de quienes se apropian de la riqueza. También, autores como Amartya Sen en su libro *On Economic Inequality* (1973), plantea reflexiones sobre este tipo de asuntos, y además, ayuda a calcular y a explicar el porqué de esta desigualdad. Por lo tanto, los indicadores miden el nivel de desigualdad que existe en una distribución de ingresos, para así, encontrar índices útiles acerca del bienestar de una población, que pueden ser de gran importancia, ya que pueden servir como pautas, para diseñar nuevas políticas económicas.

Este documento muestra de manera sencilla 9 de los indicadores más empleados en estudios económicos de desigualdad.
Comenzamos con una descripción de cada uno de los indicadores (El Alcance, Desviación Media Relativa, La Varianza, Coeficiente de Variación, Desviación Estándar de los Logaritmos, Curva de Lorenz, Coeficiente de Gini, Entropía de Theil, y Medida de Atkinson) y la manera de cómo se calculan. Además, hablaremos sobre las ventajas y desventajas del uso de cada una de éstas medidas.
También aplicamos estos indicadores en cuatro escenarios hipotéticos que explicamos más adelante, para calcular niveles de desigualdad en diferentes tipos de distribuciones. Y finalmente, con los resultados de Ingles de las pruebas SaberPro 2011, calculamos en 33 departamentos de Colombia la desigualdad en el bilingüismo.

Tenemos que tener en cuenta que existen ciertas características que los indicadores deben contener. Foster (1985) propone cuatro cualidades deseables en las medidas de desigualdad, por lo tanto es importante que cada indicador cumpla con lo siguiente:

1. ***Principio de transferencia de Pigou-Dalton*:** Planteado por Dalton (1920), basado en Pigou (1912). Si existe una transferencia del más rico al más pobre, la medida de desigualdad debería disminuir.

2. ***Independencia de la escala de medición del ingreso*****:** Si existe un cambio en la misma proporción de los ingresos de todos los individuos, la medida de desigualdad no debería cambiar. (Foster, 1985)
3. ***Principio de la Población de Dalton*****:** Si existe una adición proporcional de individuos a todos los niveles de ingresos, la medida de desigualdad no debería cambiar. (Dalton, 1920)
4. ***Simetría o anonimato*****:** Si existe un intercambio de ingresos entre dos individuos, la medida de desigualdad no debería cambiar.(Foster, 1985)

Para analizar estos indicadores en cuatro escenarios hipotéticos (Distribución de cola izquierda, de cola derecha, distribución normal y por último una distribución uniforme), calculamos sus niveles de desigualdad con los diferentes indicadores, y graficamos en histogramas utilizando el software estadístico R.

Para esto, definiremos la siguiente notación:

- representa el ingreso de cada individuo donde es el número de individuos en la población.
- representa la media muestral de la distribución del ingreso.
- representa la desviación estándar de la distribución del ingreso.

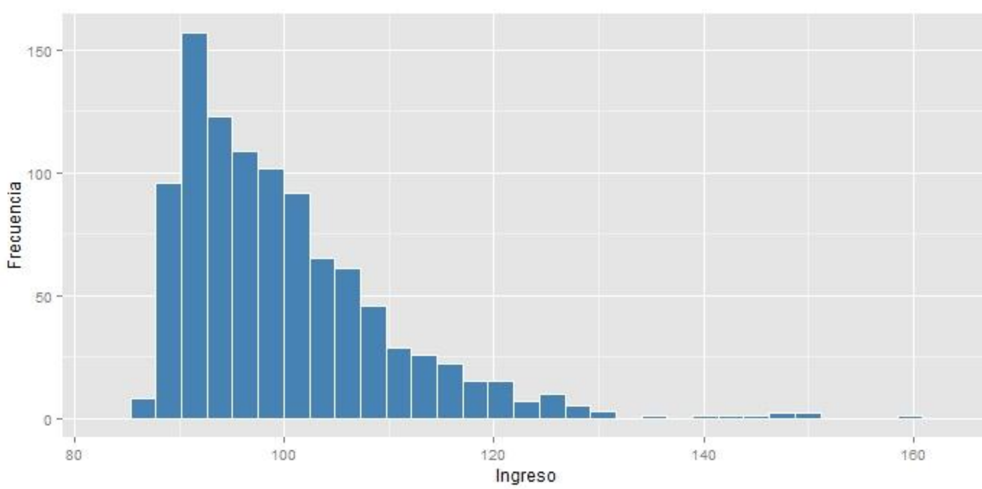
El primer escenario es una distribución de cola derecha donde una gran proporción de individuos obtienen ingresos bajos. Los datos fueron construidos a partir de la función `rchisq` en R como lo muestra la Ilustración 1:

**Ilustración 1: Datos distribución de cola Derecha**

```
Console ~/
> library(ggplot2) # Ejecuta el paquete ggplot para crear gráficos.
> set.seed(12345) # Fija valores aleatorios, que se pueden reproducir despues.
> i<-rchisq(1000,3) # Genera una distribución chi-cuadrado con mil datos aleatorios
y 3 grados de libertad.
> i<-as.data.frame(i) # Cambia el formato de i a "data frame"
> D<-((((i-colMeans(i))/sapply(i,sd))+10)*10) # Estandarización de los datos.
> D<-as.data.frame(D) # Cambia el formato de D a "data frame"
> colMeans(D) # Calcula la media muestral.
  i
100
> sapply(D,sd) # Calcula la desviación estandar.
 i
10
> max(D) # Calcula el valor máximo.
[1] 160.751
> min(D) # Calcula el valor mínimo.
[1] 87.6093
>
```

El ingreso mínimo es de 87.6093 y el máximo de 160.751 unidades monetarias. Tamaño muestral de 1000, media muestral =100, y desviación estándar 10.

**Gráfico 1: Primer Escenario - Cola Derecha.**



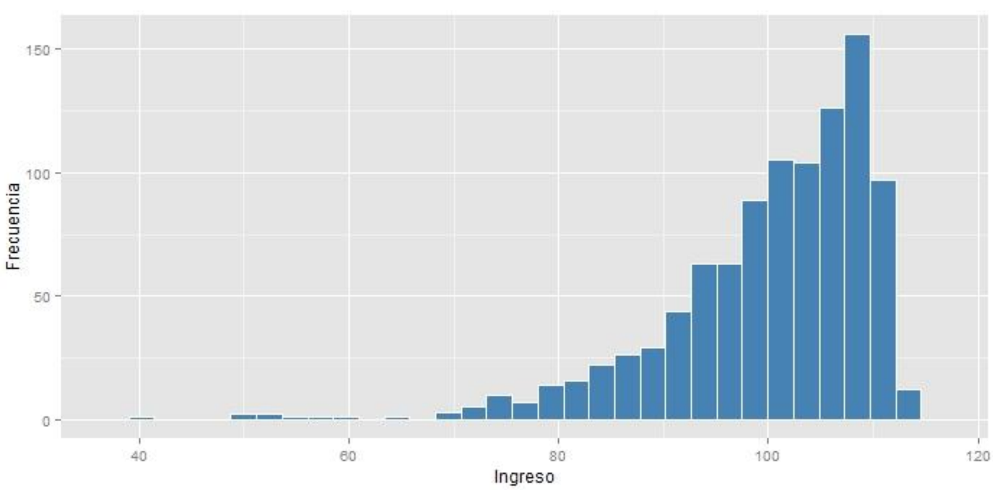

El segundo escenario hipotético es una distribución de cola izquierda, donde una gran proporción de individuos obtiene ingresos altos. Los datos fueron construidos mediante la función `rchisq` en R, lo cual genera una muestra aleatoria con una distribución chi-cuadrado. Además, se manipularon algebraicamente los datos de la distribución de cola derecha, para que la cola quedara al lado izquierdo. En la Ilustración 2 vemos como se crearon los datos:

**Ilustración 2: Datos distribución cola Izquierda.**

```
Console ~/
> library(ggplot2) # Ejecuta el paquete ggplot para crear gráficos.
> set.seed(12345) # Fija valores aleatorios, que se pueden reproducir despues.
> i<-rchisq(1000,3) # Genera una distribución chi-cuadrado con mil datos aleatorios
y 3 grados de libertad.
> d<-(((i-10))*(-1)) # Se resta 10 a cada dato i, y se multiplica por -1, para
cambiar el sentido de la cola.
> d<-as.data.frame(d) # Cambia el formato de d a "data frame"
> I<-((((d-colMeans(d))/sapply(d,sd))+10)*10) # Estandarización de los datos.
> I<-as.data.frame(I) # Cambia el formato de I a "data frame"
> colMeans(I) # Calcula la media muestral.
  d
100
> sapply(I,sd) # Calcula la desviación estandar.
 d
10
> max(I) # Calcula el valor maximo.
[1] 112.3907
> min(I) #Calcula el valor minimo.
[1] 39.24898
> 
```

El ingreso mínimo es de 39.24898, y el ingreso máximo es 112.3907 unidades monetarias. Se utilizó una muestra de tamaño 1000, con una media muestral =100, y desviación estándar 10.

