

SERIE: ¿CÓMO SE COMPORTA LA INVESTIGACIÓN EDUCATIVA EN CUBA, CINCO AÑOS DESPUÉS?

NÚMERO No.11: LOS GRANDES ALIADOS DE LA INVESTIGACIÓN CUANTITATIVA EN EL ANÁLISIS DE LOS DATOS: LA ESTADÍSTICA MULTIVARIANTE Y EL ENTORNO 'R-PROJECT' (1ra. Parte)

Dr. Cs. Paul A. Torres Fernández
Profesor e Investigador Titular, ICCP

orcid.org/0000-0002-7862-2737

[linkedin.com/in/paul-antonio-torres-fernandez-4684048a](https://www.linkedin.com/in/paul-antonio-torres-fernandez-4684048a)

<https://www.facebook.com/ptorresfernandez>

<https://twitter.com/paintelectual>

<https://paulantoniotorresfernandez.blogspot.com>



3.07.2022; 15:10hs.

Introducción

En los últimos dos posts de la presente Serie sobre Metodología de la Investigación Científica [y cuyo propósito fundamental ha sido –recuérdese– llevar a efecto una nueva *introducción de resultados* de mi tesis de Doctor en Ciencias (Torres, 2016)] hemos estado reflexionando acerca del papel de la Estadística Aplicada en la investigación educativa. En el Número 9 abordamos las posibilidades de la Estadística Descriptiva y en el Número 10 –el post anterior– las acotadas potencialidades de la Estadística Inferencial ‘clásica’.

En relación con la primera –la Estadística Descriptiva– valoramos su utilidad tanto para investigaciones del *enfoque cuantitativo*, como del *cuantitativo*; y lo mismo para estudios de naturaleza *explicativa*, como *correlacional*, como incluso *descriptiva*. La Estadística Inferencial sí queda restringida a investigaciones del *enfoque cuantitativo*, esencialmente de naturaleza *explicativa*, aunque pudiera emplearse también en las de carácter *correlacional*.

Para poder penetrar en el complejo *mundo* de la Estadística Inferencial –aún *clásica*– tuvimos que introducir –en el post anterior– al menos cuatro nociones básicas: (i) *inferencia estadística*, (ii) *nivel de confianza* (para la *inferencia por intervalo*), (iii) *pruebas de hipótesis*, y (iv) *diferencias significativas*. Puesto que las dos primeras intervienen en el proceso de las *pruebas de hipótesis* y el *estadístico de prueba* suele

suponer una diferencia entre *parámetros*, he tratado de ilustrar su presencia en una única representación gráfica, con el propósito de fijarlas. Ya sabemos que en ellas cuatro subyace –inevitablemente– la noción de *probabilidad estadística*.

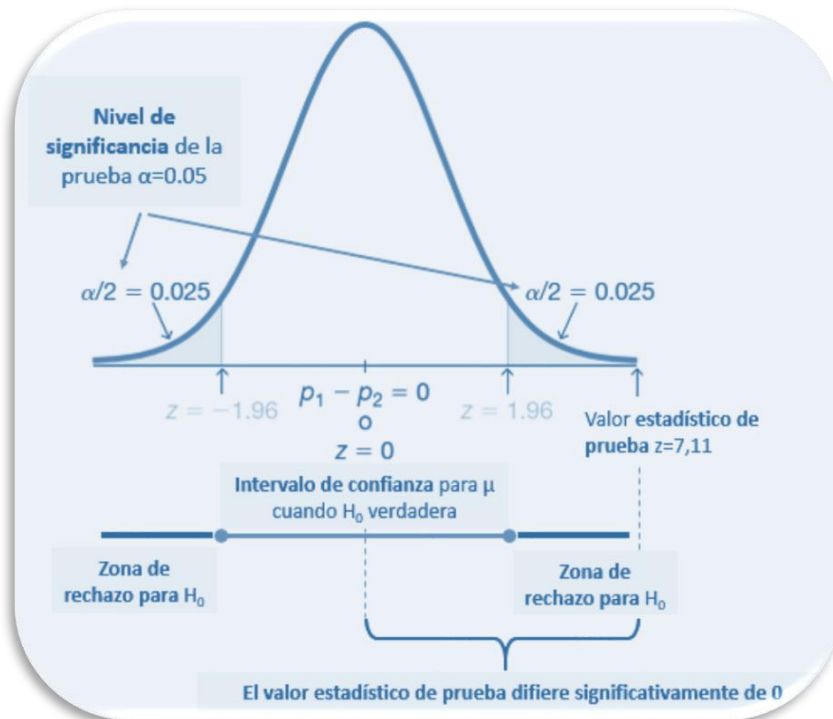


Figura No.2: Presencia de las cuatro nociones básicas del post No.10 en la representación de una prueba de hipótesis. [Elaboración propia a partir de (Triola, 2018, p.419)]

En la Figura No.2 aparece reflejado lo anteriormente señalado en un ejemplo concreto. A partir de conocerse las *proporciones muestrales* de un mismo atributo en dos *poblaciones* diferentes, se asume como *hipótesis nula* que ambas *proporciones* son iguales ($H_0: p_1=p_2$, o lo que es lo mismo: $H_0: p_1-p_2=0$, o sea la distancia es cero).

