

# Midiendo el desarrollo entre hombres y mujeres

*María Fernanda Puddy*

*Erik Angel*

*Andrés Bermúdez*

*Daniel Sandoval*

*Gabriel A. Fuentes*

*Abril, 2022*

- a) Compara la media de ingreso entre hombres y mujeres y comenta acerca de las diferencias que encontraste (presenta los valores y agrega como máximo dos párrafos explicativos del por qué consideras que existe la diferencia, utiliza soporte académico).

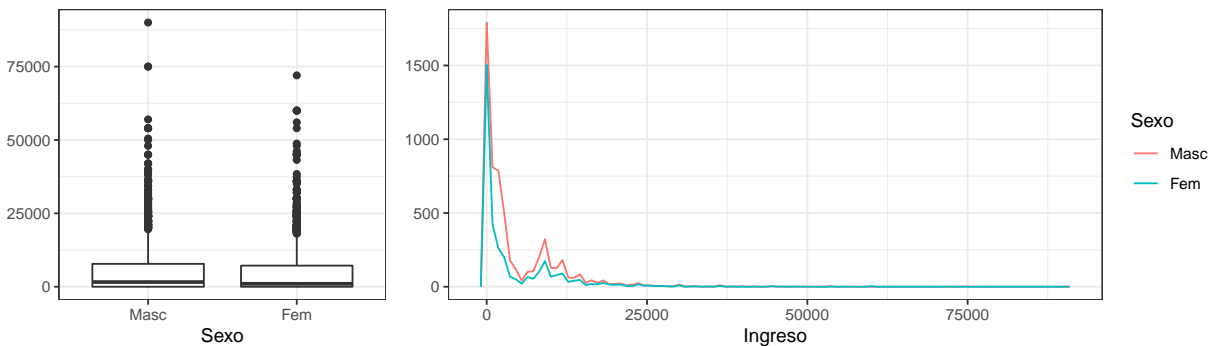
En nuestro análisis, los sujetos se definen como todos los individuos que cuentan con la edad mínima para trabajar (15 años) o más y que permanecen ocupados en uno o más trabajos. Además, el ingreso de las personas está definido como la sumatoria de su salario base más sus ingresos por horas laborales extras, comisiones, ingresos por trabajar durante algún periodo vacacional, bono 14, aguinaldo, bono vacacional, décimo quinto sueldo y bono de productividad, así como cobertura de viáticos por motivo de alimentación, vivienda y transporte.

El cuadro 1 presenta las estadísticas descriptivas de esta variable cuando la muestra es segregada por sexo. De tal forma, es posible apreciar que el ingreso medio es más elevado para el sexo masculino por aproximadamente 300 quetzales. Gamboa (2021)—quien también utiliza los datos de la Encuesta Nacional de Empleo e Ingresos (2019)—encuentra que la marcada brecha salarial por género en el país se explica debido a que la participación de la fuerza laboral femenina se concentra dentro de ocupaciones caracterizadas por tener bajas y medias remuneraciones salariales.

**Cuadro 1:** Estadísticas descriptivas de ingresos entre hombres y mujeres

sexo	Media	Mediana	Desv. Est.	Mínimo	Máximo	Asimetría	Curtosis	Obs.
Masc	4436.01	1600	6479.63	0	90000	3.05	17.70	5937
Fem	4136.75	1000	6845.39	0	72000	3.03	14.58	3489

**Figura 1:** Boxplot e histograma de frecuencias poligonales de ingresos entre hombres y mujeres



- b) Compara ahora las medias y determina si estadísticamente son distintas (el hecho de que los valores absolutos sean distintos no necesariamente significa que estadísticamente lo sean). Presenta tu cálculo y la evidencia para afirmar tu respuesta.

Una prueba  $t$  de Welch es utilizada en este caso debido a su mayor robustez durante situaciones donde tanto los tamaños como las varianzas de ambas muestras difieren entre sí. El estadístico  $t$  de las medias muestrales se computa como:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_{\overline{X}_1}^2 - s_{\overline{X}_2}^2}} \quad (1)$$

donde  $\overline{X}_i$  y  $s_{\overline{X}_i}$  representan la media y error estándar de la muestra  $i$ , respectivamente. El cuadro 2 presenta los resultados de esta prueba, los cuales confirman la significancia estadística en la diferencia de las medias de ambos grupos a un nivel de confianza del 95%.