**Gráfico 2: Segundo Escenario – Cola Izquierda.**



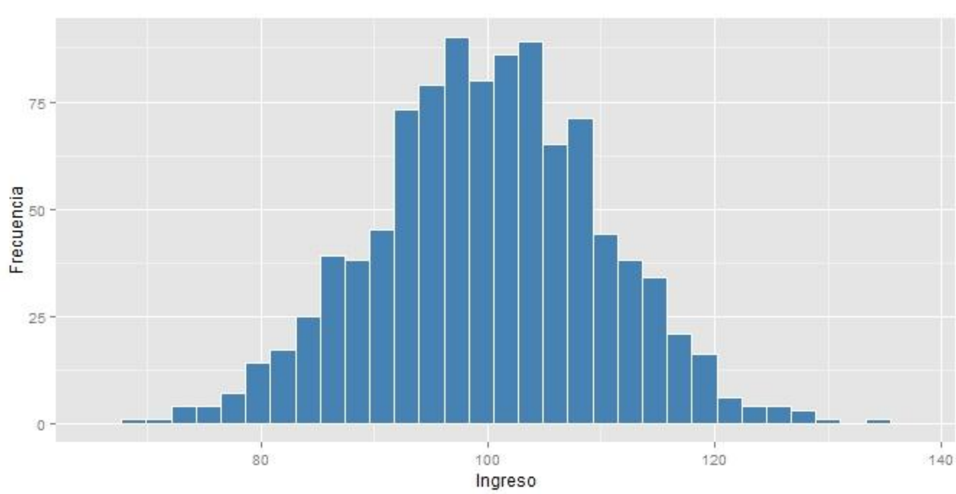
El tercer escenario es una distribución normal, donde el ingreso de la mayoría de los individuos se acerca al ingreso medio de la muestra. La función utilizada en R fue `rnorm`, que genera los datos aleatoriamente con una distribución normal, como lo muestra la Ilustración 3:

**Ilustración 3: Datos Distribución Normal.**

```
Console ~/
> library(ggplot2) # Ejecuta el paquete ggplot para crear gráficos.
> set.seed(12342) # Fija valores aleatorios, que se pueden reproducir despues.
> n<-rnorm(1000) # Genera una distribución normal con mil datos aleatorios.
> n<-as.data.frame(n) # Cambia el formato de n a "data frame"
> N<-((((n-colMeans(n))/sapply(n,sd))+10)*10) # Estandarización de los datos.
> N<-as.data.frame(N) # Cambia el formato de N a "data frame"
> colMeans(N) # Calcula la media muestral.
  n
100
> sapply(N,sd) # Calcula la desviación estandar.
 n
10
> max(N) # Calcula el Valor maximo.
[1] 134.87
> min(N)# Calcula el valor minimo.
[1] 69.31299
>
```

El ingreso mínimo es 69.31299, y el máximo es 134.87, unidades monetarias. El tamaño de la muestra es 1000, media muestral =100, y desviación estándar 10.

**Gráfico 3: Tercer Escenario - Distribución Normal**



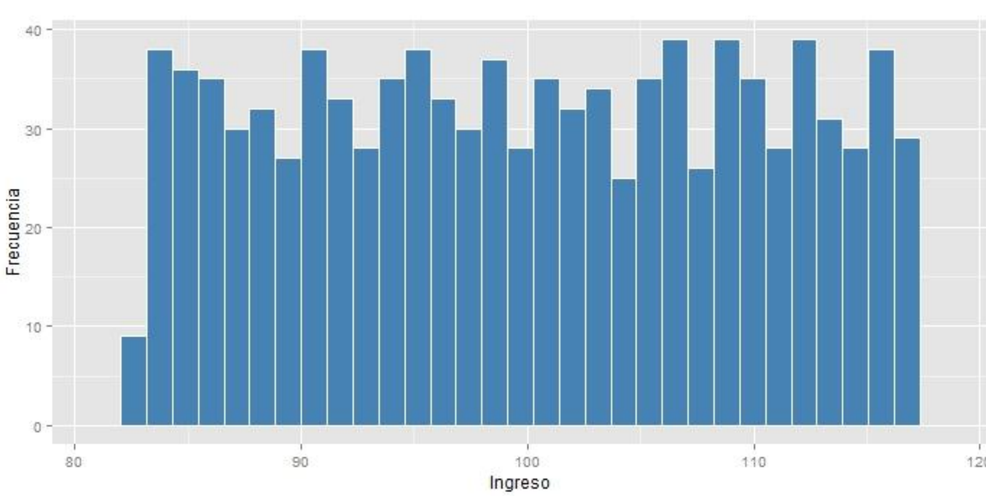
Por último, en el cuarto escenario, tenemos una distribución uniforme, donde existen cantidades similares de individuos con ingresos bajos, medios y altos. Los datos fueron generados aleatoriamente mediante la función runif en R, como se muestra en la Ilustración 4:

**Ilustración 4: Datos Distribución Uniforme.**

```
Console ~/
> library(ggplot2) # Ejecuta el paquete ggplot para crear gráficos.
> set.seed(12342) # Fija valores aleatorios, que se pueden reproducir despues.
> u<-runif(1000) # Genera una distribución uniforme con mil datos aleatorios.
> u<-as.data.frame(u) # Cambia el formato de u a "data frame"
> U<-((((u-colMeans(u))/sapply(u,sd))+10)*10) # Estandarización de los datos.
> U<-as.data.frame(U) # Cambia el formato de U a "data frame"
> colMeans(U)# Calcula la media muestral.
  u
100
> sapply(U,sd)# Calcula la desviación estandar.
 u
10
> max(U) # Calcula el valor maximo.
[1] 117.1248
> min(U) #Calcula el valor minimo.
[1] 82.94492
>
```

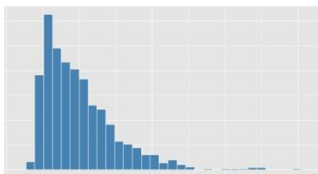
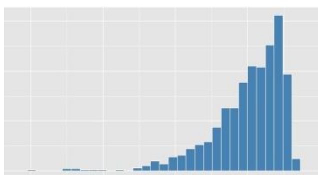
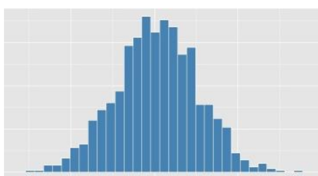
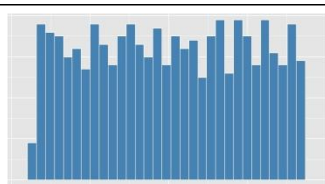
El ingreso mínimo es de 82.94492, y el máximo es de 117.1248, unidades monetarias. Tamaño de muestra 1000, media muestral = 100, y desviación estándar 10.

**Gráfico 4: Cuarto Escenario - Distribución Uniforme.**



En la Tabla 1 se resumen las estadísticas descriptivas de los 4 escenarios:

**Tabla 1: Resumen estadísticas descriptivas de los escenarios.**

| **Características.** | **Distribución.** | **Tamaño ( )** | **Media muestral ( )** | **Desviación estándar ( )** | **Valor mínimo** | **Valor máximo** |
|---|---|---|---|---|---|---|
| Concentración en ingresos bajos. |  | 1000 | 100 | 10 | 87.6093 | 160.751 |
| Concentración en ingresos altos. |  | 1000 | 100 | 10 | 39.24898 | 112.3907 |
| Concentración en ingresos medios. |  | 1000 | 100 | 10 | 69.31299 | 134.87 |
| No hay concentración. |  | 1000 | 100 | 10 | 82.94492 | 117.1248 |

Nuestro propósito del trabajo no es decir cual distribución es mejor que otra, ya que como lo dice S. Kuznetz (1950) cuando hablamos de desigualdad de renta, solo nos referimos a la diferencia de ésta, sin tener en cuenta su deseabilidad. Por lo tanto, estos indicadores no miden lo “adecuado” de la distribución, sino cuan cerca o lejos se encuentra de la situación en donde todos los individuos obtienen una renta idéntica, sin que esto signifique un fin en especial. (Núñez Velázquez, 2006)
Ahora, con los 4 escenarios y las características deseables, procederemos a explicar los 9 indicadores.

***INDICADORES***

1. ***EL ALCANCE.***

Es una comparación entre el valor máximo y mínimo de la distribución de ingresos. El alcance se puede definir como la diferencia de estos valores extremos como proporción del ingreso medio.(Sen, 1973). Por lo tanto, se formula de la siguiente manera:

De acuerdo a las cualidades que se esperan de las medidas de desigualdad, el alcance no cumple con el *Principio de transferencia de Pigou-Dalton* ya que una transferencia de un rico a un pobre no modificaría la medida de desigualdad, ya que solo se toman en cuenta los valores extremos de la distribución del ingreso.

Por otro lado, este indicador sí cumple la cualidad sobre independencia de escala. Si los ingresos de todos los individuos cambian en una misma proporción , entonces:

Por lo tanto, la medida de desigualdad no cambia.

Además, el alcance cumple con el principio de población. Por ejemplo, dada una distribución de ingresos . Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es . Como se puede verificar fácilmente, la media y los valores máximos de la distribución de ingresos no cambian. Por lo tanto, la medida de desigualdad es la misma.

También, esta medida de desigualdad posee la cualidad de simetría. Si se efectúa una reordenación de los ingresos de los individuos de una muestra, el valor máximo y mínimo , y la media de los ingresos no se alterarán. Por lo tanto, el resultado de la medida seguirá siendo el mismo. Es decir, se cumple que la medida no depende de características diferentes a los ingresos de los individuos y un intercambio entre los individuos no variará el nivel de desigualdad.

Los resultados de las cualidades esperadas del Alcance se resumen en la Tabla 2:

**Tabla 2: Cualidades esperadas del Alcance.**

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | X |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

La ventaja de este indicador es que es sencillo de calcular, ya que solo se requiere la información sobre los valores extremos de la distribución de los ingresos, y la media. Además permite con facilidad su interpretación. Por ejemplo:

Si se divide el ingreso total en partes iguales entonces:

Por lo tanto, , que corresponde a una distribución donde cada individuo pose el mismo ingreso .

