

# Correlación estadística entre el IDH y otros indicadores macroeconómicos registrados en 2020

Luis Diego Esquivel Cruz<sup>1</sup>, Ester Andrea Gutiérrez Elizondo<sup>2</sup>, Ana Laura López Murillo<sup>3</sup>  
[luis.esquivel@ucr.ac.cr](mailto:luis.esquivel@ucr.ac.cr), [ester.gutierrezelizondo@ucr.ac.cr](mailto:ester.gutierrezelizondo@ucr.ac.cr), [ana.lopezmurillo@ucr.ac.cr](mailto:ana.lopezmurillo@ucr.ac.cr)

## RESUMEN

Existen estudios que vinculan al Índice de Desarrollo Humano (IDH) con el Producto Interno Bruto (PIB) por países. El objeto de estudio de este proyecto es estudiar la relación entre el Índice de Desarrollo Humano (IDH) y el Ingreso Nacional Bruto per cápita, la tasa de inflación, la tasa de desempleo y la esperanza de vida al nacer, mediante un coeficiente de correlación adecuado. Este proyecto pretende entender si existe una asociación significativa entre el desarrollo humano y las condiciones macroeconómicas en cada continente. Para determinar las relaciones se utilizaron los coeficientes de correlación de Pearson, Spearman y Pearson Winsorizado. Y se calcularon intervalos de confianza mediante los métodos delta y bootstrap para verificar la hipótesis nula de independencia de los coeficientes. Se observó que: el PIB per cápita mantiene una correlación positiva y fuerte con el IDH; la EVN cuenta con una relación similar con el IDH, más pierde fuerza en África y América; el desempleo muestra variaciones significativas en magnitud, alcanzando una relación inversa en Asia y Europa; y el deflactor del PIB mantiene una relación inversa y con grandes variaciones en todos los continentes, principalmente en Europa. Por su parte, el desempleo y el deflactor del PIB no rechazan la hipótesis de independencia, por falta de información. De manera que Europa y Asia son los continentes que cuentan con las relaciones más fuertes, más también presentan la incongruencia de tener la relación inversa del IDH con el desempleo. Seguidamente están África y América, donde el PNB per cápita resalta como el más relacionado fuertemente con el IDH.

**PALABRAS CLAVE:** Indicadores económicos, coeficiente de correlación, intervalos de confianza.

## INTRODUCCIÓN

Inicialmente, el desarrollo de un país se medía según su Producto Interno Bruto per cápita, sin embargo, más adelante se concluiría que esta medida no era suficiente para conocer el desarrollo social, cultural, económico, educativo y salubre de una nación, como lo explica Chowdhury (1991). Desde su punto de vista, el Índice de Desarrollo Humano era un indicador más riguroso, sin embargo, este indicador aún contaba con deficiencias. Uno de los mayores desafíos es que al querer medir el desarrollo, las variables consideradas dependen del criterio de desarrollo del analista. Otro aspecto es el requerimiento de justificar el peso ponderado que se le otorgaría a cada variable, considerando la posible existencia de componentes correlacionados, de manera que su aporte al indicador sería inútil. Por último, algunos países

---

<sup>1</sup> Estudiante de la Escuela de Matemática de la Universidad de Costa Rica

<sup>2</sup> Estudiante de la Escuela de Matemática de la Universidad de Costa Rica

<sup>3</sup> Estudiante de la Escuela de Matemática de la Universidad de Costa Rica

con similitud relativa en algunos indicadores pueden presentar variaciones significativas en los indicadores que componen el IDH.

En 2015, Yakunina y Bychkov realizaron un estudio en el cual se correlacionó el Índice de Desarrollo Humano (IDH) de un país con la educación, la esperanza de vida al nacer (EVN) y el Producto Nacional Bruto (PNB) per cápita del mismo. Este estudio mostró que existe una relación alta entre el IDH y dichas variables. Por su parte, otro estudio realizado por Deb (2015), en el cual se consideran 4 años (1990, 2000, 2010 y 2013) y 140 países, los cuales fueron separados según sus niveles de ingreso en bajos, medios y altos, permitió observar una alta correlación, entre el IDH y el PIB, en los países de bajos ingresos, más esta correlación disminuye conforme aumentan los ingresos de los países. Lo cual hizo que surgiera una idea: en vez de agruparlos por ingresos y años, se agruparon por regiones en un año específico. Esto presentaría una variación similar tanto del IDH como de la correlación con otras variables. E incluso se puede expandir a nivel de continentes, permitiendo obtener más información sobre las realidades de cada continente. El presente trabajo busca estudiar la correlación entre el IDH y los datos del PNB per cápita, la inflación, el desempleo y la esperanza de vida al nacer, para los continentes de Asia, América, Europa y África durante el 2020, desde un enfoque estadístico.