Dado que lo que se tiene son los cálculos de las *proporciones muestrales* (además de los *tamaños* de cada una de las dos *muestras*, y del *número de éxitos* [o presencia del atributo de interés] en cada una de ellas), y la *hipótesis de investigación* se debe formular en relación con las *poblaciones* y no con las *muestras* (pues en ellas podrían calcularse directamente las *proporciones* y compararse sin mayor esfuerzo), entonces hay que considerar una (i) *inferencia estadística* subyacente a la realización de la (iii) *prueba de hipótesis*.

Para una *buena* precisión al *inferir*, fijaremos el (ii) *nivel de confianza* del *intervalo* a construir (aquí para el *estadígrafo de prueba*, como ya vimos en el post anterior); en nuestro ejemplo, el *nivel de confianza* asumido es de un 95%, con lo que, –al mismo tiempo– estaríamos estableciendo un *nivel de significancia de la prueba* de $\alpha=0,5$ y las *zonas de rechazo* de H_0 en ambas *colas* de la *distribución teórica* que corresponde a dicho *estadígrafo de prueba* (aquí, la *curva normal estándar*).

De obligarnos el *valor estadístico de prueba* calculado a *rechazar* a H_0 , por situarse en una de las *zonas de rechazo* (y por tanto, fuera del *intervalo de confianza* antes descrito), entonces podríamos afirmar (con una *probabilidad de error al rechazar*

menor a 0,05) que las *proporciones* de ambas *poblaciones* son distintas, tanto que (iv) *difieren significativamente*.

Pero, bueno, recuerden que no hay que alarmarse. Si se han comprendido esas cuatro nociones básicas (y dos o tres conceptos subordinados más), entonces podemos olvidarnos de las fórmulas y de los cálculos numéricos con ellas, así como de las tablas de valores de las distribuciones teóricas requeridas, pues ahí están los actuales paquetes estadísticos capaces de ocuparse de ello.

A nosotros –los investigadores– nos quedarían solo las tareas de proporcionarles los datos, elegir el recurso estadístico conveniente e interpretar las salidas que nos proporcione el software. Es decir, mucho menos trabajo que el requerido en los años '70, '80 y '90.

Por otra parte, recordarán que terminamos el post anterior valorando críticamente la incongruencia existente entre el número limitado de *variables* que intervienen en una *prueba de hipótesis* (como mismo en una *inferencia por intervalo*) y el marcado *carácter multivariante* del proceso educativo, punto de referencia de cualquier investigación educativa. Para ilustrar esa naturaleza compleja del *objeto de estudio* general de la actividad científica de nuestra comunidad, acudíamos a Mafalda y al sistema de influencias que intervienen en su formación personalógica.

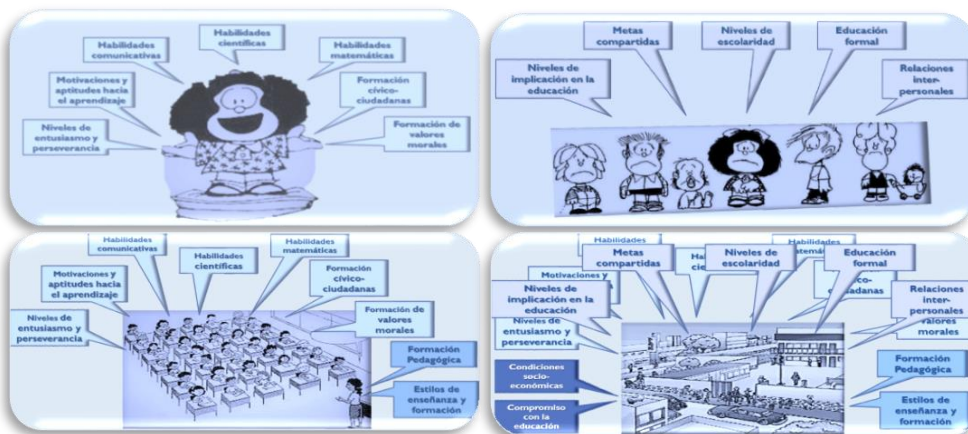


Figura No.3: Representación del carácter dinámico, complejo y multifactorial del proceso educativo, a través del personaje de Mafalda. (Elaboración propia)

En este último posts dedicado al tema del análisis de los datos de investigaciones educativas con el auxilio de la Estadística, les hablaré de un tercer grupo de *recursos estadísticos* de gran valía y actualidad. Pero no hay que alarmarse, pues difícilmente puedan trabajarse con lápiz y papel, ni con calculadoras tradicionales; pues los ingenieros informáticos y los llamados ‘científicos de los datos’ (*data scientific*) se han ocupado de facilitarnos mucho esos procesos, a tal grado que el conocimiento estadístico necesario es realmente mínimo, aunque ello les resulte paradójico.