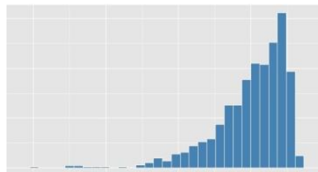
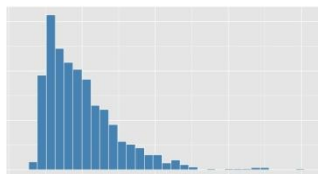
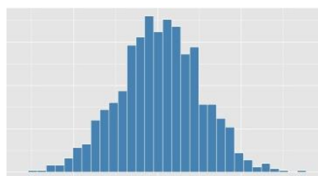
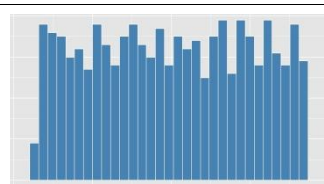
Por otra parte, si una persona recibe todo el ingreso entonces:

En este orden de ideas, El alcance se encuentre entre 0 y , cuando existe una distribución igual del ingreso total, y cuando el ingreso total pertenece a un solo individuo.

La desventaja de emplear El alcance como medida de desigualdad, es que omite la distribución entre los extremos, es decir, no toma en cuenta los valores intermedios de la muestra. Como solo destaca los valores extremos, se puede llegar a conclusiones equivocadas de los niveles de desigualdad. (Sen, 1973).

A continuación, calculamos en R El Alcance según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 3:

**Tabla 3: Resultados en R del Alcance.**

| **Características.** | **Distribución.** | **Alcance ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 0.7314171 |
| Concentración en ingresos bajos. |  | 0.7314171 |
| Concentración en ingresos medios. |  | 0.6555699 |
| No hay concentración. |  | 0.3417985 |

Con los resultados, vemos que tanto la concentración en ingresos altos, como en ingresos bajos, se obtiene la misma medida de desigualdad. Utilizando el Alcance, el escenario que se encuentra más cerca a una situación donde todos los individuos obtienen una renta idéntica, es el de la distribución uniforme, donde no hay concentración.

## 2. DESVIACION MEDIA RELATIVA.

Esta medida compara el nivel de ingresos de cada individuo con respecto al ingreso medio de la distribución. Una ventaja de la desviación media relativa es que agrega estas distancias individuales al sumar los valores absolutos de todas las diferencias, es decir no solo toma en cuenta los valores extremos (como el Alcance) sino que también lo que ocurre entre estos valores. Se formula de la siguiente manera:

Una desventaja de este indicador es que no es sensible a las transferencias entre pobres y ricos mientras ambos individuos se encuentren en el mismo lado del ingreso medio. Si suponemos que se transfiere una cantidad del ingreso de una persona más rica a una más pobre:

Lo que ocurre con una transferencia de una persona (que tiene un ingreso menor a la media), a una más pobre, es que aumenta una brecha, pero disminuye otra en la misma cantidad. Esto quiere decir que el valor de permanece igual, y por lo tanto se viola el *Principio de transferencia de Pigou-Dalton*.(Sen & Foster, 1997).

El indicador sí cumple con la cualidad de la independencia de la escala de medición del ingreso. Si los ingresos de todos los individuos cambian en una misma proporción , entonces:

Ante el aumento en la misma cantidad de los ingresos, la desviación media relativa no cambia.

También se cumple con el principio de población. Dada una distribución de ingresos . Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es

Para

Para

Por lo tanto, la desviación media relativa no cambia.

Además, tiene la cualidad de simetría, como se puede observar en la fórmula, el indicador solo depende de las características del ingreso de los individuos.

Los resultados de las cualidades esperadas de la Desviación Media Relativa se resumen en la Tabla 4:

**Tabla 4: Cualidades esperadas de la Desviación Media Relativa.**

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | X |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

Para su interpretación, analicemos que ocurre en los casos extremos. Primero, si el ingreso total se divide en parte iguales:

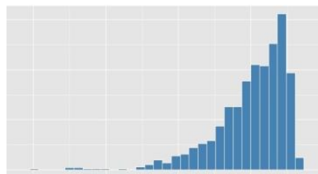
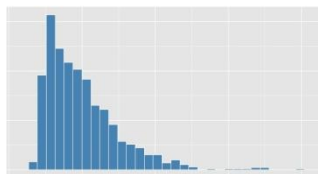
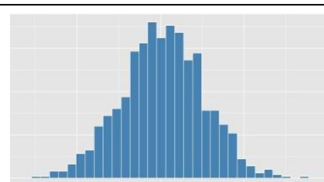
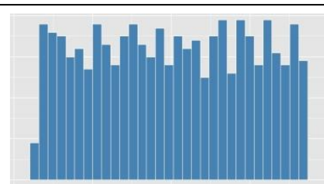
Por lo tanto, , que corresponde a una distribucion donde cada individuo pose el mismo ingreso .

Por otra parte, si una persona recibe todo el ingreso, llamemos a esa persona :

En este orden de ideas, la Desviación media relativa se encuentra entre 0 y .

A continuación, calculamos en R La Desviación Media Relativa según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 5:

**Tabla 5: Resultados en R de la Desviación Media Relativa.**

| **Características.** | **Distribución.** | **Desviación Media Relativa ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 0.07585467 |
| Concentración en ingresos bajos. |  | 0.07585467 |
| Concentración en ingresos medios. |  | 0.07944809 |
| No hay concentración. |  | 0.08666019 |

Utilizando la Desviación Media Relativa, los escenarios que se encuentran más cerca a una situación donde todos los individuos tienen una renta idéntica, son la distribución de cola derecha e izquierda.

## 3. *LA VARIANZA.*

Se define como una medida de dispersión de una variable aleatoria respecto a su media, o valor central.

Como medida de desigualdad, resulta de la elevación al cuadrado de las desviaciones con respecto al nivel medio de los ingresos. Una ventaja es que no solo evita el problema del signo de las diferencias, sino que al elevar al cuadrado cada brecha acentúa sobre las diferencias más lejanas (Cortés & Rubalcava, 1982).

Para este indicador, cualquier transferencia de una persona más pobre a una más rica aumenta el valor de la varianza. Por lo tanto, cumple con la condición de Pigou-Dalton.

Una desventaja es que no cumple con el principio de independencia de escala ya que si se aumenta el ingreso de todos los individuos en la misma proporción el valor de la varianza aumenta.

Es evidente que el numerador de la fórmula de la varianza se hace más grande cuando aumenta el ingreso de todos los individuos.

Por otro lado, si se unen dos poblaciones con igual distribución, la varianza no cambia (Principio de población). Dada una distribución de ingresos . Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es

Para

Para

Además, si se realiza una reordenación de los ingresos de los individuos en una muestra es claro que no afecta el valor de la varianza, ya que solo depende de cada observación, del ingreso medio, y del número de observaciones.
Como se puede observar en la formula, el indicador solo depende de las características del ingreso, por lo tanto cumple el principio de simetría.

Los resultados de las cualidades esperadas de La Varianza se resumen en la Tabla 1:

**Tabla 6: Cualidades esperadas de La Varianza.**

| **Cualidades esperadas** | **¿Se cumple?** |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | ✓ |
| • *Independencia de la escala de medición del ingreso* | X |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

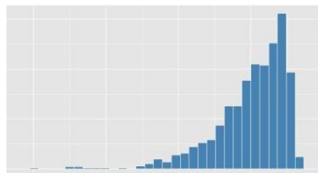
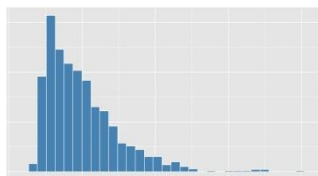
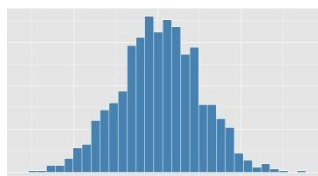
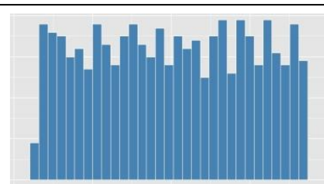
Para su interpretación, analicemos que ocurre en los casos extremos. Primero, si el ingreso total se divide en parte iguales:

Por lo tanto, , que corresponde a una distribución donde cada individuo obtiene el mismo ingreso .
Cuando el ingreso total lo obtiene un solo individuo, el valor de la Varianza es

.

A continuación, calculamos en R La Varianza según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 7:

**Tabla 7: Resultados en R de la Varianza**

| **Características.** | **Distribución.** | **Varianza ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 99.9 |
| Concentración en ingresos bajos. |  | 99.9 |
| Concentración en ingresos medios. |  | 99.9 |
| No hay concentración. |  | 99.9 |

Vemos en los resultados, que como las 4 distribuciones poseen la misma media muestral, y la misma desviación estándar, todas se encuentran a la misma distancia de la situación en donde todos los individuos obtienen una renta idéntica.

### 4. *COEFICIENTE DE VARIACIÓN*

Resulta de la división de la raíz cuadrada de la varianza entre la media de la distribución de los ingresos.

Al igual que la varianza, el coeficiente de variación es sensible a las transferencias de ingresos, pero valora igual cualquier tipo de transferencias entre individuos ricos y pobres. Entonces, no cumple con la condición de Pigou-Dalton. En este caso, el coeficiente de variación aumenta.

Además, a diferencia de la varianza, el coeficiente de variación es independiente del nivel medio del ingreso. Para evitar que la medida de desigualdad cambie ante transformaciones proporcionales de los ingresos de los individuos (Principio de independencia de escala), el coeficiente de variación supone la división entre la media de los ingresos.

Es evidente que si se unen dos poblaciones con la misma distribución, al igual que la varianza, el valor del coeficiente de variación no varía. Además, las únicas variables que tiene en cuenta dependen solo del ingreso de los individuos (Anonimato o simetría).