Para ello se requiere una selección adecuada de coeficientes de correlación para las variables en cuestión, para lo cual Chen y Popovich (2002) brindan una guía para la escogencia de índices de correlación según el tipo de variables con las que se cuente. Para variables continuas se sugiere utilizar la  $r$  de Pearson. Dado que este coeficiente requiere de la normalidad de los datos, entonces es relevante considerar métodos de verificación de esta característica. En esta corriente, Shapiro y Wilk (1965) proponen un método sencillo para la evaluación de normalidad.

Sin embargo, los datos no suelen seguir una distribución normal. Para obtener las correlaciones se estudiarán tres coeficientes distintos: el coeficiente de Pearson, de Spearman y de Pearson Winsorizado. Estos coeficientes mostraron mantener buenos resultados en discrecionalidad de errores de tipo I, manteniendo una tasa del 5%, independientemente del tamaño de la muestra, según Temizhan et al. (2022). No obstante, es importante tener presente que el mismo estudio indicó que los resultados de poder estadístico de los coeficientes de Spearman y Pearson son influenciados negativamente cuando se incumple la normalidad, para lo cual se realizará la prueba Shapiro-Wilk con el fin de verificar la normalidad de los datos.

Ahora bien, para Li (2022) al implementar los coeficientes de correlación es importante estudiar el error muestral que rodea al resultado, esto se realiza mediante los intervalos de confianza (IC) que pueden dar información respecto a la hipótesis nula de los coeficientes de correlación. Estos intervalos se calculan de manera distinta según el coeficiente de correlación, para el coeficiente de Pearson y de Spearman se utiliza el método delta para estabilizar la varianza de dichos coeficientes mediante la  $z$  de Fisher, como lo indican Van Der Vaart (2000) y Fieller et al. (1957). Por su parte, Li (2022) denota que el IC de Pearson Winsorizado se calcula a través de bootstrapping dado que es lo más apropiado para estimadores robustos. Por lo anterior, se emplea el método Bootstrap del percentil (BPI) debido a que los extremos del intervalo dependen de los rangos de las correlaciones del bootstrap y, por ello, el BPI es más robusto ante distribuciones no normales, según Li (2022).

## METODOLOGÍA

Los estadísticos computados en esta sección se obtuvieron a través de R en su versión 4.3.1. Los datos sobre el Índice de Desarrollo Humano, Esperanza de Vida al Nacer, Desempleo total (% del total de la fuerza laboral), Producto Nacional Bruto per cápita y el Deflactor de PIB, indicadores que se reportan de manera anual, fueron obtenidos del Banco Mundial (2023) y Roser (2014). Cada una de estas variables son usadas para medir su correlación con el IDH. En particular, se eligen los datos del año 2020 para los países de los continentes de África, América, Asia y Europa. Contabilizando un total de 198 países, de los cuales 54 son africanos, 50 son de Asia, 48 de Europa y 46 americanos. No obstante, como la base de datos es el resultado de la unión de 5 bases distintas, la cantidad de observaciones por indicador difiere entre sí. A continuación, se detallan el total de datos faltantes por continente:

**Cuadro 1**

*Cantidad de datos faltantes por continente*

Indicadores	Datos faltantes			
	África	América	Asia	Europa
Índice de desarrollo humano	1	11	2	6
Esperanza de vida al nacer	1	11	2	6
Desempleo total	1	13	0	8
PNB per cápita	2	6	3	5
Deflactor del PIB	2	3	2	3

Fuente: Elaboración propia con datos del Banco Mundial y Roser

Al analizar las distribuciones y los valores anómalos de los datos, considerando los datos en escala ordinal y logarítmica, se observan variaciones en las distribuciones de los datos según el continente y la variable en cuestión. Por lo que no se logra determinar una distribución en particular. En cuanto los valores anómalos, estos se detectan en escala ordinaria en los cuatro continentes, mientras en escala logarítmica en Asia y América.

Asimismo, con el uso de gráficos de dispersión del paquete ggplot2 se identificó la relación inicial de los datos, aplicando nuevamente las dos escalas antes mencionadas. En la escala ordinal se obtuvo una relación lineal y positiva entre las variables del IDH y la esperanza de vida al nacer; mientras que para el IDH y el desempleo total, a simple vista, la relación parece ser casi inexistente y posee algunos valores atípicos; la relación entre el PNB per cápita y el IDH muestra ser creciente y curva con pocos valores atípicos; y en el caso del deflactor del PIB, también es lineal pero esta no parece ser positiva, ni negativa, contando con valores anómalos que destacan sobre los de las demás relaciones.

Al cambiar de escala tanto la esperanza de vida al nacer como el desempleo total mantienen un comportamiento similar al descrito anteriormente. El PNB per cápita todavía posee cierta curvatura, pero es mucho menor a la escala ordinal; y, por último, el deflactor del PIB parece tener cierta relación lineal y conserva varios valores anómalos significativamente mayores a los demás.