Técnicas multivariantes de notable utilidad para la investigación educativa: ‘regresiones’, ‘análisis de la varianza’ y ‘análisis canónico’

Es literalmente imposible intentar explicar aquí las numerosas técnicas de análisis estadístico que componen este tercer grupo avanzado de recursos auxiliares de la investigación educativa. En la Figura No.4 he presentado una buena parte de ellas, pero me limitaré aquí a solo tres: las ‘regresiones’, los ‘análisis de la varianza’ y el

‘análisis canónico’. La imagen la hemos tomado de un resumen introductorio de la obra (Pérez, 2004), que asume como soporte digital a SPSS.

En cambio, una ilustración gráfica de casi todas esas técnicas aparece expuesta en la portada del presente post, perteneciente a la interface *RWinzard* (Universidad de Vigo, 2014). A ella he añadido el ícono de *RStudio*, la interface de *R-Project* más popular en la actualidad.

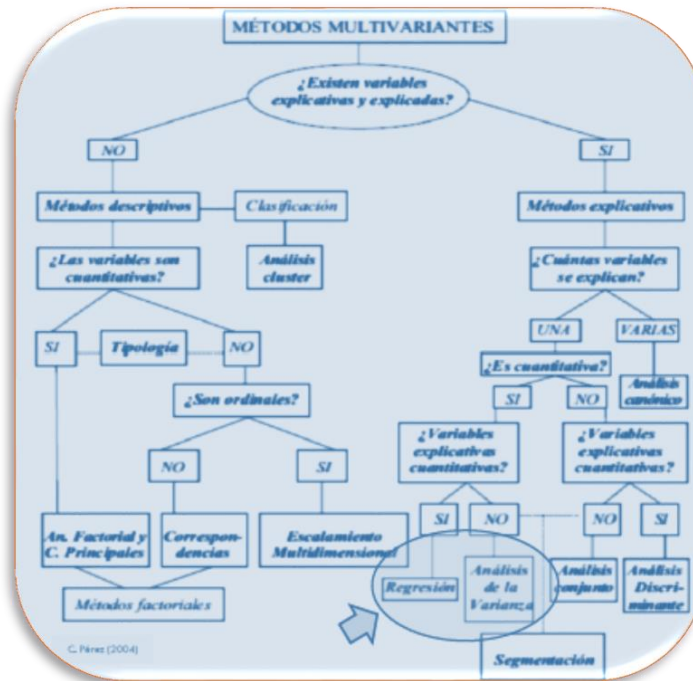


Figura No.4: Sistematización de los principales métodos estadísticos multivariantes y de sus condiciones de elección. (Pérez, 2004, p.3)

Como he resaltado en la Figura anterior, las técnicas de ‘regresión’ y de ‘análisis de la varianza’ se utilizan cuando se requiere explicar la *variabilidad* de una *variable numérica*, a partir de una o más *variables explicativas numérica(s) o categórica(s)*, respectivamente. De modo que son muy útiles en *investigaciones explicativas* en las que se sigue el *enfoque cuantitativo* de investigación; lo mismo de *corte experimental* que de *naturaleza ex-post-facto* (‘después de sucedidos los hechos’, recuerden).

Por su parte, el ‘análisis canónico’ –ciertamente, menos conocido que las otras dos técnicas multivariantes anteriores– se debe emplear cuando se quiere explicar la *variabilidad*, no de una sino de varias *variables explicadas*; para cuyo caso se emplearían lo mismo *variables explicativas numéricas* que *categóricas*, dando así mayor flexibilidad a las condiciones de partida del análisis estadístico.

La gran potencia de todas estas técnicas –que ya vimos no tenía ni la *estimación puntual*, ni *por intervalo*, como tampoco las *pruebas de hipótesis*– es que proporcionan una *estimación del efecto* de las *variables explicativas* sobre las *explicadas*; es decir, una magnitud concreta del llamado ‘*impacto*’ de unas *variables previas* (de *contexto*, *entrada* o *proceso*) sobre otras de *salida* (o de *producto*, del decursar del *objeto de estudio* en la práctica).