Los resultados de las cualidades esperadas del Coeficiente de Variación se resumen en la Tabla 8:

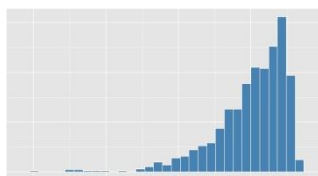
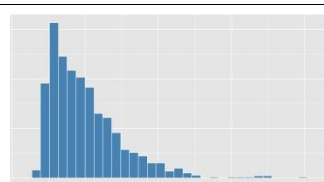
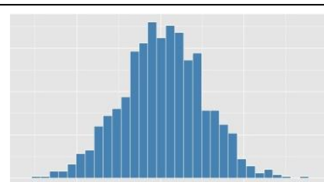
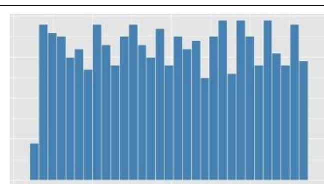
**Tabla 8: Cualidades esperadas del Coeficiente de Variación.**

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | X |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

Una ventaja del coeficiente de variación es que este indicador es independiente del valor medio de la distribución de los ingresos, sin embargo, una desventaja es el inconveniente al adjudicar el mismo peso a las transferencias realizadas en niveles distintos. Al igual que la varianza, la medida usa como referencia la media de los ingresos, en vez de comparar todos los ingresos. Así, se puede dar el caso en que la media no es el ingreso de ningún individuo de la población en estudio. (Sen & Foster, 1997)

A continuación, calculamos en R el Coeficiente de Variación según el tipo de distribución (ver Anexo), y mostramos los resultados en la Tabla 9:

**Tabla 9: Resultados en R del Coeficiente de Variación.**

| **Características.** | **Distribución.** | **Coeficiente de Variación ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 0.09994999 |
| Concentración en ingresos bajos. |  | 0.09994999 |
| Concentración en ingresos medios. |  | 0.09994999 |
| No hay concentración. |  | 0.09994999 |

Al igual que la varianza, todas las distribuciones se encuentran a la misma distancia de la situación en donde todos los individuos obtienen una renta idéntica.

### 5. *DESVIACION ESTANDAR DE LOS LOGARITMOS.*

Es una medida de desigualdad que marca las diferencias entre distribuciones del ingreso de una misma cantidad, pero realizada en diferentes niveles de ingreso, y asigna mayor importancia a las trasferencias de ingresos en el extremo inferior.

Es posible que la desviación estándar de los logaritmos aumente cuando haya transferencias de una persona rica a una persona pobre. Aunque esto ocurre sólo en los niveles de ingresos muy altos, esta medida de desigualdad puede violar el principio de Pigou-Dalton.

Además, posee la propiedad para reflejar una tendencia mayor hacia la igualdad cuando la transferencia se realiza desde una persona con ingresos altos a una con ingresos bajos.

Una ventaja es que la desviación estándar de los logaritmos no requiere corrección por efecto de escala, sin embargo una desventaja seria que el indicador se indetermina si uno de los valores es cero.

Si los ingresos de todos los individuos cambian en la misma proporción, originan el mismo término en cada componente de la diferencia del numerador.

En cuanto al principio de población. Dada una distribución de ingresos . Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es

Para

Para

La medida de desigualdad no cambia.

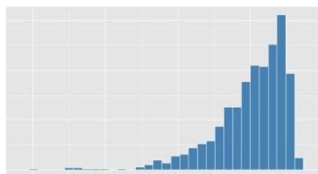
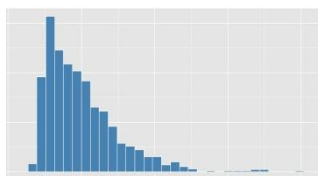
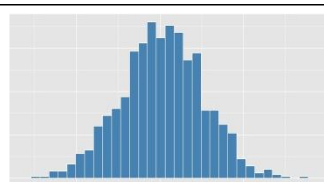
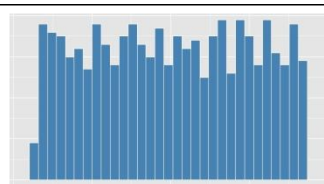
Los resultados de las cualidades esperadas de la Desviación Estándar de los Logaritmos se resumen en la Tabla 10:

**Tabla 10: Cualidades esperadas de la Desviación Estándar de los Logaritmos.**

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | X |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

A continuación, calculamos en R la Desviación Estándar de los Logaritmos según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 11:

**Tabla 11: Resultados en R de la Desviación Estándar de los Logaritmos.**

| **Características.** | **Distribución.** | **Desviación estándar de los Logaritmos ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 0.1124528 |
| Concentración en ingresos bajos. |  | 0.09419439 |
| Concentración en ingresos medios. |  | 0.1013677 |
| No hay concentración. |  | 0.1008012 |

Utilizando la Desviación Estándar de los Logaritmos, el escenario que se encuentra más cerca a una situación donde todos los individuos tienen una renta idéntica es distribución de cola derecha.

## *6. CURVA DE LORENZ.*

La curva de Lorenz es una herramienta gráfica que se realiza mediante la ordenación de menor a mayor de los ingresos de la población.
En el eje de las abscisas se establece el porcentaje acumulado de la población, y en el eje de las ordenadas el porcentaje acumulado de los ingresos.

En cuanto a las propiedades básicas que una medida de desigualdad debe satisfacer, un índice que puede ser calculado a partir de la curva de Lorenz satisface las propiedades de independencia de escala y el principio de población. Puesto que el nivel medio de los ingresos y el volumen total de la población son innecesarios para calcular el indicador. Por lo tanto, ante variaciones de este tipo, el indicador de desigualdad no cambia.(Goerlich, 2001)
Además, la curva de Lorenz satisface el principio de transferencias de Pigou-Dalton, puesto que dichas transferencias nos acercan a la línea de igualdad en algún tramo de la distribución.
Por otro lado, se puede observar que una reordenación de los ingresos de los individuos de una muestra no variaría ni el porcentaje acumulado de la población ni el porcentaje acumulado de los ingresos. Por lo tanto la curva de Lorenz no cambia.

Los resultados de las cualidades esperadas de la Curva de Lorenz se resumen en la Tabla 12:

**Tabla 12: Cualidades esperadas de la Curva de Lorenz**

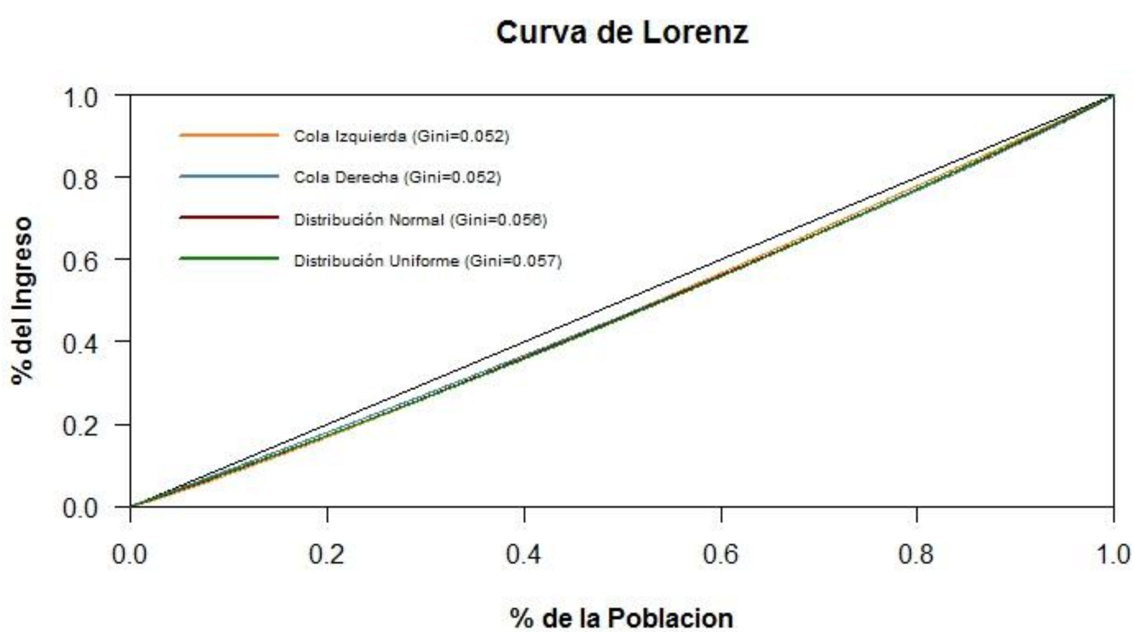
| **Cualidades esperadas** | **¿Se cumple?** |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | ✓ |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

Una interpretación sencilla de la curva de Lorenz es que si la distribución de ingresos en la sociedad es totalmente igualitaria, el 10% de la población que posee menos ingresos

recibe el 10% del total de los ingresos, la mitad del porcentaje acumulado de la población debe recibir la mitad del total de los ingresos, hasta que el 100% de la población recibe el 100% de los ingresos, es decir, todo se distribuye.(Sen, 1973). Por otro lado, es importante señalar que entre más alejada esté la curva de Lorenz de la línea diagonal de equidistribución mayor será la desigualdad en la distribución de ingresos.
Una ventaja de este gráfico es que permite ordenar dos distribuciones, por ejemplo si se tiene un gráfico en la que se presentan dos curvas de Lorenz la curva situada más arriba es la que presenta menores niveles de desigualdad. Una desventaja es cuando dos curvas de Lorenz se cortan entre sí, ya que no se puede establecer cual distribución tiene menor valor de medición de desigualdad, por lo tanto la curva de Lorenz provee un criterio de ordenamiento incompleto.