### Coefficiente de Correlación de Pearson

Kutner et al. (2005) definen dicho coeficiente para una muestra de datos como:

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (x_i - \bar{x})^2}}$$

donde  $\bar{x}$  y  $\bar{y}$  representan la media muestral de las variables. Dicho coeficiente da como resultado un valor numérico entre -1 y 1, lo que indica la dirección de la relación que mantienen dichas variables. La hipótesis nula de Pearson dice que las variables son independientes, es decir que  $r_{x,y} = 0$ . Al realizarse el cálculo del coeficiente se puede rechazar dicha hipótesis o aceptarla. De acuerdo con Chen y Popovich (2002) las hipótesis del coeficiente de correlación de Pearson son: linealidad y normalidad. La primera se refiere a que el coeficiente es adecuado únicamente para variables continuas que mantengan una relación lineal. Además, entre las presunciones del coeficiente se asume que las variables mantienen una distribución normal bivariada.

### **Coefficiente de correlación de Spearman**

Chen y Popovich (2002) argumentan que este coeficiente explicaría una relación entre dos variables que pueden ser tanto continuas como discretas, basándose en una comparación de rangos. Esto quiere decir que se ordenan las observaciones de mayor a menor y a cada una se le asigna un rango, donde la menor observación recibe el rango 1. Este ordenamiento de las variables permite deshacerse de la hipótesis de normalidad y de relación lineal entre las variables. Luego se calcula el cuadrado de las diferencias de los rangos de cada conjunto de variables y finalmente se calcula el estadístico de la siguiente manera:

$$r_{rank} = 1 - \frac{\sum d^2}{n^3 - n} \quad \text{donde } 0 \leq r_{rank} \leq 1$$

Donde  $d$  es la diferencia entre  $X$  y  $Y$ , y  $n$  es el número de observaciones de la muestra. En cuanto a su interpretación,  $r_{rank}$  se interpreta de igual manera que el coeficiente de correlación de Pearson. Es decir que en caso de que  $r_{rank} = 0$  se acepta la hipótesis nula y se dice que las variables no guardan ninguna relación. En caso de que sea contrario se acepta la hipótesis alternativa donde se explica que  $X$  y  $Y$  estarían correlacionadas.

### **Coefficiente de Pearson Winsorizado**

Temizhan et al. (2022) explican el cálculo de esta medida como el coeficiente de correlación de Pearson pero aplicado a datos winsorizados. Según Tukey (1962) winsorizar los datos consiste en ordenar los datos de mayor a menor, seleccionar aquellos valores que se consideran anómalos y reemplazarlos por el valor más cercano que permanezca dentro del rango de valores “normales” para el resto de los datos. A estos datos “normales” se les conoce como valores de corte y suelen tomarse umbrales como los percentiles, en particular, lo tradicional es utilizar el percentil 5 y el percentil 95.

### **Prueba de Shapiro - Wilk**

Esta es una prueba de normalidad para un conjunto de datos, según Shapiro y Wilk (1965) el estadístico se obtiene tras “dividir el cuadrado de combinación lineal apropiada de del estadístico de orden de la muestra entre la estimación simétrica habitual”. La justificación tras dicho estadístico recae en que este cociente es invariante ante la escala y el origen. De esta manera se plantea el estadístico  $W$ :

$$W = \frac{(\sum_{i=0}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

donde  $x_{(i)}$  es la  $i$ -ésima variable de la muestra,  $\bar{x}$  es el promedio de muestra y  $a_i$  son los coeficientes dados por:  $(a_1, \dots, a_n) = \frac{m^T V^{-1}}{C}$ . En esta ecuación  $C$  es un vector tal que:  $C =$

$\|V^{-1}m\| = (m^T V^{-1}V^{-1}m)^{1/2}$  y  $m = (m_1, \dots, m_n)^T$  es el vector de los valores esperados del estadístico de orden de las variables aleatorias independientes e idénticamente distribuidas y además  $V$  es la matriz de covarianza de estadísticos de orden normal. Donde  $W \in [0,1]$ , se rechaza la hipótesis que los datos poseen distribución normal (hipótesis nula) si dicho valor es más pequeño que el nivel de significancia  $\alpha = 0,05$ .

