

# Diseño y validación de un instrumento de medida para el razonamiento estadístico de estudiantes de las Ciencias de la Salud \*

**Jaime A Gaviria Bedoya**    *Universidad de Antioquia*  
**Difariney González Gómez**    *Universidad de Antioquia*  
**Jhony A Villa Ochoa**    *Universidad de Antioquia*

---

**Resumen:** Se presenta el diseño y validación de un instrumento de medición para evaluar el razonamiento estadístico de estudiantes de posgrado de Ciencias de la Salud de una universidad pública colombiana. Se utilizó un formato de prueba objetiva de opción múltiple de 27 ítems. Los ítems se adaptaron a partir de instrumentos validados y se construyeron nuevos ítems a partir de resultados de artículos publicados en Ciencias de la Salud. La versión final del instrumento fue respondida por 149 estudiantes de posgrado de maestría y doctorado en las áreas de Salud. Se presentan resultados descriptivos de la muestra y la puntuación total del instrumento. Además, se muestran los resultados de un modelo de ecuaciones estructurales para la validez de constructo. Debido al tamaño reducido de la muestra, se utilizaron métodos de estimación especiales como el enfoque estructural after measurement (SAM) (Rosseel and Loh, 2022). Se usó principalmente la librería de R *lavaan*.

**Keywords:** razonamiento estadístico, validación instrumentos, Ciencias de la Salud, Modelo de ecuaciones estructurales

---

## Introducción

El razonamiento estadístico se considera un objetivo de aprendizaje que es importante promover y desarrollar como parte de la formación de los estudiantes que toman cursos de estadística en diferentes niveles escolares (Ben-Zvi, Makar and Garfield, 2018). En particular, este razonamiento es esencial para los estudiantes de Ciencias de la Salud, ya que les permite comprender la información cuantitativa que encuentran en su contexto, cómo interpretar los riesgos y las probabilidades en los estudios médicos, así como tomar decisiones en situaciones en las que la variabilidad y la incertidumbre están presentes. A pesar de la reconocida importancia del razonamiento estadístico, especialmente en la educación universitaria, los investigadores sugieren realizar más estudios que aborden este razonamiento y su medición en estudiantes de Ciencias de la Salud (Fioravanti, Greca and Meneses, 2019). Además, diferentes revisiones bibliográficas coinciden en la importancia del uso de instrumentos validados para medir este constructo teórico (Langrall et al., 2017). Debido a que no existe consenso sobre la multidimensionalidad del razonamiento estadístico (Ferrerías, 2014, Zieffler et al. (2008)), esta investigación tiene como pregunta de investigación: ¿cuál es la estructura de dimensionalidad del razonamiento estadístico de los estudiantes de las Ciencias de la Salud?. El objetivo de este estudio es describir el proceso de diseño y validación de un instrumento de medida para el razonamiento estadístico en estudiantes de posgrado de Ciencias

---

\* Autor de contacto: [jaime.gaviria@udea.edu.co](mailto:jaime.gaviria@udea.edu.co).

de la Salud de una universidad pública colombiana.

La estructura del documento es la siguiente: En la segunda sección se presentará el diseño y validación del instrumento propuesto. En la sección de resultados se presentará la caracterización sociodemográfica de la muestra de estudiantes, las medidas de resumen de sus puntajes totales de razonamiento estadístico y los resultados de los modelos de ecuaciones estructurales SEM. Finalmente, en la última sección se presentarán las conclusiones principales del estudio y las implicaciones para la investigación en educación estadística.

## **Diseño y validación del instrumento**

### *Diseño del instrumento de medida para el razonamiento estadístico*

El diseño del instrumento se llevó a cabo siguiendo las recomendaciones de los estándares para las pruebas en educación y psicología ([American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014](#)). Se optó por un formato de prueba objetivo de respuesta única, verdadero-falso y de opción múltiple porque este tipo de ítems es adecuado para evaluar las habilidades cognitivas. Luego, se construyó el test de especificaciones (blueprint) y se elaboró un banco inicial de 55 ítems que, luego del proceso de revisión de expertos y las entrevistas en voz alta a los estudiantes, se redujo a 27 ítems en la versión final. Para la selección de los ítems se tomaron dos tipos de fuentes: 1) instrumentos reconocidos, publicados y validados para el razonamiento estadístico y 2) se adaptaron ítems de investigaciones sobre el uso de la estadística en Ciencias de la Salud. Las áreas de contenido a evaluar corresponden al plan de estudios de los cursos de bioestadística de primer año para estudiantes de maestría y doctorado de las Ciencias de la Salud de la universidad, e incluyen estadística descriptiva, probabilidad, inferencia estadística paramétrica y no paramétrica. Siguiendo la Teoría Clásica de los Test (TCT), la puntuación total del instrumento se calculó como la suma de las respuestas correctas al instrumento.