En R, mediante el paquete `(ineq)`, se calcularon las curvas de Lorenz para cada distribución (Ver Anexo), que se muestran a continuación:

**Gráfico 5: Curva de Lorenz de las 4 distribuciones.**




Es difícil concluir únicamente a partir de la gráfica de la curva de Lorenz, ya que las 4 distribuciones se encuentran a una distancia muy similar de la línea diagonal de equidistribución. Además, como se explica en el siguiente indicador, La curva de Lorenz provee un criterio de ordenamiento incompleto cuando se trata de dos o más curvas que se cortan entre sí, por lo tanto, es mejor analizar la desigualdad con el coeficiente de Gini.

## 7. *COEFICIENTE DE GINI*.

El coeficiente de Gini compensa esta deficiencia de ordenamiento incompleto de la Curva de Lorenz ya que toma una medida de la distancia de la diagonal de la curva de Lorenz en todo su recorrido (Lora, 1987). Se define como la razón de la diferencia entre la diagonal equidistribución, y el área total debajo de esta diagonal.

Si el área entre la diagonal y la curva de Lorenz es llamado X, y Z es el área total bajo la diagonal, el coeficiente de concentración de Gini es igual a X/Z.  Entonces:

El coeficiente de Gini satisface las propiedades de independencia de escala y el principio de población ya que es definido a partir de la curva de Lorenz.
Además, cumple con el principio de las transferencias de Pigou-Dalton; ya que una transferencia de una persona rica a una persona más pobre eleva toda la curva de Lorenz, por lo tanto reduce el valor del índice.

En cuanto al principio de simetría se puede observar que el coeficiente de Gini solo depende de los ingresos, y un intercambio de la cantidad total de ingresos de los individuos no variará ninguno de los ejes de la gráfica de la curva de Lorenz, por lo tanto el coeficiente de Gini no varía.

Los resultados de las cualidades esperadas del Coeficiente de Gini se resumen en la Tabla 13:

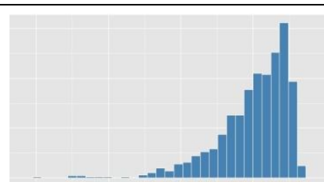
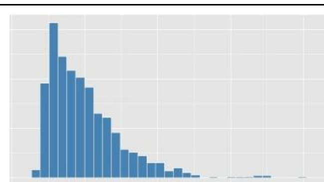
**Tabla 13: Cualidades esperadas del Coeficiente de Gini.**

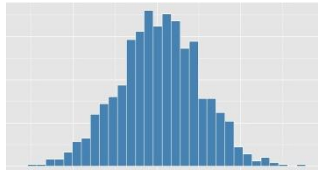
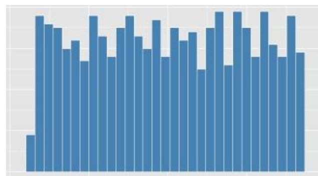
| **Cualidades esperadas** | **¿Se cumple?** |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | ✓ |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

Tiene una ventaja en relación con la varianza, el coeficiente de variación y la desviación estándar de los logaritmos. Como toma en cuenta las desviaciones entre todos los pares de ingresos, evita la concentración en las diferencias con respecto a la media. Por lo tanto, es una medida muy directa de la diferencia del ingreso.

A continuación, calculamos en R el Coeficiente de Gini según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 14:

**Tabla 14: Resultados en R del Coeficiente de Gini.**

| **Características.** | **Distribución.** | **Coeficiente de Gini ( )** |
|---|---|---|
| Concentración en ingresos altos. |  | 0.05228711 |
| Concentración en ingresos bajos. |  | 0.05228711 |

| | | |
|---|---|---|
| Concentración en ingresos medios. |  | 0.05632145 |
| No hay concentración. |  | 0.05770213 |

Como habíamos dicho con la gráfica de la Curva de Lorenz, estas 4 distribuciones se encuentran a una distancia muy similar de la situación en donde todos los individuos obtendrían una renta idéntica, sin embargo, gracias al Coeficiente de Gini, podemos observar esta diferencia y nos damos cuenta que las distribuciones de cola derecha e izquierda tienen el coeficiente de Gini más bajo, lo que indica una menor desigualdad.

### *8. ENTROPIA DE THEIL.*

Esta medida de desigualdad, desarrollada por Henry Theil, se basa en el concepto de entropía de la teoría de la información. Según Sen (1973), la entropía o el contenido de información esperado de la situación, puede considerarse como la sumatoria del contenido de información de cada hecho ponderada por sus respectivas probabilidades:

donde:

- representa la probabilidad de que ocurra cierto hecho.
- es el contenido de información de que el hecho haya ocurrido.

Si se interpreta como la parte del ingreso que recibe cada individuo, se puede utilizar como una medida de desigualdad.

Entonces, si se resta la entropía de una distribución del ingreso de su valor máximo de , se obtiene un indicador que mide la desigualdad. (Sen, 1973)

Una ventaja de este indicador es que permite medir la desigualdad entre subgrupos y dentro de los distintos subgrupos, lo cual resulta en una herramienta útil para análisis regionales de desigualdad. Una desventaja de este indicador consiste en que el resultado puede estar influenciado por el tamaño de la población. (Yalta, 1999).
Para esta medida es cierto que una transferencia de una persona con ingresos más altos a una con ingresos más bajos disminuye el valor de la desigualdad. Por lo tanto cumple con la condición de Pigou-Dalton. Además, al igual que el coeficiente de Gini es invariante con respecto a la escala de medición del ingreso (Principio de independencia de escala).

En cuanto al principio de población. Dada una distribución de ingresos . Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es
Para

Para

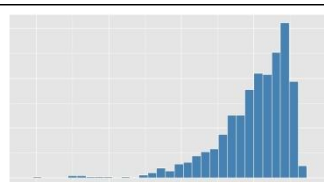
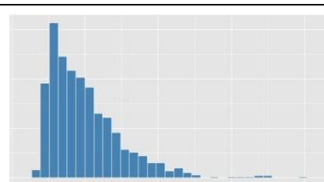
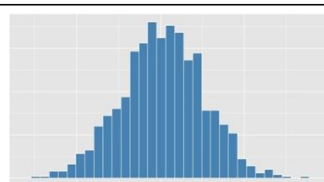
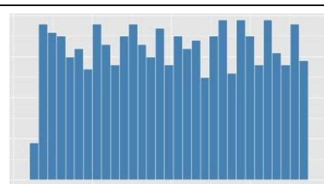
Por lo tanto, la medida de desigualdad de la segunda distribución cambia según el número de réplicas de la primera distribución.
Como se puede observar en la formula, el indicador solo depende de las características del ingreso, por lo tanto cumple el principio de simetría.
Los resultados de las cualidades esperadas de la Entropía de Theil se resumen en la Tabla 15:

Tabla 15: Cualidades esperadas de la Entropía de Theil.

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | ✓ |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | X |
| • *Simetría o anonimato* | ✓ |

A continuación, calculamos en R la Entropía de Theil según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 18:

Tabla 16: Resultados en R de la Entropía de Theil.

| Características. | Distribución. | Entropía de Theil ( ) |
|---|---|---|
| Concentración en ingresos altos. |  | 0.005338988 |
| Concentración en ingresos bajos. |  | 0.004777127 |
| Concentración en ingresos medios. |  | 0.005020451 |
| No hay concentración. |  | 0.005010685 |

Utilizando la Entropía de Theil, el escenario que se encuentra más cerca a una situación donde todos los individuos tienen una renta idéntica, es la distribución de cola derecha.

### 9. MEDIDA DE ATKINSON.

Atkinson parte de su idea del *ingreso equivalente igualitariamente distribuido*. Si el nivel de ingreso por persona es igual entre todos los individuos, el bienestar total sería igual al bienestar generado por la distribución efectiva del ingreso. (Atkinson, 1970)

Los valores que alcanza pueden ser interpretados como la proporción del ingreso al que se estaría dispuesto a renunciar a cambio de la igualdad distributiva. Se define como

Una ventaja es la interpretación sencilla de este indicador, ya que cuando se distribuyen los ingresos de manera uniforme, entonces *el ingreso equivalente igualitariamente distribuido* será igual a la media de la distribución de los ingresos. Por lo tanto, el valor de la medida de Atkinson será cero. Una desventaja de esta medida, implica que se debe escoger el parámetro épsilon, que no es realmente conocido. Además, requiere de un juicio de valor por parte del investigador.

En cuanto a las cualidades que se esperan de un buen indicador. El alcance cumple con el principio de transferencia de Pigou-Dalton.

Dada una distribución. 2,3,5). Si el individuo con mayor ingreso le transfiere una unidad al individuo con menos ingresos, entonces
Para 2,3,5).

Por lo tanto, se puede alcanzar el mismo nivel de bienestar con el 96.56% del ingreso existente, si estuviera igualmente distribuido.

Para

Por lo tanto, se puede alcanzar el mismo nivel de bienestar con el 99.53% del ingreso existente, si estuviera igualmente distribuido.

Entonces, se puede observar que dada una transferencia de una persona rica a una persona pobre, la medida de desigualdad disminuye.

Ahora, dada una distribución. 2,3,5). Si existe un cambio en la misma proporción de ingresos de los individuos. Por ejemplo si aumentan los ingresos de los individuos un 2%. 2.4,3.6,6).

Para 2,3,5)

Para 2.4,3.6,6)

Por lo tanto, cumple con el principio de independencia de escala. La medida de desigualdad no cambia.

Por otro lado, dada una distribución. 2,3,5). Si existe una adición proporcional de individuos a los niveles de ingresos tal que la nueva distribución es 2,3,5,2,3,5).