### Cálculo de intervalos de confianza mediante el método delta

Según Van der Vaart (2000), una sucesión de coeficientes de correlación de Pearson, convergen a la distribución normal de media 0 y varianza  $(1 - \rho)^2$  cuando  $n \rightarrow \infty$ . Al tomar la función diferenciable:  $\phi(\theta) = \arctan(\rho)$ . Por el método delta se puede asegurar que

$$\sqrt{n}(\arctan(r_n) - \arctan(\phi(\rho))) \rightarrow \phi'(\theta)N(0, (1 - \rho^2)^2) = N(0,1)$$

### Estimación de intervalos por bootstrapping basado en el método percentil

Flores (2005) describe el método Bootstrap como un proceso de varios pasos: Primero, con una muestra  $X_1, X_2, \dots, X_n$  se realiza una selección con reemplazo de la muestra original con el objetivo de conseguir una nueva muestra  $X^*_1, X^*_2, \dots, X^*_n$  del mismo tamaño. Segundo, “para la nueva muestra se calcula el valor del estadístico  $\hat{\theta}$  que se utiliza como estimador del parámetro  $\theta$ ” (Flores, 2005). Tercero, Flores describe que: “se repiten los dos pasos previos hasta obtener un número elevado de estimaciones de  $\theta$ ” (2005). Para finalizar, “se construye una distribución empírica del estadístico  $\hat{\theta}$ , este representa una aproximación a la verdadera distribución de  $\theta$ ” (Flores, 2005)

Para efectos de este trabajo, el parámetro de interés es el coeficiente de correlación entre las variables  $X, Y$  denótelos como  $r_{XY}$  y a su estimador como  $\hat{r}_{XY}$ . Entre los métodos que utilizan bootstrapping para calcular los intervalos de confianza están los intervalos basados en percentiles. El intervalo de confianza  $(1 - \alpha)\%$ , donde  $\alpha$  es el error de tipo I, para el coeficiente de correlación  $r_{XY}$  se calcula como:  $[r_{XY}(\alpha/2), r_{XY}(1 - \alpha/2)]$ . Para un número  $N$  de simulaciones, se toma como límite superior e inferior el  $N \cdot (1 - \alpha/2)$ -ésimo y el  $N \cdot \alpha/2$ -ésimo  $\hat{r}_{XY}$ , respectivamente, de la lista ordenada de todos los  $\hat{r}_{XY}$ . Es decir, depende de los rangos.

## RESULTADOS

Inicialmente se aplica la prueba de normalidad de datos para cada continente, cuyos resultados pueden ser observados en el Anexo A, Cuadro 4. Tomando un nivel de significancia del 3%, se concluye que en cada continente existe al menos una variable cuyos datos no son normales. En particular, el deflactor del PIB no tiene distribución normal para África, Asia y América. En Europa, el IDH no presenta distribución normal y en Asia el desempleo no cumple con dicha hipótesis. Por su parte, al obtener los coeficientes de correlación, Pearson, Spearman y Pearson Winsorizado, se agruparon estos por continente.

### Cuadro 2

#### *Coefficientes de correlación en África y América*

Variables	África			América		
	Pearson	Spearman	Winsorizado	Pearson	Spearman	Winsorizado
IDH-EVN	0,71	0,64	0,59	0,72	0,83	0,82
IDH-Desempleo	0,65	0,63	0,65	0,26	0,16	0,22

IDH-PNB per cápita	0,88	0,89	0,90	0,90	0,91	0,90
IDH-Deflactor del PIB	-0,13	-0,24	-0,19	-0,43	-0,29	-0,39

Nota. Todos los coeficientes anteriores fueron calculados con la función cor() y la función wincor() del paquete WRS2 en R

El caso del continente africano se puede observar en el Cuadro 2. La correlación más fuerte es entre el IDH y el PNB per cápita en el coeficiente de Pearson Winsorizado con un 0,90, a su vez, Pearson y Spearman también arrojaron resultados altos y positivos, concluyendo que las tres medidas de correlación coinciden en el comportamiento de la relación entre el IDH y el PNB per cápita. De manera similar, tanto el desempleo como la esperanza de vida al nacer demuestran mantener una correlación positiva con el IDH. En el caso del desempleo, se obtiene que la correlación más baja fue con el coeficiente de Spearman con un 0,63, mientras que los coeficientes de Pearson y de Pearson Winsorizado coinciden en 0,65. La correlación entre la esperanza de vida y el IDH, registra que el menor resultado fue con el coeficiente de correlación de Pearson Winsorizado con un 0,59 y el más alto fue Pearson con 0,71. Finalmente, el deflactor del PIB no aparenta mantener una correlación importante con el IDH, donde con valor absoluto cuenta con un mínimo de 0,13 con Pearson y un máximo de 0,24 bajo Spearman.

En América, se presenta una correlación positiva y fuerte entre el IDH con la esperanza de vida al nacer y el PNB per cápita bajo las tres medidas, como se observa en el Cuadro 2. Para el primero, es Pearson el que logra un mínimo con 0,72 y Spearman un máximo de 0,83, mientras para el segundo caso, tanto Pearson como Pearson Winsorizado indican una correlación de 0,90, y a su vez, Spearman indica una correlación de 0,91. Siendo destacable para el análisis de correlación que dos de las tres medidas de correlación coinciden de tal manera.