Antes de avanzar con una caracterización más singular de cada uno de ellas, y ejemplificarlas en situaciones de investigación educativa, es muy importante hacer énfasis en la noción de *variabilidad* de los datos. Si realizan una mirada retrospectiva de lo estudiado hasta ahora en torno al tema de la Estadística Aplicada (la Estadística Descriptiva y la Estadística Inferencial

'clásica') allí nos preocupábamos por *describir y/o estimar* una pareja de *medidas de resumen estadístico* interrelacionadas: una de *tendencia central* (por ejemplo, \bar{X} o μ) junto a otra de *dispersión* (por ejemplo, s o σ); o una *proporción* de las *frecuencias* de ocurrencia de ciertos atributos (generalmente, dado en *por cientos*).

Pero aquí, al hacer énfasis en la *variabilidad* (o *dispersión de los datos*), se presupone el *axioma* de que '*lo que varía explica y lo que es constante, no*'.

Ahora bien, como ambos grupos de valores (*variables explicativas y explicadas*) cambian, si las primeras lo hacen relativamente poco y las segundas no, entonces la *capacidad explicativa* del *modelo estadístico* que las asocia suele ser muy baja.

Es decir, estas técnicas tratan de identificar –en una o entre varias– aquellas *variables explicativas* cuyos *rangos de valores* coinciden mejor con el de las *variables de salida* (de *producto*, o *explicadas*; como prefiera decirse).

Es por ello que en el inicio de todo *análisis multivariante* se deben hacer exploraciones previas del comportamiento de los datos (con recursos de la Estadística Descriptiva, esencialmente), de manera que se elijan aquellas *variables explicativas* que parezcan ser más '*potentes*' (porque pretender abarcar muchas de ellas a la vez tampoco es bueno, pues es mucho más difícil de explicar el *efecto combinado* que generan).

También las *pruebas de hipótesis* (de la Estadística Inferencial '*clásica*') son de ayuda en los *análisis multivariantes*, sobre todo para verificar ciertas condiciones obligatorias de los *modelos* correspondientes a cada técnica, además de para elegir el mejor de los *modelos estadísticos* que se pueden generar; ello, sobre la base del precepto de que '*ningún modelo estadístico es perfecto*', pues todos tienen un margen de error.

Dicho esto, pasaremos a describir y ejemplificar cada una de las tres técnicas multivariantes elegidas para este post.

La *regresión lineal* puede ser definida como "(...) *un modelo estadístico para estimar el efecto de una variable sobre otra. (...) Brinda la oportunidad de predecir las puntuaciones de una variable tomando las puntuaciones de la otra variable. Entre mayor sea la correlación entre las variables (covariación), mayor capacidad de predicción.*" (Hernández-Sampieri, Fernández & Baptista, 2010; p. 314).

En realidad, ella puede ser *simple* (cuando la *variable explicada* ['y'] es contrastada por una única *variable explicativa* ['x']), como también *múltiple* (que es cuando la *variable explicada* o *producto* ['y'] es contrastada por dos o más *variables explicativas* ['x_i']). Este último caso es muy útil para intentar *atrapar* el carácter marcadamente *multifactorial* del proceso educativo y de su contexto socio-económico.

Otro tipo de *regresión* de provecho para el ámbito educativo es la *regresión binomial*, que es en la que está sustentada la *Teoría de Respuesta al Ítem* (TRI), con ayuda de la cual se pueden diseñar y pilotear las llamadas '*pruebas objetivas*' de los '*estudios a gran escala*', como lo es el ERCE-2019 (OREALC-UNESCO-Santiago, 2021). Pero su estudio escapa a las posibilidades y propósitos fundamentales del presente post.

Retornando a la *regresión lineal*, nótese como en (Hernández-Sampieri, Fernández & Baptista, 2010) se resalta –con acierto– que posee no solo un *carácter explicativo* (o sea, de *estimación* de la magnitud de los *impactos* de las *variables explicativas* sobre la *explicada*), sino también otro *predictivo*, lo cual es igual de importante. Veamos un ejemplo de esas dos *potencialidades* de la *regresión lineal*.

En la Figura No.5 (debajo) se presenta una tabla donde se muestran los resultados de la aplicación de una *regresión poblacional* a cuatro *índices estadísticos* construidos con agrupaciones de *factores asociados a los logros del aprendizaje* en el recién concluido cuarto Estudio Regional Comparativo y explicativo (ERCE-2019), de la OREALC-UNESCO. Se trata de resultados específicos de Cuba.