Para 2,3,5,2,3,5)

La medida de desigualdad no cambia. Entonces, cumple con el principio de población.

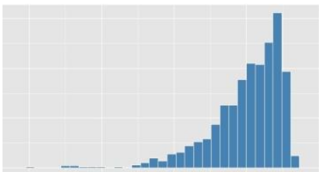
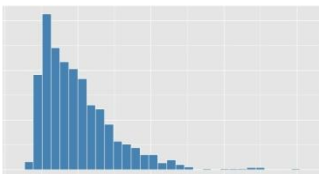
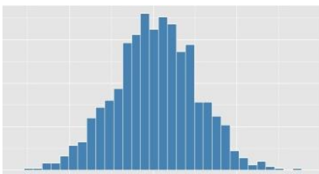
Los resultados de las cualidades esperadas de la Medida de Atkinson se resumen en la Tabla 16:

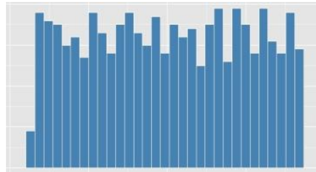
**Tabla 17: Propiedades esperadas de la Medida de Atkinson.**

| Cualidades esperadas | ¿Se cumple? |
|---|---|
| • *Principio de transferencia de Pigou-Dalton* | ✓ |
| • *Independencia de la escala de medición del ingreso* | ✓ |
| • *Principio de la Población de Dalton* | ✓ |
| • *Simetría o anonimato* | ✓ |

A continuación, calculamos en R la Medida de Atkinson según el tipo de distribución (Ver Anexo), y mostramos los resultados en la Tabla 18:

**Tabla 18: Resultados en R de la Medida de Atkinson.**

| Características. | Distribución. | Medida de Atkinson ( ) |
|---|---|---|
| Concentración en ingresos altos. |  | 0.002773225 |
| Concentración en ingresos bajos. |  | 0.00234023 |
| Concentración en ingresos medios. |  | 0.002519905 |

| | | |
|---|---|---|
| No hay concentración. |  | 0.002510545 |

La situación con la menor medida de Atkinson, corresponde a la distribución de cola derecha, donde la concentración se da en los ingresos bajos. 0.00234023 quiere decir que con el 99,76% del ingreso total igualmente distribuido, se podría alcanzar el mismo nivel de bienestar.

### *DESIGUALDAD EN EL BILINGÜISMO.*

Para calcular los niveles de desigualdad en el bilingüismo, utilizamos los 9 indicadores vistos anteriormente, en los resultados del puntaje de inglés de las pruebas SaberPro del año 2011 de 33 departamentos del país, para un total de 249.244 datos. Estos resultados se resumen en la tabla del Anexo 2.
Presentamos los resultados más relevantes obtenidos por cada indicador, y partir de estos, un resultado general sobre la desigualdad en el bilingüismo.

### EL ALCANCE:

| DEPARTAMENTO | ALCANCE |
|---|---|
| CAQUETA | 1,566768 |
| CESAR | 1,561923 |
| CORDOBA | 1,542448 |
| NORTE DE SANTANDER | 1,541474 |
| TOLIMA | 1,535079 |
| HUILA | 1,533942 |
| NARIÑO | 1,530852 |
| CUNDIMARCA | 1,522452 |
| SUCRE | 1,521435 |
| MAGDALENA | 1,521116 |
| CAUCA | 1,520453 |
| BOYACA | 1,510208 |
| BOLIVAR | 1,49774 |
| CALDAS | 1,489655 |
| SANTANDER | 1,489602 |
| VALLE | 1,488597 |

| | |
|---|---|
| META | 1,488164 |
| RISARALDA | 1,476481 |
| ATLANTICO | 1,476124 |
| QUINDIO | 1,474299 |
| CASANARE | 1,472598 |
| ANTIOQUIA | 1,471405 |
| CHOCO | 1,428305 |
| BOGOTA | 1,425674 |
| LA GUAJIRA | 1,420467 |
| ARAUCA | 1,303389 |
| PUTUMAYO | 1,293663 |
| AMAZONAS | 1,24741 |
| SAN ANDRES | 0,6080154 |
| GUAVIARE | 0,5333864 |
| GUAINIA | 0,342393 |
| VAUPES | 0,2319643 |
| VICHADA | 0,1081833 |

Según el Alcance, los departamentos donde más diferencias se encontraron en los puntajes obtenidos fueron Caquetá, Cesar y Córdoba. La diferencia entre la calificación más alta y la más baja, fue más amplia en estos tres departamentos. Guainía, Vaupés y Vichada fueron los departamentos que más se acercaron a una situación donde todos los individuos obtienen la misma calificación.

**DESVIACIÓN MEDIA RELATIVA:**

| DEPARTAMENTO | DESVIACION MEDIA RELATIVA |
|---|---|
| BOGOTA | 0,0920315 |
| SAN ANDRES | 0,09100133 |
| ANTIOQUIA | 0,08720536 |
| ATLANTICO | 0,08682695 |
| RISARALDA | 0,07816314 |
| VALLE | 0,07698413 |
| QUINDIO | 0,07653505 |
| CALDAS | 0,07578968 |
| SANTANDER | 0,07500102 |
| BOLIVAR | 0,07405249 |
| PUTUMAYO | 0,06846093 |
| CAUCA | 0,06715127 |
| MAGDALENA | 0,06577719 |
| CUNDIMARCA | 0,06378111 |
| CORDOBA | 0,06112744 |
| BOYACA | 0,06106031 |
| HUILA | 0,05902381 |
| TOLIMA | 0,05863069 |
| NARIÑO | 0,05571581 |
| META | 0,0553429 |
| NORTE DE SANTANDER | 0,0514257 |
| GUAINIA | 0,05042342 |
| CESAR | 0,0497102 |
| CAQUETA | 0,04928913 |
| SUCRE | 0,04735965 |
| CHOCO | 0,04621969 |
| AMAZONAS | 0,04573634 |
| LA GUAJIRA | 0,04276463 |
| CASANARE | 0,04162586 |
| ARAUCA | 0,03838405 |
| GUAVIARE | 0,03545155 |
| VAUPES | 0,03008463 |
| VICHADA | 0,0268968 |

Según la Desviación Media Relativa, los departamentos donde más diferencias se encontraron en los puntajes obtenidos fueron Bogotá, San Andrés y Antioquia. Guaviare, Vaupés y Vichada fueron los departamentos que más se acercaron a una situación donde todos los individuos obtienen la misma calificación.

**VARIANZA:**

| DEPARTAMENTO | VARIANZA |
|---|---|
| PUTUMAYO | 2,405755 |
| BOGOTA | 1,648578 |
| ANTIOQUIA | 1,575751 |
| ATLANTICO | 1,529741 |
| SAN ANDRES | 1,504589 |
| RISARALDA | 1,273219 |
| VALLE | 1,272153 |
| QUINDIO | 1,25276 |
| CALDAS | 1,201474 |
| BOLIVAR | 1,172581 |
| CAUCA | 1,128982 |
| SANTANDER | 1,121291 |
| CORDOBA | 0,9861406 |
| CUNDIMARCA | 0,9303723 |
| HUILA | 0,9048004 |
| MAGDALENA | 0,8974022 |
| META | 0,8575034 |
| BOYACA | 0,7990133 |
| TOLIMA | 0,7928169 |
| NARIÑO | 0,7904434 |
| AMAZONAS | 0,7312745 |
| CAQUETA | 0,7084154 |
| CHOCO | 0,6245472 |
| LA GUAJIRA | 0,6179731 |
| NORTE DE SANTANDER | 0,607043 |
| SUCRE | 0,5837434 |
| CESAR | 0,5602947 |
| GUAINIA | 0,4402673 |
| ARAUCA | 0,3978665 |
| CASANARE | 0,3845454 |
| GUAVIARE | 0,2689182 |
| VAUPES | 0,1321609 |
| VICHADA | 0,08622296 |

Según la Varianza, los departamentos donde más diferencias se encontraron en los puntajes obtenidos fueron Putumayo, Bogotá y Antioquia. Guaviare, Vaupés y Vichada (nuevamente) fueron los departamentos que más se acercaron a una situación donde todos los individuos obtienen la misma calificación.

### COEFICIENTE DE VARIACIÓN:

| DEPARTAMENTO | COEFICIENTE DE VARIACION |
|---|---|
| PUTUMAYO | 0,1686164 |
| ANTIOQUIA | 0,1235477 |
| BOGOTA | 0,1224429 |
| ATLANTICO | 0,1221211 |
| SAN ANDRES | 0,1165316 |
| VALLE | 0,1123067 |
| RISARALDA | 0,1114393 |
| QUINDIO | 0,1107473 |
| CALDAS | 0,1092199 |
| BOLIVAR | 0,1088483 |
| CAUCA | 0,1084252 |
| SANTANDER | 0,1055087 |
| CORDOBA | 0,1028001 |
| CUNDIMARCA | 0,09822701 |
| HUILA | 0,09792621 |
| MAGDALENA | 0,09670958 |

| | |
|---|---|
| META | 0,09563236 |
| AMAZONAS | 0,09195844 |
| TOLIMA | 0,09173414 |
| NARIÑO | 0,09134447 |
| BOYACA | 0,09059988 |
| CAQUETA | 0,08850393 |
| CHOCO | 0,08467857 |
| LA GUAJIRA | 0,08376946 |
| SUCRE | 0,08072387 |
| NORTE DE SANTANDER | 0,08060457 |
| CESAR | 0,07846602 |
| GUAINIA | 0,0707747 |
| ARAUCA | 0,06711301 |
| CASANARE | 0,06569666 |
| GUAVIARE | 0,05554215 |
| VAUPES | 0,0392224 |
| VICHADA | 0,03176667 |

Según el Coeficiente de Variación, los departamentos donde más diferencias se encontraron en los puntajes obtenidos fueron Putumayo, Antioquia y Bogotá. Guaviare, Vaupés y Vichada (nuevamente) fueron los departamentos que más se acercaron a una situación donde todos los individuos obtienen la misma calificación.