El resultado de la correlación entre el IDH y el desempleo es positiva pero débil, pues el resultado más alto fue bajo el coeficiente Pearson con un 0,26 y la menor correlación fue bajo Spearman con 0,16. Finalmente, para los tres coeficientes se obtiene una correlación negativa y débil, mas no despreciable, entre el IDH y el deflactor del PIB, donde la mayor correlación, en valor absoluto, se dio mediante Pearson con un 0,43 y la menor fue con Spearman con 0,29.

### Cuadro 3

#### *Coefficientes de correlación en Asia y Europa*

Variables	Asia			Europa		
	Pearson	Spearman	Winsorizado	Pearson	Spearman	Winsorizado
IDH-EVN	0,85	0,84	0,81	0,86	0,59	0,88
IDH-Desempleo	-0,13	-0,20	-0,23	-0,34	-0,30	-0,30
IDH-PNB per cápita	0,91	0,93	0,92	-0,97	0,96	0,97
IDH-Deflactor del PIB	-0,33	-0,50	-0,43	-0,52	-0,56	-0,65

Nota. Todos los coeficientes anteriores fueron calculados con la función cor() y la función wincor() del paquete WRS2 en R

El continente asiático muestra un comportamiento similar, en tres de las cuatro correlaciones, con respecto a África y América, mientras la correlación con el desempleo destaca, como se observa en el Cuadro 3. Manteniendo una correlación positiva fuerte entre el IDH y la esperanza de vida al nacer y el PNB per cápita. Para la correlación IDH-EVN, se obtiene el mayor resultado bajo el coeficiente de correlación de Pearson con un 0,85 y el menor resultado fue el de Pearson Winsorizado con 0,81. En lo referente al IDH y el PNB per cápita, se obtuvo la correlación más fuerte entre todas las variables. Además, se puede describir la correlación entre el IDH y el deflactor del PIB como inversa. Por su parte, si bien la correlación entre el IDH y el desempleo es negativa y débil.

Por último, se presenta el caso del viejo continente en el Cuadro 3, un comportamiento similar al visualizado en el caso asiático. Resaltando la correlación entre el IDH y el PNB per cápita, por ser la más fuerte en Europa, donde el mayor resultado es bajo Pearson y Pearson Winsorizado con un 0,97, mientras el más bajo es por Spearman con un 0,96. La relación entre el IDH y la esperanza de vida al nacer es positiva y la segunda más fuerte, con un máximo bajo Pearson Winsorizado al alcanzar el 0,88, mientras bajo Spearman se llega a un 0,85. Por otro lado, la correlación entre el IDH y el deflactor del PIB es inversa bajo las tres medidas de y correlación, donde con valor absoluto la mayor cifra es con el coeficiente de Pearson Winsorizado con un 0,65, siendo acotada por abajo por el coeficiente de correlación de Pearson con un 0,52. En cuanto al IDH y el desempleo, esta es la relación más débil de las cuatro y se mantiene negativa, al analizarlo con valor absoluto, el menor valor obtenido fueron los resultados de Spearman y Pearson Winsorizado con un 0,30 mientras el mayor valor registrado fue para Pearson con un 0,34.

Ante esto, se rechazan las hipótesis nula de independencia y se determina que existen distintos tipos de correlación entre el IDH y el resto de variables. De manera que la relación entre el IDH y el deflactor del PIB es inversa para todos los continentes, donde se puede observar una variación de la magnitud de esta relación, mientras que existe una relación fuerte tanto entre el IDH y la EVN, así como entre IDH y el PNB per cápita en los cuatro casos, lo cual calza en la literatura y es de esperarse, pues estos índices forman parte del cálculo del Índice de Desarrollo Humano. Así mismo, la relación entre el IDH y el desempleo presentó variaciones destacables, pues se mantiene positiva en dos de los cuatro continentes y negativa en el resto, siendo este último caso el que más resalta.

Por su parte, y como aspecto importante en la validación de los resultados previamente mencionados, se realizaron intervalos de confianza para cada continente bajo un porcentaje del 95% y bajo dos métodos: método delta y método bootstrap. Estos resultados se pueden observar en el Cuadro 5 del Anexo B.1 y en el Cuadro 6 del Anexo B.2, respectivamente.

Para el método delta africano, se trabajó con una muestra de 50 datos para los coeficientes de Spearman y Pearson, salvo para la correlación entre el deflactor del PIB y el IDH, donde se utilizó una muestra de 49 datos. Exceptuando el deflactor del PIB, ninguno de los intervalos de confianza contiene el 0, por lo que existe una fuerte evidencia para rechazar la hipótesis nula de ambos coeficientes. En América, bajo el método delta, se emplearon 29 datos para todas las variables bajo el coeficiente de correlación de Spearman y 28 bajo Pearson, salvo

por el deflactor del PIB, donde se emplearon 27. De los cuales, los únicos con los que no se podría rechazar la hipótesis nula sería con el desempleo y del deflactor del PIB para Spearman.