Fuentes	Factores asociados	Sexto grado		
		Lectura	Matemática	Ciencias
Cuestionario de estudiantes de 6°	ORG_: Organización de la clase de... (media: 0, desviación estándar: 1)	12.73***	7.84***	18.44***
	AAE_: Apoyo al aprendizaje en... (media: 0, desviación estándar: 1)	13.49***	9.98***	18.15***
	AEMAT: Autoeficacia en Matemática (media: 0, desviación estándar: 1)	—	16.42***	—
	ISECF: Nivel socioeconómico de la familia (media: 0, desviación estándar: 1)	39.78***	19.25***	20.62***
Módulo Nacional	MN: Empleo de preguntas e impulsos didácticos en clases. (0-1)	32.74***	16.98	25.20

Figura No. 5: Estimaciones poblacionales de las efectos de los índices estadísticos de reactivos del cuestionario de estudiantes de 6° grado en el ERCE- 2019, correspondientes las clases, sobre los logros del aprendizaje (Torres et al., 2022; p. 68).

El primer *índice* (ORG) hace referencia a atributos de las clases bajo un *enfoque de enseñanza* del tipo *explicativo-ilustrativo* (Torres, 2013), y fue construido a partir de reactivos de los cuestionarios regionales de docentes y estudiantes.

El último *índice* (MN) corresponde a reactivos similares, pero del *módulo nacional cubano* (un grupo reducido de preguntas incorporadas a los cuestionarios regionales, con foco en un interés investigativo propio); en este otro caso se indaga sobre el conocimiento y empleo de atributos de una *enseñanza* del tipo *desarrolladora* (Torres, 2013). Los otros tres *índices* (AAE, AEMAT e ISECF) aparecen solo con fines comparativos.

Obsérvese que para las tres áreas curriculares evaluadas al cierre del nivel primario (Lectura, Matemática y Ciencias Naturales, 6° grado) la *variable predictora* 'ORG' tiene un *efecto positivo* y además con una alta *significación estadística* ($p < 0.001$) sobre los *logros del aprendizaje*; similares a los 'impactos' generados por 'AAE_'.

En cambio, los *efectos* de la *variable predictora* 'MN' sobre los *logros del aprendizaje* son mucho más altos, a la altura de la *variable no educativa* de mayor *capacidad predictiva* a escala mundial (aquí, representada por el *índice estadístico* 'ISECF'), y con resultados similares a los de la *habilidad socioemocional* 'AEMAT', un área donde los escolares cubanos obtuvieron muy buenos resultados en el ERCE-2019.

Ahora bien, estos resultados no solo son *explicativos* sino también *proyektivos*, porque los *modelos estadísticos de regresión poblacional* empleados *estiman* los valores de la *variable explicada* 'y' (aquí: '*logros del aprendizaje*') como *resultados esperados*; y no los *obtenidos* con la *muestra de ocasión* (o incluso con las otras 100 *muestras artificiales* generadas con la técnica de *replicación repetida balanceada* [RRB] para *estimar el error* de las *estimaciones calculadas* con el *modelo*).

De modo que este extraordinario hallazgo científico (el más importante obtenido para Cuba con el ERCE-2019, en mi modesta opinión) pudiera ser presentado a un público no especializado con un simple *gráfico de barras* (propio de la elemental Estadística Descriptiva), sin transparentar los complejos procesos estadísticos de orden superior que están detrás de él.

Si se fijan en el gráfico de referencia (la Figura No.6, debajo) podrán observar que, de acuerdo con los rigurosos resultados del ERCE-2019, se puede afirmar que para los más de 227 mil estudiantes cubanos que integraron la *población* del ERCE-2019 en el curso escolar 2018-2018, los *'impactos'* de la *enseñanza desarrolladora*, como norma, duplican los proyectados por la mayoritariamente empleada *enseñanza explicativo-ilustrativa* (L3 significa Lectura 3°, M3 es Matemática 3°; C6 es Ciencias 6°, etc.).