### DESVIACIÓN ESTÁNDAR DE LOS LOGARITMOS:

| DEPARTAMENTO | DESVIACIÓN ESTÁNDAR DE LOS LOGARITMOS |
|---|---|
| GUAINIA | 0,06767252 |
| GUAVIARE | 0,05234385 |
| SAN ANDRES | 0,11179140 |
| VAUPES | 0,03918644 |
| VICHADA | 0,03182521 |

Solo 5 departamentos tuvieron un puntaje mínimo diferente a cero. Es por eso que en los demás departamentos el índice de desigualdad no pudo ser calculado mediante la desviación estándar de los logaritmos. Guainía fue el departamento donde se encontró mayor diferencia en los puntajes obtenidos, mientras que Vichada fue el departamento más cercano a la situación donde todos los individuos obtienen la misma calificación.

**COEFICIENTE DE GINI:**

| DEPARTAMENTO | COEFICIENTE DE GINI |
|---|---|
| BOGOTA | 0,06388237 |
| SAN ANDRES | 0,062871 |
| ANTIOQUIA | 0,06132216 |
| ATLANTICO | 0,0608287 |
| RISARALDA | 0,05535624 |
| VALLE | 0,05485477 |
| QUINDIO | 0,05440933 |
| CALDAS | 0,05390257 |
| PUTUMAYO | 0,0533464 |
| SANTANDER | 0,05300709 |
| BOLIVAR | 0,05270815 |
| CAUCA | 0,04955149 |
| MAGDALENA | 0,04698024 |
| CUNDIMARCA | 0,04615907 |
| CORDOBA | 0,04534686 |
| BOYACA | 0,04387423 |
| HUILA | 0,04382096 |
| TOLIMA | 0,04274389 |
| META | 0,04114726 |
| NARIÑO | 0,04098864 |
| NORTE DE SANTANDER | 0,03769345 |
| CAQUETA | 0,03714768 |
| CESAR | 0,03652311 |
| GUAINIA | 0,03616533 |
| AMAZONAS | 0,03520613 |
| CHOCO | 0,03498267 |
| SUCRE | 0,03496832 |
| LA GUAJIRA | 0,03220774 |
| CASANARE | 0,03054135 |
| ARAUCA | 0,02817695 |
| GUAVIARE | 0,02609018 |
| VAUPES | 0,02147604 |
| VICHADA | 0,01798271 |

Según el coeficiente de Gini, los departamentos de Bogotá, San Andrés y Antioquia, son los que más se alejan de la diagonal de equidistribucion. Guaviare, Vaupés y Vichada son los departamentos donde se encontraron menos diferencias en los puntajes obtenidos.

**ENTROPÍA DE THEIL:**

| DEPARTAMENTO | ENTROPÍA DE THEIL |
|---|---|
| BOGOTA | 0,006745412 |
| SAN ANDRES | 0,006566406 |
| ATLANTICO | 0,006342158 |
| ANTIOQUIA | 0,006212632 |
| RISARALDA | 0,005238127 |
| QUINDIO | 0,005136484 |
| VALLE | 0,005092721 |
| CALDAS | 0,00496148 |
| SANTANDER | 0,004885299 |
| BOLIVAR | 0,004666288 |
| CAUCA | 0,004077851 |
| MAGDALENA | 0,003912999 |
| CUNDIMARCA | 0,003867152 |
| CORDOBA | 0,003504754 |
| BOYACA | 0,00333092 |
| HUILA | 0,003323837 |
| TOLIMA | 0,003183176 |
| NARIÑO | 0,002876735 |
| CAQUETA | 0,002712748 |
| META | 0,002643137 |
| NORTE DE SANTANDER | 0,002614865 |
| GUAINIA | 0,002425406 |
| CESAR | 0,002383067 |
| SUCRE | 0,00198299 |
| CHOCO | 0,001978513 |
| AMAZONAS | 0,001725882 |
| PUTUMAYO | 0,001578927 |
| CASANARE | 0,001565324 |
| LA GUAJIRA | 0,00153766 |
| GUAVIARE | 0,001475122 |
| ARAUCA | 0,001237728 |
| VAUPES | 0,000768084 |
| VICHADA | 0,000505039 |

Según la Entropía de Theil, los departamentos donde se encontraron mayor diferencia en los puntajes obtenidos fueron Bogotá, San Andrés y Atlántico. Arauca, Vaupés y Vichada fueron los departamentos que más se acercaron a una situación donde todos los individuos obtienen la misma calificación.

**MEDIDA DE ATKINSON:**

Según la Medida de Atkinson, todos los departamentos obtuvieron el mismo resultado (Ver Anexo 2) que se aproximó a cero. Es decir, que con este indicador, no existe diferencia en los puntajes obtenidos en los diferentes departamentos.

***CONCLUSIÓN***

Vemos una tendencia[1] en los resultados de los cálculos anteriores, donde Bogotá, San Andrés y Antioquia, se repiten en la mayoría de indicadores como los departamentos con las mayores diferencias en los puntajes de inglés obtenidos en las pruebas SaberPro. Esto significa que en éstos departamentos, la diferencia del individuo con el mejor puntaje, y el que menor puntaje, es relativamente más grande que en los demás departamentos, y que la media muestral de los puntajes obtenidos se encuentra por encima del promedio nacional, que es 9,72. Por el otro lado, Vaupés y Vichada demostraron ser los departamentos con las menores diferencias en los puntajes obtenidos. Es decir, que la diferencia del individuo con mayor puntaje, y el de menor puntaje, es relativamente menor que en los demás departamentos. También vemos que la media muestral de los puntajes obtenidos en estos departamentos, se encuentra por debajo del promedio nacional.

Lo que nos permite concluir que en los departamentos con puntajes medios por encima del promedio nacional, existe una mayor desigualdad en el bilingüismo. Mientras que en los departamentos con puntajes medios por debajo del promedio nacional, ocurre lo contrario.

[1] Sin contar la medida del Alcance, ya que este indicador como solo toma en cuenta los valores extremos, mientras que los demás tienen en cuenta cada calificación, arrojó tendencia fue diferente.

## *REFERENCIAS.*


Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory, 2*(3), 244-263.

Cortés, F., & Rubalcava, R. M. (1982). *Técnicas estadísticas para el estudio de la desigualdad social*. Mexico.

Dalton, H. (1920). *The measurement of the inequality of incomes*.

Foster, J. E. (1985). Inequality measurement. Fair Allocation, Proceedings of Symposia in Applied Mathematics.

Goerlich, F. J. (2001). *Desigualdad, diversidad y convergencia: (más) instrumentos de medida, modelos de regresión*: Instituto Valenciano de Investigaciones Económicas.

Kuznets, S. S. (1950). *Shares of upper income groups in income and savings*: National Bureau of Economic Research.

Lora, E. (1987). *Técnicas de Medición Económica: Metodología y Aplicaciones en Colombia*: Fundacion para la Educacion Superior y el Desarrollo, Fedesarrollo.

Núñez Velázquez, J. J. (2006). La desigualdad económica medida a través de las curvas de Lorenz. Métodos Cuantitativos para la Economía y la Empresa.

Pigou, A. C. (1912). *Wealth and welfare*. London,: Macmillan and co., limited.

Rousseau, J. J. (1755). *Discours sur l'origine & les fondements de l'inégalité parmi les hommes*: chez Marc Michel Rey.

Sen, A. (1973). *On economic inequality*. Oxford,: Clarendon Press.

Sen, A., & Foster, J. E. (1997). *On economic inequality* (Enlarged ed.). Oxford

New York: Clarendon Press ;

Oxford University Press.

Yalta, N. S. (1999). *La pobreza, la desigualdad y la educación en el Perú de hoy: una aproximación cuantitativa*: Universidad de Chile, Facultad de Ciencias Físicas, Departamento de Ingeniería Industrial.

*ANEXOS.*

**1)**

- **Calculo del Alcance en R de las 4 distribuciones:**

```
> ((max(x)-min(x))/(colMeans(x)))
```

donde

  - : Es el *data frame* que pertenece a cada distribución.

- **Calculo de la Desviación Media Relativa en R de las 4 distribuciones.**

```
> (sum(abs(colMeans(x)-x)))/(nrow(x)*colMeans(x))
```

donde

  - : Es el *data frame* que pertenece a cada distribución.

- **Calculo de la Varianza en R de las 4 distribuciones.**

```
> ((sum((colMeans(x)-x)^2))/ (nrow(x))
```

donde

  - : Es el *data frame* que pertenece a cada distribución.

- **Calculo del Coeficiente de Variación en R de las 4 distribuciones.**

```
> ((((sum((colMeans(x)-x)^2))/(nrow(x))^(1/2))/colMeans(x))
```

donde

  - : Es el *data frame* que pertenece a cada distribución

- **Calculo de la Desviación Estándar de los Logaritmos en R de las 4 distribuciones.**

```
> ((sum(((log(mean(x)))-((log(x))))^2))/nrow(x))^(1/2)
```

donde

- : Es el *data frame* que pertenece a cada distribución

- **Calculo del Coeficiente de Gini en R de las 4 distribuciones.**

```
ineq(x$y, parameter=NULL, type=" ")
```

donde

- : Es el *data frame* que pertenece a cada distribución.
- : El nombre de la columna de datos en el *data frame* .
- : Parámetro de la medida de desigualdad (Si se ajusta en `NULL`, se utiliza el parámetro por defecto de la respectiva medida.
- : Cadena de caracteres, que da la medida utilizada para calcular la desigualdad. En este caso “`Gini`”.