Sobre el caso delta en el continente asiático, se empleó un tamaño de la muestra de 46 países. Las correlaciones entre los indicadores del PNB per cápita, el deflactor del PIB y la esperanza de vida al nacer no contienen al 0 en sus intervalos de confianza, por lo que existe evidencia para rechazar la hipótesis nula para ambos coeficientes, caso opuesto a lo que sucede con el desempleo, donde no se puede rechazar con ninguno de los coeficientes. En el caso europeo, se utilizó una muestra de 39 países, resaltando 3 tipos de comportamientos: el referente al PNB per cápita y al de la EVN se mantiene positivo, mientras que el referente al desempleo mantiene una parte positiva y otra negativa (por lo que incluye al cero y no se puede rechazar la hipótesis nula). El deflactor del PIB mantiene intervalos totalmente negativos.

En el caso bootstrap se emplearon 1000 simulaciones con el fin de calcular el intervalo de confianza mediante percentiles. En África este proceso incluyó 50 países, salvo para el correspondiente con el deflactor de PIB bajo Pearson, donde se emplearon 49 países, obteniendo resultados similares a los del método delta, al no poder rechazar la hipótesis nula para el deflactor del PIB, más ahora en los tres coeficientes. En el caso de América, se emplearon 29 países con excepción del deflactor del PIB bajo Pearson, y manteniendo resultados similares a los del método delta, al no poder rechazar la hipótesis nula para el desempleo en los tres coeficientes y sin lograr rechazarlo para el deflactor bajo Spearman y Pearson Winsorizado.

Continuando con Asia, bajo el método bootstrap se observó un intervalo de confianza positivo para los referentes a la esperanza de vida al nacer y el PNB per cápita, totalmente negativo para el caso del deflactor del PIB y, donde incluye el cero en el caso del desempleo para los tres casos, lo cual resulta similar al caso delta. Sin embargo, al emplear 39 observaciones para Europa, no se logra rechazar la hipótesis nula en los casos del desempleo bajo Spearman y Pearson Winsorizado al existir el 0 en sus intervalos de confianza.

De manera general: África no puede rechazar la hipótesis nula en ninguno de los métodos para el caso del deflactor del PIB; América no rechaza la hipótesis nula para el desempleo con ningún coeficiente, ni el deflactor del PIB, bajo Spearman y Pearson Winsorizado, en ambos métodos. Igualmente, en Asia no se rechaza la hipótesis nula del desempleo con ningún coeficiente. No obstante, la hipótesis nula es rechazada con el coeficiente de Pearson entre el IDH y el desempleo para el continente europeo. Se concluye que existe poca evidencia para rechazar la hipótesis de independencia para los casos correspondientes al deflactor del PIB y el desempleo. Sin embargo, se logra rechazar dicha hipótesis para la esperanza de vida al nacer y el PNB per cápita. Por su parte, cada una de las magnitudes de las correlaciones entre el IDH y las variables, se encuentran dentro de los intervalos de confianza calculados.

Teniendo presente dichas verificaciones, se pueden analizar las diferencias en las magnitudes de las correlaciones según los continentes. En el caso de la EVN, se encontró cierta similitud entre Europa y Asia al presentarse de manera fuerte, más esta diverge para América y África, donde pierde fuerza y se genera una brecha de magnitud considerable; para el desempleo, existen grandes diferencias, África y América mantienen una relación positiva, más

con una importante brecha de magnitud, mientras Asia y Europa presentan una correlación negativa y una brecha leve entre ellas; el PNB per cápita es el que mayor similitud presenta entre los cuatro continentes; y por último, el deflactor del PIB presenta grandes variaciones en sus magnitudes.

## CONCLUSIONES

En este trabajo se buscaba determinar la existencia de diferencias en las magnitudes en las correlaciones, entre el IDH y la esperanza de vida al nacer, el desempleo, el deflactor del PIB y el PNB per cápita, según se comparaban entre América, África, Asia y Europa. Pues en caso de existir diferencias en dichas magnitudes, se puede evidenciar cómo cada una de dichas variables se relaciona con el IDH y, de esta manera, poder exponer características de cada continente. Para poder comparar las magnitudes de las correlaciones, se planteó el cálculo de tres coeficientes de correlación: de Pearson, de Spearman y de Pearson Winsorizado. El primer reto que se presentó fue el requerimiento de normalidad en los datos recolectados para los coeficientes de Pearson, pues nuestros datos no mostraron dicha normalidad al aplicarse la prueba Shapiro-Wilk.