**Cuadro 2:** Prueba t de dos muestras de Welch

Estadístico t	Valor p
2.090006	0.0366533

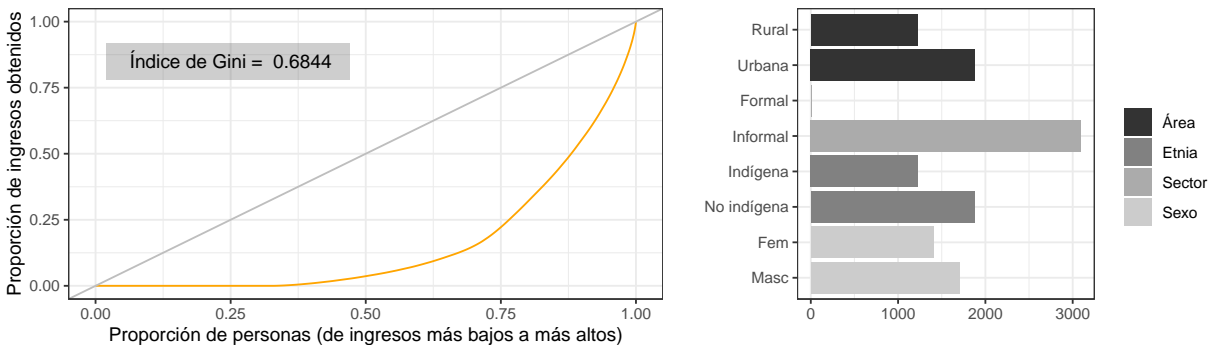
- c) Calcula el coeficiente de Gini de toda la muestra contenida en la base de datos y comenta cuál es tu apreciación del coeficiente de Gini encontrado (un párrafo como máximo).

El coeficiente de Gini, una medida que demuestra el grado de desigualdad en una distribución de ingresos/riqueza, se define como:

$$Gini = \frac{1}{2n^2\mu} \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|, \quad (2)$$

donde  $\mu$  es el ingreso medio,  $n$  el número de individuos en la muestra, mientras que  $y_i$  y  $y_j$  representan el ingreso del individuo  $i$  y  $j$ , respectivamente. Dado a que valores elevados en el índice corresponden a un mayor nivel de desigualdad, es posible apreciar que la desigualdad en ingresos es considerablemente elevada en la muestra general. Sin embargo, es importante hacer notar que el cálculo de este coeficiente se ve afectado por el considerable número de observaciones con un nivel nulo de ingresos en la base de datos (32.93%).

**Figura 2:** Curva de Lorentz para la muestra general y distribución de zeros en la base de datos



- d) Calcula el coeficiente de Gini únicamente para la población masculina y para la femenina, compara los valores absolutos y determina si estadísticamente son distintos (analiza este punto con dos párrafos explicativos máximo).

Se realizó un proceso de bootstrapping, de remuestreo, para los índices de GINI para poder generar los estadísticos necesarios como el error estándar y así, poder realizar una prueba t. Analizando los resultados obtenidos en el cuadro 3 se puede observar desde el inicio una media del coeficiente de Gini superior para mujeres con 0.5478 que para hombres con 0.5357, al igual que en sus intervalos, donde las mujeres poseen un 0.5081 de mínimo mientras los hombres tienen 0.5044, y un máximo para mujeres de 0.5809 mientras para hombres es de 0.5721. Demostrando en términos absolutos una diferencia, aunque pequeña, existente en términos de desigualdad económica con respecto al sexo.

Por otro lado, en términos estadísticos se determinó a través de una prueba t, que si existe una diferencia estadística entre el coeficiente de Gini para hombres y para mujeres. Lo cual quiere decir que estadísticamente no hay una distribución de las riquezas igualitaria en Guatemala con respecto al sexo.

**Cuadro 3:** Bootstrapping del coeficiente de Gini por género

Grupo	Media	Desv. Est.	Mínimo	Máximo	Obs.		Valor
Hombres	0.6615	0.0105	0.6325	0.6970	1000	Estadístico t	-126.3206
Mujeres	0.7209	0.0105	0.6840	0.7604	1000	Valor p	0.0000

- e) Calcula el anterior inciso d) pero ahora segmentando a los hombres indígenas de los no indígenas y a las mujeres indígenas de las no indígenas. Comenta las diferencias de las medias de ingreso y de los Gini de los cuatro grupos segmentados.

De la misma manera se realizó el método de Bootstrap para el análisis a continuación. En primer lugar, en el cuadro 4 se presentan los datos para el coeficiente de Gini para hombres por autodenominación étnica donde se separa a la muestra entre indígena (lo cual incluye a aquellos que se autodenominan mayas, xincas, y garífunas) y no indígena. Lo que se observa es una clara desigualdad inferior al ser no indígena en el caso de los hombres con una media de Gini de 0.494 para no indígenas y un 0.5603 para indígenas, con una diferencia de 0.0663 lo cual puede llegar a ser hasta preocupante el ver una diferencia significativa tomando en cuenta las condiciones en las que se realiza la ENEI (no llega a las personas del interior o más aisladas, mayoritariamente indígenas). Es importante ver la diferencia entre los intervalos para los hombres indígenas y no indígenas donde se ve nuevamente una clara desigualdad inferior para los no indígenas. Asimismo, por una prueba de t estadística, se demuestra que estadísticamente son distintos los coeficientes de Gini para hombres indígenas y no indígenas.