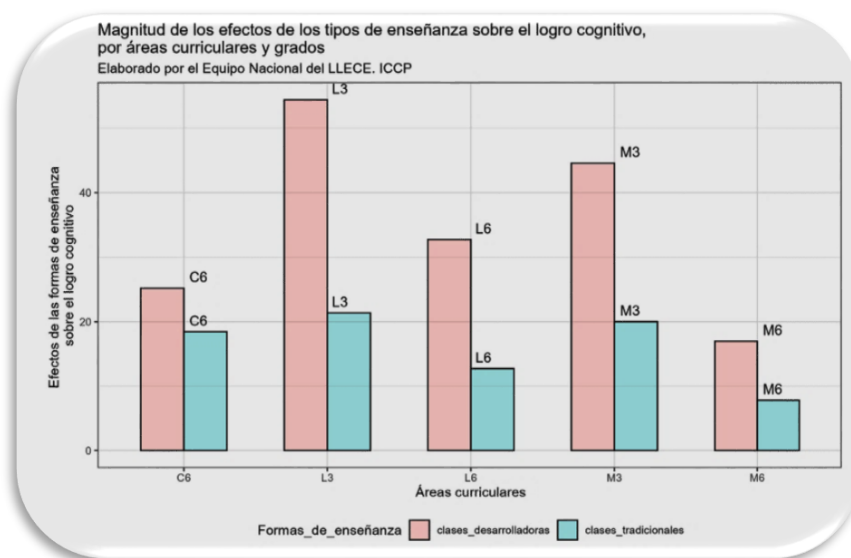


Figura No. 6: Estimaciones poblacionales de las efectos de las clases sobre los logros del aprendizaje, contrastando las correspondientes a la enseñanza del tipo explicativo-ilustrativa y desarrolladora, respectivamente (Torres et al., 2022; p. 68).

Estas potencialidades de los *modelos de regresión lineal* (la *explicativa* y la *predictiva*) se pueden aprovechar mejor si los análisis estadísticos correspondientes se efectúan en un *entorno informático* tan revelador (¡a la vez que software libre!) como es R-Project (Torres, 2018). Ello es posible no solo por la profundidad de los *paquetes estadísticos* (programaciones) que él posee, sino además por la posibilidad que tiene de proporcionar *salidas* de análisis estadísticos en forma de mapas e, incluso, de imágenes satelitales.

En la Figura No.7 aparece un ejemplo de ello que he desarrollado con las bases de datos liberadas por la OREALC-UNESCO, Santiago con los resultados del segundo Estudio Regional Comparativo y Explicativo (ERCE-2006). Me he limitado a los *logros del aprendizaje* en Matemática 6° grado de los escolares cubanos de la provincia de Holguín, que participaron entonces en ese estudio como parte de la *muestra nacional*.

La Figura está compuesta por dos imágenes mapeadas y coloreadas por R-Project a partir de los datos proporcionados al *guion de programación* que realicé. El mapa de la izquierda proyecta las *puntuaciones reales* de los escolares holguineros *promediados* por municipios (claro, algunos de ellos no fueron seleccionados dentro de la *muestra aleatoria y representativa* del país, luego no aparecen todos aquí).

Préstese atención a la *escala tricolor* que aparece a su derecha, debajo ('PUNTAJE ESTÁNDAR FINAL'). Los municipios fueron coloreados automáticamente por el software, en dependencia de que su *promedio* se situara en torno al *promedio* de los municipios participantes (blanco), o por debajo (azul) o por encima (rojo).

Así, por ejemplo, el municipio 'Rafael Freyre' se encuentra en torno al '*promedio provincial*' que es superior a los 600 puntos (una *unidad de desviación estándar* de la *media regional* en aquel estudio). 'Mayarí', con el color azul más intenso entre todos ellos, obtuvo los resultados *factuales* (reales) más bajos de toda '*la provincia*' e incluso por debajo de la *media regional* (que fue de 500 puntos, en aquella ocasión).

Por último, 'Cueto' obtuvo el mejor resultado por mucho, con cerca de 800 puntos de la *escala SERCE-2006* (es decir, '*dos pistas*' por delante del '*pelotón*' regional, si se asume el símil del posicionamiento de los *logros del aprendizaje* de los países participantes como una carrera atlética en una pista de 400m).

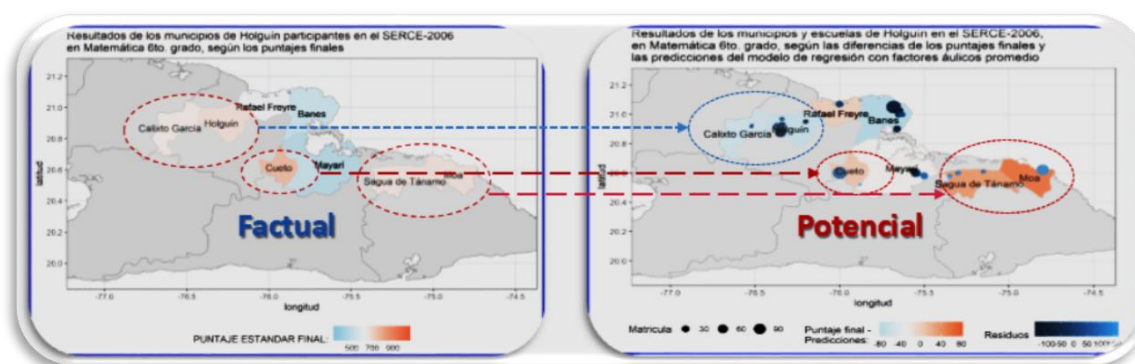


Figura No. 7: Salidas mapeadas y coloreadas de los resultados de una submuestra nacional en el ERCE-2006; contrastando datos factuales con sus predicciones (Elaboración propia).