Una vez recopilados las magnitudes de las correlaciones de Pearson, Spearman y Pearson Winsorizado para cada uno de los continentes se encontró que todos de los coeficientes rechaza la hipótesis nula, es decir, que sí mantienen una relación significativa, más para corroborar este resultado se calcularon los intervalos de confianza de estos mismos coeficientes de correlación, siendo empleados dos métodos: el delta y bootstrap. Ambos métodos dieron resultados similares entre ellos, donde concluyen que no se puede garantizar una relación lineal significativa entre el IDH y el desempleo, así como entre el IDH y el deflactor del PIB, más en el caso de la esperanza de vida con el IDH y el PNB per cápita con el IDH sí se puede rechazar esta hipótesis nula, es decir que sí existe evidencia suficiente que relacione cada una de esas variables con el IDH, lo cual mantiene sentido pues son variables que se emplean en el cálculo del IDH.

Al corroborarse el comportamiento de la hipótesis nula en cada una de las variables con respecto al IDH, se procedió a analizar los resultados obtenidos de los coeficientes de correlación, concluyendo que sí existen variaciones en las magnitudes en las correlaciones entre los continentes americano, europeo, asiático y africano. Lo cual permite comprender de una manera más amplia cómo se relaciona cada variable en cada continente, siendo el PNB per cápita el más influyente en el IDH en todos los continentes y el que menos varía en cada uno de los mismos, mientras la esperanza de vida influye fuertemente en Asia y Europa, más pierde relevancia en el IDH de América y pierde aún más relevancia en el IDH africano. El desempleo mostró mantener una relación positiva en África y una relación débil en América, así como una relación inversa para Asia y Europa. Y por último, el deflactor del PIB es inversamente relacionado con el IDH en todos los continentes, con mayor fuerza en Europa, seguido de Asia, América y África.

Se evidencia la importancia del PNB per cápita en la medición del IDH al existir una fuerte correlación en todos los continentes. Por su parte, no se puede concluir lo mismo para la EVN en países africanos al no mantener una relación fuerte con el IDH. Si bien estos índices

presentaron grandes variaciones en sus magnitudes en cada continente, se puede concretar que el deflactor del PIB funciona inverso al IDH.

### **AGRADECIMIENTOS**

Este grupo de trabajo desea externar su más profundo agradecimiento al profesor Maikol Solís Chacón, quien brindó su apoyo inconmensurable a lo largo de la elaboración del presente proyecto y nos introdujo al mundo de la estadística. Su pasión por la enseñanza y trayectoria académica fueron pilares fundamentales del éxito de este trabajo. Atesoraremos estos conocimientos durante nuestras trayectorias profesionales. También deseamos destacar el valioso aporte de Joshua Cervantes Artavia, quien nos nutrió con su retroalimentación constante. Reconocemos que sus aportes ayudaron a elevar la calidad de nuestro trabajo.

### **BIBLIOGRAFÍA**

Banco Mundial (2023) GNI per capita, Atlas method (current US\$)

Chen, P. Y., & Popovich, P. M. (2002). *Correlation: Parametric and nonparametric measures* (No. 139). Sage.

Chowdhury, O. H. (1991). Human Development Index : A Critique. *The Bangladesh Development Studies*, 19(3), 125–127. JSTOR.  
<https://www.jstor.org/stable/40795411>

Deb, S. (2015). Gap Between GDP and HDI: Are the Rich Country Experiences Different from the Poor?

Fieller, E. C., Hartley, H. O., & Pearson, E. S. (1957). Tests for Rank Correlation Coefficients. I. *Biometrika*, 44(3/4), 470. <https://doi.org/10.2307/2332878>

Flores, J. G. (2005). Aplicación del método bootstrap al contraste de hipótesis en la investigación educativa. *Revista de Educación*, 251–265.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. McGraw-hill.

- Li, J. C. (2022). Bootstrap confidence intervals for 11 robust correlations in the presence of outliers and leverage observations. *Methodology*, 18(2), 99–125. <https://doi.org/10.5964/meth.8467>
- R Core Team (2023). R: A Language and Environment for Statistical Computing. R. Foundation for Statistical Computing, Vienna, Austria.
- Roser, M. (2014). Human development index (HDI). *Our world in Data*.
- Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591–611. <https://doi.org/10.2307/2333709>
- Temizhan, E., Mirtagioglu, H., & Mendes, M. (2022). Which Correlation Coefficient Should Be Used for Investigating Relations between Quantitative Variables? *American Scientific Research Journal for Engineering, Technology, and Sciences*, 85(1), 265–277. [https://asrjetsjournal.org/index.php/American\\_Scientific\\_Journal/article/view/7326](https://asrjetsjournal.org/index.php/American_Scientific_Journal/article/view/7326)
- Tukey, J. W. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics*, 33(1), 1–67. <https://doi.org/10.1214/aoms/1177704711>
- Vaart, A. W. Van Der. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.
- Yakunina, R. P., & Bychkov, G. A. (2015). Correlation Analysis of the Components of the Human Development Index Across Countries. *Procedia Economics and Finance*, 24, 766–771. ScienceDirect. [https://doi.org/10.1016/s2212-5671\(15\)00692-9](https://doi.org/10.1016/s2212-5671(15)00692-9)

## Anexos

**Anexo A:** Resultados de la prueba Shapiro-Wilk:

A continuación, se presentan los resultados de la prueba Shapiro-Wilk empleada para medir la normalidad de los datos. Esto se realizó mediante la función `shapiro.test(data)` en R.