**Cuadro 4:** Bootstrapping del coeficiente de Gini entre hombres por autodenominación étnica

Grupo	Media	Desv. Est.	Mínimo	Máximo	Obs.		Valor
Indígena	0.7255	0.0109	0.6895	0.7615	1000	Estadístico t	235.1054
No indígena	0.6075	0.0115	0.5735	0.6396	1000	Valor p	0.0000

En segundo lugar, en el cuadro 5 se encuentra el coeficiente de Gini para mujeres con la misma clasificación por autodenominación étnica. En este caso se repite el patrón donde una mujer no indígena presenta un coeficiente de Gini inferior, donde la media de una mujer no indígena se encuentra en 0.5083 mientras que para una mujer indígena es de 0.6545, los intervalos igualmente demuestran esta desigualdad con respecto a la autodenominación étnica. Y de igual forma se comprueba estadísticamente a partir de una prueba t la diferencia estadísticamente significativa entre un coeficiente de Gini con respecto al otro, en este caso de mujeres indígenas y no indígenas. Una conclusión a la que se puede llegar estadísticamente por estos resultados es que es conveniente tanto el ser hombre como el ser no indígena para afrontar una desigualdad inferior, lo cual es realmente preocupante.

**Cuadro 5:** Bootstrapping del coeficiente de Gini entre mujeres por autodenominación étnica

Grupo	Media	Desv. Est.	Mínimo	Máximo	Obs.		Valor
Indígena	0.8340	0.0086	0.8107	0.8609	1000	Estadístico t	356.7282
No indígena	0.6737	0.0113	0.6382	0.7087	1000	Valor p	0.0000

- f) Con base a la Teoría del Capital Humano (puedes revisar el estudio de Becker disponible en el portal y la presentación), arma una ecuación minceriana para evaluar los efectos de cada variable del Capital Humano en la muestra total.

$$\ln w_i = f(s_i, x_i) = w_0 + \rho s_i + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \quad (2)$$

donde  $w$  representa el ingreso mensual de la persona  $i$ ,  $s_i$  sus años de educación y  $x_i$  sus años de experiencia laboral potencial (edad menos años de educación menos seis).

La ecuación Minceriana surge a partir de la teoría del capital humano. La popularidad de esta metodología reside en la relación que existe entre los años de escolaridad con la tasa de retorno de la educación, es decir, el retorno de la inversión en educación para los individuos. Esta utiliza un método de mínimos cuadrados ordinarios usando como variable dependiente el logaritmo de los ingresos y como variables independientes los años de educación, la experiencia laboral y el cuadrado de la misma.

La ecuación toma en cuenta algunos supuestos, entre ellos; el retorno de la educación es independiente de los años de experiencia ya que no se toma en cuenta la capacitación y educación que las personas reciben al estar empleadas. Se espera una relación positiva entre años de educación y experiencia con los ingresos, sin embargo, al considerar la experiencia al cuadrado esta tendrá un coeficiente negativo. Esto se explica ya que cada año adicional de experiencia tiene un rendimiento marginal menor sobre los ingresos.

Al plantear la ecuación Minceriana, según los resultados obtenidos los años de educación y experiencia tienen una relación positiva con el ingreso y son estadísticamente significativas. La tasa de rentabilidad de la educación es 13.3% , lo que indica que al aumentar los años de educación los ingresos de los individuos aumentan.

**Cuadro 6:** Función de ingresos de Mincer para la muestra completa

Coefficiente	Valor estimado	Error estándar	Estadístico t	Valor p
Intercepto	6.0979728	0.0596687	102.19711	< 2e-16
Educación	0.2022493	0.0068617	29.47531	< 2e-16
Experiencia	0.0905056	0.0037307	24.25997	< 2e-16
Experiencia acumulada	-0.0012363	0.0000596	-20.74159	< 2e-16

$R^2$  ajustado: 0.1837375 | Estadístico  $F$ : 475.277557 (< 2.2e-16)

- g) Procede a evaluar el inciso f) realizando una ecuación para hombres y otra para mujeres, asegurándote de incluir en tu modelo variables extra (además de las variables propias del Capital Humano) que permitan evaluar el efecto de pertenecer al sector formal o informal; de estar en el área urbana o rural; y de ser indígena o no indígena (presenta tus cuadros de salida para ambas regresiones y analiza estadísticamente los resultados que obtuviste).