### *Validación y análisis de la evaluación*

La población de estudio se definió como estudiantes de posgrado activos pertenecientes a alguno de los programas de las facultades de las áreas de Ciencias de la Salud de la Universidad. Se definió una muestra por conveniencia. Se incluyeron en el estudio estudiantes de posgrado de maestría y doctorado de una de las diez áreas de salud de la Universidad. Para estimar las relaciones estructurales entre las dimensiones del razonamiento estadístico, se siguieron tres etapas ([Kline, 2023](#)), aunque por cuestiones de espacio en este documento solo se presentan los resultados de la última etapa. En la etapa 1 se realizó un Análisis Factorial Confirmatorio (AFC). Luego, en la etapa 2 se estableció la confiabilidad y validez del instrumento de medición. La confiabilidad se evaluó a través de la consistencia interna, utilizando el  $\alpha$  de Cronbach y los coeficientes G6 de Guttman de forma global y para cada dimensión encontrada en la etapa 1. Finalmente, en la etapa 3 se estimó un modelo estructural siguiendo el enfoque de ecuaciones estructurales (SEM). El modelo de medición propuesto fue un modelo reflexivo de tres factores y, debido al pequeño tamaño de la muestra, se utilizaron métodos de estimación especiales, como el enfoque estructural después del de medida (SAM) ([Rosseel and Loh, 2022](#)). Todos los análisis estadísticos se realizaron en el software @R versión 4.3.1 y la librería lavaan ([Rosseel, 2012](#)).

## Resultados

### *Descriptivos de la muestra y el puntaje*

El trabajo de campo se realizó en línea utilizando Microsoft Forms entre los meses de noviembre de 2022 y abril de 2023. Un total de 150 estudiantes de posgrado de diez facultades diferentes de Ciencias de la Salud de la Universidad respondieron el instrumento para el razonamiento estadístico. El número de casos completos fue elevado (95.97%), por lo que el manejo de datos faltantes mediante el método pairwise deletion. Sin embargo, se realizó un análisis de sensibilidad en el que se realizaron 50 imputaciones múltiples con algoritmo de imputación multivariante por ecuaciones encadenadas (MICE) de la librería mice (Buuren and Groothuis-Oudshoorn, 2011) y al comparar los coeficientes de regresión y sus errores estándar se observó que eran muy similares.

Se encontró que el 53% de los estudiantes eran mujeres, el 54% de los estudiantes eran de la Facultad Nacional de Salud Pública, seguida de la Facultad de Medicina (25%). El nivel de posgrado predominante fue maestría (52.1%) y la media y desviación estándar de la edad fueron 36 años (8.8) respectivamente. La puntuación mínima fue de cuatro puntos y la máxima de 27 puntos. La media y la desviación estándar del instrumento fueron de  $19.21 \pm 5.58$  puntos, respectivamente. Estos resultados indican una gran variabilidad en la puntuación total de razonamiento estadístico de los estudiantes de posgrado.