En la representación de la derecha se tienen los resultados de *ajustar* los *logros del aprendizaje* anteriores por tres *índices estadísticos* que agrupan *factores asociados* referidos a las '*condiciones áulicas de la enseñanza*', con la ayuda de un *modelo de regresión lineal múltiple*. Es decir, se *estimaron* los efectos sobre el *logro del aprendizaje* de la Matemática 6° grado por tres *variables explicativas* a la vez.

Le pido al lector que se concentre primero en las escalas numéricas que aparecen debajo, a la derecha, de esa segunda imagen. Está la de las predicciones de los puntajes de los estudiantes como resultado del *efecto combinado* de aquellos *índices estadísticos* construidos ('Puntaje final-Predicciones'); como se trata de incrementos o decrecimientos oscilan entre -80 y 80 puntos. La otra es la escala de los *residuos*, que son las diferencias entre el puntaje real que habían obtenido los estudiantes y sus correspondientes *predicciones*, de acuerdo con el *modelo* elegido y validado.

Solicité a R-Project agregar una tercera escala, esta *discreta*, que expresa rangos de matrículas de las escuelas de la submuestra (a la izquierda, con tres clases: 30, 60 y 90). Al comparar los resultados factuales (a la izquierda) con los potenciales (a la derecha), se aprecia que los municipios de 'Cueto', 'Sagua de Tánamo' y 'Moa' continúan en una posición destacada, no solo por haber obtenido puntuaciones originales por encima de la *media* de la submuestra, sino también por ser superiores a los *valores esperados* de acuerdo con las percepciones de sus estudiantes sobre las condiciones de la enseñanza, y *predichos* por el *modelo* final.

Por su parte, los municipios de ‘Holguín’ y ‘Calixto García’ obtuvieron resultados por debajo de lo esperado, según las agrupaciones de *factores asociados de naturaleza áulica*; mientras que los de ‘Banes’ y ‘Rafael Freyre’ no sufren cambios importantes, tras *ajustar el modelo diseñado con esas variables predictivas*.

Y ‘Moa’ se mejora mucho así mismo; es decir, si las escuelas primarias de este último municipio hubieran fortalecido más el trabajo pedagógico en torno a las tres *variables explicativas* consideradas como potencialmente importantes (desde la Pedagogía), entonces las puntuaciones de sus estudiantes primarios hubieran sido, por mucho, superiores a las realmente alcanzadas.

Un detalle más... R-Project nos *dice* también con su ‘*mapeo coloreado*’ que la variable ‘*tamaño de la matrícula de las escuelas*’ no parece ser una *buena predictora* del *logro del aprendizaje* en Matemática 6° grado.

Aunque no la introduje en el *modelo de regresión lineal múltiple* (para no complicar más la interpretación de sus resultados), podemos observar que lo mismo asume valores muy grandes en municipios con *promedios* bajos (como ‘Banes’ y ‘Holguín’), que en municipios con *promedios de puntajes* muy altos (como ‘Cueto’ y ‘Moa’). Pasa otro tanto si se aprecia la ubicación de las escuelas de baja matrícula. Vamos, ¡que la cantidad de educandos matriculados en la escuela parece ser *invariante* a los resultados en Matemática 6° grado! ¡Son otros los *factores* que lo *explican mejor*!