Tomando en cuenta el planteamiento anterior, se realizó un modelo para contrastar los resultados entre hombres y mujeres. De la misma forma se incluyeron 3 variables dicotómicas, definidas como: Área Urbana (1), Sector Formal (1), No Indígena (1)

En ambos modelos, todas las variables son estadísticamente significativas. Sin embargo, el modelo de mujeres presenta un  $R^2$  mayor. Se evidenció que la tasa de rentabilidad de la educación es mayor para las mujeres que para los hombres, 8% y 6% respectivamente. Esto se puede explicar ya que en Guatemala, las mujeres tienen un menor acceso a educación que los hombres. Caso contrario para la experiencia, esta variable tiene un mayor peso para el ingreso de los hombres que de las mujeres.

Respecto al resto de variables estudiadas en el modelo, todas tienen una relación positiva con los ingresos. Sin embargo, el pertenecer al sector formal es la variable con mayor peso dentro del modelo. Al ser parte del sector formal los individuos tienen acceso a un salario mínimo, aunque en Guatemala este no es suficiente para cubrir todas las necesidades básicas, ofrecen estabilidad en los ingresos y algunas prestaciones adicionales.

**Cuadro 7:** Función de ingresos de Mincer para hombres (controlando por área, sector y autodenominación étnica)

Coefficiente	Valor estimado	Error estándar	Estadístico t	Valor p
Intercepto	6.3506518	0.0479666	132.39739	< 2e-16
Educación	0.0741099	0.0055554	13.34027	< 2e-16
Experiencia	0.0439350	0.0029478	14.90440	< 2e-16
Experiencia acumulada	-0.0006122	0.0000465	-13.17523	< 2e-16
Área (urbana)	0.2878691	0.0223220	12.89620	< 2e-16
Sector (formal)	1.4509827	0.0229813	63.13748	< 2e-16
Autodenominación (no indígena)	0.2781599	0.0244000	11.40001	< 2e-16

$R^2$  ajustado: 0.6366784 | Estadístico  $F$ : 1237.8901792 (< 2.2e-16)

**Cuadro 8:** Función de ingresos de Mincer para mujeres (controlando por área, sector y autodenominación étnica)

Coefficiente	Valor estimado	Error estándar	Estadístico t	Valor p
Intercepto	6.0310946	0.0803504	75.059881	< 2e-16
Educación	0.1023565	0.0092319	11.087273	< 2e-16
Experiencia	0.0407729	0.0049369	8.258746	2.59e-16
Experiencia acumulada	-0.0005951	0.0000793	-7.504822	9.07e-14
Área (urbana)	0.2754603	0.0411450	6.694872	2.77e-11
Sector (formal)	1.6259108	0.0359074	45.280669	< 2e-16
Autodenominación (no indígena)	0.3585666	0.0437176	8.201878	4.10e-16

$R^2$  ajustado: 0.6390504 | Estadístico  $F$ : 616.2383552 (< 2.2e-16)

h) Con base a los cálculos que realizaste en los incisos previos, utiliza dos párrafos por ítem para comentar lo siguiente:

— ¿Es Guatemala un país con alta desigualdad?

— ¿Existen diferencias significativas por sexo en el mercado laboral?

— ¿Es más alta la desigualdad entre hombres o entre mujeres y por qué?

— ¿Consideras que todas las diferencias de ingreso que encontraste en tu análisis son producto de la discriminación?

— ¿Cuáles son tus sugerencias para reducir las brechas de ingreso entre hombres y mujeres y fortalecer el camino al desarrollo?

## Referencias

- Björklund, A., & Kjellström, C. (2002). Estimating the return to investments in education: How useful is the standard Mincer equation? *Economics of Education Review*, 21(3), 195–210. [https://doi.org/10.1016/S0272-7757\(01\)00003-6](https://doi.org/10.1016/S0272-7757(01)00003-6)
- Giles, D. E. A. (2004). Calculating a Standard Error for the Gini Coefficient: Some Further Results\*. *Oxford Bulletin of Economics and Statistics*, 66(3), 425–433. <https://doi.org/10.1111/j.1468-0084.2004.00086.x>
- Hoover, G. & Yaya, M. (2010). *Racial/Ethnic Differences in Income Inequality across US Regions*. Retrieved April 23, 2022, from <<https://journals.sagepub.com/doi/abs/10.1007/s12114-010-9057-z>>
- Mirzaei, S., Borzadaran, G. R. M., Amini, M., & Jabbari, H. (2017). A comparative study of the Gini coefficient estimators based on the regression approach. *Communications for Statistical Applications and Methods*, 24(4), 339–351. <https://doi.org/10.5351/CSAM.2017.24.4.339>
- Gamboa, D. (2021). *Diferencia salarial entre hombres y mujeres: Un estudio de Guatemala desde la perspectiva de las ocupaciones en el 2019*. *Revista Académica ECO*, (24), 21-37.