**Cuadro 4**  
*Resultados de la prueba Shapiro-Wilk*

Continente	Indicadores	Variables	
		W	p-valor
África	IDH	0,977	0,434
	EVN	0,951	0,038
	Desempleo	0,964	0,137
	PNB per cápita	0,973	0,316
	Deflactor del PIB	0,727	1,327e-10
Asia	IDH	0,969	0,261
	EVN	0,982	0,682
	Desempleo	0,877	0,877
	PNB per cápita	0,949	0,046
	Deflactor del PIB	0,727	6.478e-08
América	IDH	0,983	0,909
	EVN	0,957	0,299
	Desempleo	0,979	0,841
	PNB per cápita	0,965	0,448
	Deflactor del PIB	0,845	0,0009
Europa	IDH	0,927	0,015
	EVN	0,908	0,004
	Desempleo	0,953	0,106
	PNB per cápita	0,954	0,115
	Deflactor del PIB	0,954	0,115

Fuente: Elaboración propia

## Anexo B: Intervalos de confianza

### Anexo B.1: Método Delta

Para calcular los intervalos de confianza a través del método delta se usaron funciones de elaboración propia en R.

**Cuadro 5**  
*Intervalos de confianza al 95% por el método delta*

Continente	Indicadores	Pearson		Spearman	
		[Inf, Sup]	[Inf, Sup]	[Inf, Sup]	[Inf, Sup]
África	EVN	0,541	0,827	0,427	0,780
	Desempleo	0,449	0,784	0,414	0,773
	PNB per cápita	0,796	0,930	0,809	0,937
	Deflactor del PIB	-0,393	0,161	-0,488	0,055
América	EVN	0,480	0,864	0,665	0,921
	Desempleo	-0,122	0,579	-0,229	0,507
	PNB per cápita	0,800	0,955	0,804	0,956
	Deflactor del PIB	-0,695	-0,058	-0,601	0,096
Asia	EVN	0,744	0,915	0,721	0,910
	Desempleo	-0,409	0,162	-0,471	0,104
	PNB per cápita	0,846	0,951	0,871	0,961
	Deflactor del PIB	-0,563	-0,039	-0,697	-0,240
Europa	EVN	0,750	0,926	0,717	0,918
	Desempleo	-0,594	-0,031	-0,572	0,022

PNB per cápita	0,942	0,984	0,930	0,981
Deflactor del PIB	-0,717	-0,243	-0,750	-0,291

Fuente: Elaboración propia

### Anexo B.2: Método Bootstrap basado en percentiles

Para estimar estos intervalos de confianza se utilizó el la función `bootcor_test()` del paquete TOSTER en R

Cuadro 6

*Intervalos de confianza al 95% por el método Bootstrap*

Continente	Indicadores	Pearson		Spearman		Winsorizado	
		[Inf, Sup]	[Inf, Sup]	[Inf, Sup]	[Inf, Sup]	[Inf, Sup]	[Inf, Sup]
África	EVN	0,559	0,830	0,362	0,787	0,263	0,811
	Desempleo	0,464	0,794	0,416	0,773	0,410	0,798
	PNB per cápita	0,817	0,925	0,809	0,928	0,789	0,951
	Deflactor del PIB	-0,401	0,067	-0,490	0,034	-0,471	0,092
América	EVN	0,554	0,855	0,6598	0,9178	0,609	0,921
	Desempleo	-0,051	0,565	-0,206	0,555	-0,123	0,560
	PNB per cápita	0,817	0,953	0,788	0,952	0,782	0,961
	Deflactor del PIB	-0,678	-0,156	-0,636	0,048	-0,713	0,043
Asia	EVN	0,732	0,918	0,699	0,922	0,653	0,907
	Desempleo	-0,398	0,174	-0,472	0,087	-0,500	0,108
	PNB per cápita	0,872	0,948	0,862	0,957	0,837	0,960
	Deflactor del PIB	-0,535	-0,079	-0,703	-0,243	-0,687	-0,096
Europa	EVN	0,789	0,922	0,732	0,909	0,787	0,938
	Desempleo	-0,603	-0,031	-0,592	0,020	-0,618	0,034
	PNB per cápita	0,950	0,983	0,918	0,980	0,921	0,986
	Deflactor del PIB	-0,761	-0,358	-0,764	-0,274	-0,827	-0,262

Fuente: Elaboración propia