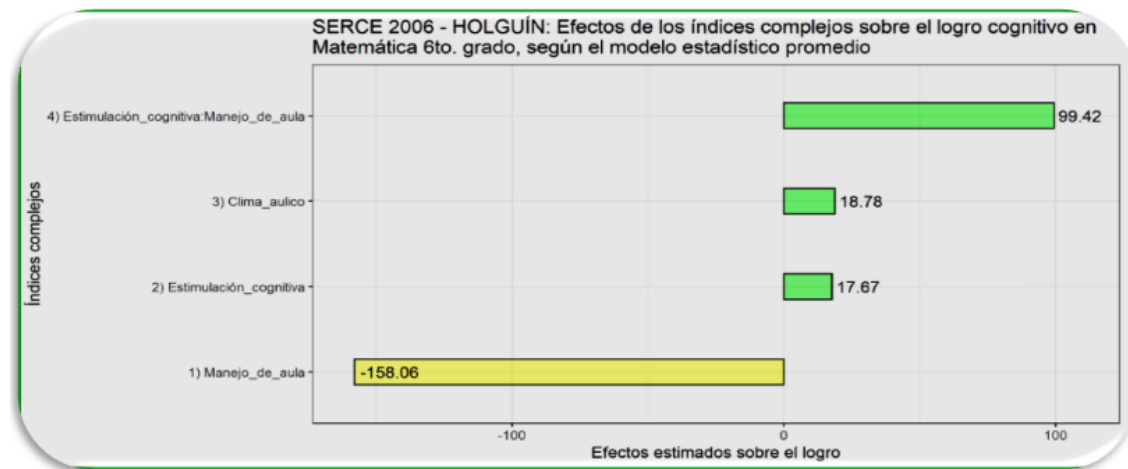


Figura No. 8: Representación de los efectos estimados de los factores áulicos sobre los logros del aprendizaje de la Matemática 6° grado, en el ejemplo (Elaboración propia).

En la Figura No.8 (arriba), aparece una representación más elemental (un *gráfico de barras*) de la *magnitud* y el *sentido* de los *efectos* sobre los *logros del aprendizaje* de la Matemática 6° grado de las tres *variables explicativas* asumidas en el ejemplo, los *índices estadísticos*: ‘*Manejo de aula*’, ‘*Estimulación cognitiva*’ y ‘*Clima áulico*’; además de un cuarto *efecto* dado por la *interacción* de ‘*Manejo de aula*’ y ‘*Estimulación cognitiva*’.

Y se aprecia, además, que cuando se *combinan* estas dos últimas *variables explicativas* se genera un *efecto positivo* mayor, aun cuando el del ‘*Manejo de aula*’, por sí solo, *no explique bien* el *logro cognitivo*. ¡Cuánta información útil y no perceptible a simple vista nos puede llegar a dar esta *técnica multivariante* de ‘*regresión lineal*’ con la ayuda de R-Project!

Me faltan por explicar dos *técnicas multivariantes* más para –al menos mínimamente– poder ilustrar las potencialidades combinadas de la Estadística Inferencial ‘*avanzada*’ y el entorno informático R-Project. Pero ya se me ha hecho muy extenso este post. Prefiero no abusar más de ustedes, mis fieles lectores, y dejar el resto para una segunda parte de este tema.

Es muy importante llegar al final de él, pues hay que responder –inevitablemente– a la pregunta de cuál de ‘las tres’ Estadísticas recomendar para la investigación educativa; si la Estadística Descriptiva, o si la Estadística Inferencial ‘*clásica*’, o si esta última que prefiero llamar Estadística Inferencial ‘*avanzada*’. No vaya a pensar el lector que la respuesta es tan evidente. De momento anticipo que –en lo que a mi modesta opinión respecta– no es única.

¡Los esperamos el próximo fin de semana! ¡Sigán con nosotros!

(Tomado de los Blogs ‘Investigación Educativa en Cuba’, de Google, y ‘Evaluación Educativa’, de CubaEduca)

Referencias bibliográficas

1. Hernández-Sampieri, R., Fernández, C. & Baptista, M. P. (2010). *Metodología de la investigación* (Quinta edición). México D. F., México: McGraw-Hill.
2. OREALC-UNESCO, Santiago (2021). *Estudio regional comparativo y explicativo (ERCE-2019). Reporte nacional de resultados. Cuba*. Chile, Santiago: UNESCO (Recuperado de <http://learningportal.iiep.unesco.org>).
3. Pérez, C. (2004). *Técnicas de Análisis Multivariante de Datos*. España, Madrid: Pearson Educación, S.A.
4. Torres, P. (2013). *El arte de enseñar científicamente. Consejos útiles para docentes noveles*. Cuba, La Habana: Editorial Pueblo y Educación.
5. _____ (2016). *Retos de la investigación educativa actual. Aportes a su tratamiento*. Universidad en Ciencias Pedagógicas “Enrique José Varona”. Cuba, La Habana (Recuperado de <https://drive.google.com/file/d/17w13EvbTYMR266KHI2UsFw9VDyCXwal/view?usp=sharing>).
6. _____ (2018). Lo que todo investigador educativo cubano debiera conocer: el entorno informático R. *Atenas, Vol. 4, Núm. 44*. (Recuperado de <https://www.redalyc.org/articulo.oa?id=478055154001>)
7. Torres, P. et al. (2022). *Cuarto Estudio Regional Comparativo y Explicativo (ERCE-2019). Informe Nacional II*. Cuba, La Habana: Instituto Central de Ciencias Pedagógicas. (Informe de investigación)
8. Triola, M. (2018). *Estadística (Decimosegunda edición)*. México, Ciudad de México: Pearson Educación (ISBN 978-607-32-4377-3).
9. Universidad de Vigo-España (2014). *RWizard software versión 4.2* (Recuperado de www.ipez.es/RWizard) (Registro No. 03/2014/1262)