- **Calculo de la Entropía de Theil.**

```
ineq(x$y, parameter=NULL, type=" ")
```

donde

- : Es el *data frame* que pertenece a cada distribución.
- : El nombre de la columna de datos en el *data frame* .
- : Parámetro de la medida de desigualdad (Si se ajusta en `NULL`, se utiliza el parámetro por defecto de la respectiva medida.)
- : Cadena de caracteres, que da la medida utilizada para calcular la desigualdad. En este caso “`Theil`”.

- **Calculo de la Medida de Atkinson.**

```
ineq(x$y, parameter=NULL, type=" ")
```

donde

- : Es el *data frame* que pertenece a cada distribución.
- : El nombre de la columna de datos en el *data frame* .
- : Parámetro de la medida de desigualdad (Si se ajusta en NULL, se utiliza el parámetro por defecto de la respectiva medida.
- : Cadena de caracteres, que da la medida utilizada para calcular la desigualdad. En este caso “Atkinson”.

**2)**

| DEPARTAMENTO | TAMAÑO | MEDIA MUESTRAL | DESVIACION ESTÁNDAR | VALOR MINIMO | VALOR MAXIMO | EL ALCANCE | LA DESVIACION MEDIA RELATIVA | LA VARIANZA | EL COEFICIENTE DE VARIACION | LA DESVIACION ESTANDAR DE LOS LOGARITMOS | COEFICIENTE DE GINI | ENTROPÍA DE THEIL | MEDIDA DE ATKINSON |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AMAZONAS** | 204 | 9,299265 | 0,8572496 | 0 | 11,6 | 1,24741000 | 0,04573634 | 0,73127450 | 0,09195844 | NaN | 0,03520613 | 0,00172588 | 0 |
| **ANTIOQUIA** | 33797 | 10,16036 | 1,2553080 | 0 | 14,95 | 1,4714050 | 0,08720536 | 1,575751 | 0,1235477 | NaN | 0,06132216 | 0,006212632 | 0 |
| **ARAUCA** | 505 | 9,398574 | 0,6313920 | 0 | 12,25 | 1,30338900 | 0,03838405 | 0,39786650 | 0,06711301 | NaN | 0,02817695 | 0,00123773 | 0 |
| **ATLANTICO** | 12699 | 10,12787 | 1,2368760 | 0 | 14,95 | 1,47612400 | 0,08682695 | 1,52974100 | 0,12212110 | NaN | 0,06082870 | 0,00634216 | 0 |
| **BOGOTA** | 73237 | 10,48627 | 1,2839780 | 0 | 14,95 | 1,425674 | 0,0920315 | 1,648578 | 0,1224429 | NaN | 0,06388237 | 0,006745412 | 0 |
| **BOLIVAR** | 9636 | 9,94832 | 1,0829140 | 0 | 14,9 | 1,49774000 | 0,07405249 | 1,17258100 | 0,10884830 | NaN | 0,05270815 | 0,00466629 | 0 |
| **BOYACA** | 6303 | 9,866188 | 0,8939464 | 0 | 14,9 | 1,51020800 | 0,06106031 | 0,79901330 | 0,09059988 | NaN | 0,04387423 | 0,00333092 | 0 |
| **CALDAS** | 5872 | 10,035880 | 1,0962110 | 0 | 14,95 | 1,48965500 | 0,07578968 | 1,20147400 | 0,10921990 | NaN | 0,05390257 | 0,00496148 | 0 |
| **CAQUETA** | 1464 | 9,510020 | 0,8419618 | 0 | 14,9 | 1,56676800 | 0,04928913 | 0,70841540 | 0,08850393 | NaN | 0,03714768 | 0,00271275 | 0 |
| **CASANARE** | 1775 | 9,439099 | 0,6202920 | 0 | 13,9 | 1,47259800 | 0,04162586 | 0,38454540 | 0,06569666 | NaN | 0,03054135 | 0,00156532 | 0 |
| **CAUCA** | 4323 | 9,799713 | 1,0626580 | 0 | 14,9 | 1,520453 | 0,06715127 | 1,128982 | 0,1084252 | NaN | 0,04955149 | 0,004077851 | 0 |
| **CESAR** | 3287 | 9,539522 | 0,7486422 | 0 | 14.9 | 1,561923 | 0,0497102 | 0,5602947 | 0,07846602 | NaN | 0,03652311 | 0,002383067 | 0 |
| **CHOCO** | 2278 | 9,332739 | 0,7904565 | 0 | 13.33 | 1,42830500 | 0,04621969 | 0,62454720 | 0,08467857 | NaN | 0,03498267 | 0,00197851 | 0 |
| **CORDOBA** | 4074 | 9,659971 | 0,9931680 | 0 | 14,9 | 1,54244800 | 0,06112744 | 0,98614060 | 0,10280010 | NaN | 0,04534686 | 0,00350475 | 0 |
| **CUNDIMARCA** | 7923 | 9,819683 | 0,9646190 | 0 | 14,95 | 1,52245200 | 0,06378111 | 0,93037230 | 0,09822701 | NaN | 0,04615907 | 0,00386715 | 0 |
| **GUAINIA** | 52 | 9,375192 | 0,6700000 | 8,39 | 11,6 | 0,34239300 | 0,05042342 | 0,44026730 | 0,07077470 | 0,06767252 | 0,03616533 | 0,00242541 | 0 |
| **GUAVIARE** | 140 | 9,336571 | 0,5204353 | 8,17 | 13,15 | 0,53338640 | 0,03545155 | 0,26891820 | 0,05554215 | 0,05234385 | 0,02609018 | 0,00147512 | 0 |
| **HUILA** | 3804 | 9,713538 | 0,9513350 | 0 | 14,9 | 1,53394200 | 0,05902381 | 0,90480040 | 0,09792621 | NaN | 0,04382096 | 0,00332384 | 0 |
| **LA GUAJIRA** | 1295 | 9,384239 | 0,7864164 | 0 | 13,33 | 1,42046700 | 0,04276463 | 0,61797310 | 0,08376946 | NaN | 0,03220774 | 0,00153766 | 0 |
| **MAGDALENA** | 3399 | 9,795443 | 0,9474525 | 0 | 14,9 | 1,521116 | 0,06577719 | 0,8974022 | 0,09670958 | NaN | 0,04698024 | 0,003912999 | 0 |
| **META** | 3847 | 9,683070 | 0,9261352 | 0 | 14,41 | 1,488164 | 0,0553429 | 0,8575034 | 0,09563236 | NaN | 0,04114726 | 0,002643137 | 0 |
| **NARIÑO** | 4711 | 9,733144 | 0,8891632 | 0 | 14,9 | 1,53085200 | 0,05571581 | 0,79044340 | 0.09134447 | NaN | 0,04098864 | 0,00287674 | 0 |
| **NORTE DE SANTANDER** | 6682 | 9,666073 | 0,7791879 | 0 | 14,9 | 1,54147400 | 0,05142570 | 0,60704300 | 0,08060457 | NaN | 0,03769345 | 0,00261487 | 0 |
| **PUTUMAYO** | 814 | 9,198686 | 1,5520030 | 0 | 11,9 | 1,29366300 | 0,06846093 | 2,40575500 | 0,16861640 | NaN | 0,05334640 | 0,00157893 | 0 |
| **QUINDIO** | 3393 | 10,106500 | 1,0106500 | 0 | 14,9 | 1,47429900 | 0,07653505 | 1,25276000 | 0,11074730 | NaN | 0,05440933 | 0,00513648 | 0 |
| **RISARALDA** | 4886 | 10,125430 | 1,1284860 | 0 | 14,95 | 1,476481 | 0,07816314 | 1,273219 | 0,1114393 | NaN | 0,05535624 | 0,005238127 | 0 |
| **SAN ANDRES** | 81 | 10,526050 | 1,2342600 | 8.5 | 14,9 | 0,60801540 | 0,09100133 | 1,50458900 | 0,11653160 | 0,11179140 | 0,06287100 | 0,00656641 | 0 |
| **SANTANDER** | 15285 | 10,036240 | 1,0589450 | 0 | 14,95 | 1,48960200 | 0,07500102 | 1,12129100 | 0,10550870 | NaN | 0,05300709 | 0,00488530 | 0 |
| **SUCRE** | 3718 | 9,464747 | 0,7641338 | 0 | 14,4 | 1,52143500 | 0,04735965 | 0,58374340 | 0,08072387 | NaN | 0,03496832 | 0,00198299 | 0 |
| **TOLIMA** | 7916 | 9,706339 | 0,8904589 | 0 | 14,9 | 1,535079 | 0,05863069 | 0,7928169 | 0,09173414 | NaN | 0,04274389 | 0,003183176 | 0 |
| **VALLE** | 21741 | 10,043020 | 1,1279240 | 0 | 14,95 | 1,48859700 | 0,07698413 | 1,27215300 | 0,11230670 | NaN | 0,05485477 | 0,00509272 | 0 |
| **VAUPES** | 75 | 9,268667 | 0,3659875 | 8.39 | 10,54 | 0,23196430 | 0,03008463 | 0,13216090 | 0,03922240 | 0,03918644 | 0,02147604 | 0,00076808 | 0 |
| **VICHADA** | 28 | 9,243571 | 0,2990258 | 8.7 | 9,7 | 0,10818330 | 0,02689680 | 0,08622296 | 0,03176667 | 0,03182521 | 0,01798271 | 0,00050504 | 0 |