### *Modelo de ecuaciones estructurales SEM*

Se estimó primero un modelo de análisis factorial confirmatorio con tres factores y debido a que todos los ítems fueron categóricos, se implementó el método de estimación diagonally weighted least squares (DWLS). Posteriormente se estimó un modelo de ecuaciones estructurales (SEM) con la librería lavaan, el cual convergió. El modelo propuesto fue un modelo reflexivo de tres factores y la sintaxis usada se muestra a continuación

```
> ##### Modelo sam local con eliminación datos faltantes
>
> library(lavaan)
> fit2.1_cb.sam<-sam(modelo2.1CB,data=x_preguntas.f,ordered=T,cmd="sem",
+                   output="lavaan",mimic="Mplus",verbose=F,
+                   warn=F,missing="pairwise")
> summary(fit2.1_cb.sam,standardized=T,estimates=F)
## This is lavaan 0.6.16 -- using the SAM approach to SEM
##
##   SAM method                                LOCAL
##   Mapping matrix M method                    ML
##   Number of measurement blocks                 3
##   Estimator measurement part                  DWLS
##   Estimator structural part                   ML
##
##   Number of observations                      149
##   Number of missing patterns                  6
##
## Summary Information Measurement + Structural:
##
```

```
##      Block      Latent Nind  Chisq Df
##        1      problem    4  1.041  2
##        2      analysis   14 91.043 77
##        3 conclusions    9 15.763 27
##
##      Model-based reliability latent variables:
##
##      problem analysis conclusions
##      0.654      0.869      0.818
##
##      Summary Information Structural part:
##
##      chisq df cfi rmsea srmr
##        0  0  1      0      0
```

```
> parameterEstimates(fit2.1_cb.sam)[1:6,]
##      lhs op      rhs est   se    z pvalue ci.lower ci.upper
## 28 conclusions ~      problem 0.867 1.037 0.836 0.403 -1.166 2.899
## 29 conclusions ~      analysis 0.288 0.435 0.661 0.509 -0.566 1.141
## 30      problem ~~      problem 0.100 0.080 1.237 0.216 -0.058 0.257
## 31      analysis ~~      analysis 0.456 0.091 5.020 0.000 0.278 0.634
## 32      problem ~~      analysis 0.190 0.080 2.371 0.018 0.033 0.347
## 33 conclusions ~~ conclusions 0.020 0.033 0.602 0.547 -0.045 0.085
```

Los resultados anteriores muestran que el modelo de ecuaciones estructurales SEM convergió normalmente apesar de tener un tamaño de muestra pequeño ( $N < 200$ ). Las medidas de pseudo-ajuste (Kline, 2023) muestran que la fiabilidad para cada dimensión del razonamiento es aceptable, siendo la dimensión 1 la de menor fiabilidad. Finalmente, de los resultados de las estimaciones de los coeficientes de regresión de la parte estructural del modelo, se obtuvo que la dimensión 3 del razonamiento estadístico es explicada en mejor medida por la dimensión 1, problema, que por la 2, análisis, aunque ninguno de estos coeficientes resultó estadísticamente significativo. Por otro lado, la covarianza entre las dimensiones 1 y 2 fue significativa, lo que corrobora la interrelación entre ambas dimensiones. Estos resultados sugieren que se debe explorar mejor las relaciones estructurales entre las dimensiones del razonamiento estadístico.

## Conclusiones e implicaciones para la investigación en educación estadística

El presente estudio busca realizar un aporte teórico al campo de la educación estadística en términos de conocimiento sobre el razonamiento estadístico de estudiantes de posgrado de Ciencias de la Salud. Los resultados del análisis factorial confirmatorio (que no se muestran por cuestiones de espacio) permitieron confirmar la multidimensionalidad del razonamiento estadístico, lo cual es consistente con los hallazgos de Ferreras (2014). A pesar de que los resultados mostrados son parciales, el diseño y validación de un instrumento de medición del razonamiento estadístico busca proporcionar a los docentes que imparten bioestadística a nivel universitario en facultades de salud información útil sobre los procesos cognitivos involucrados en el razonamiento estadístico, que permita identificar aquellos aspectos más relevantes para el aprendizaje y la adquisición de habilidades en la dimensión específica.

## References

- American Educational Research Association, American Psychological Association and Nacional Council on Measurement in Education. 2014. *The standards for educational and psychological testing*. American Educational Research Association.
- Ben-Zvi, Dani, Katie Makar and Joan Garfield. 2018. *International handbook of research in statistics education*. Springer.
- Buuren, Stef Van and Karin Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." 45(3).  
**URL:** <http://www.jstatsoft.org/v45/i03/>
- Ferreras, Soledad. 2014. "Desarrollo del razonamiento estadístico en estudiantes de enfermería [Tesis doctoral, Universidad Pontificia Comillas, Madrid].".  
**URL:** <https://repositorio.comillas.edu/xmlui/handle/11531/1300>
- Fioravanti, Rodrigo, Ileana Maria Greca and Jesus Ángel Meneses. 2019. "Camino do ensino de estatística para a área da saúde." 22(1):67–96. Publisher: Comitê Latinoamericano de Matemática Educativa.
- Kline, Rex B. 2023. *Principles and Practice of Structural Equation Modeling*. Fifth edition ed. Guilford Press.
- Langrall, Cynthia, Katie Makar, Michael Shaughnessy and Per Nilsson. 2017. Learning and Teaching Mathematics Content: Probability and Statistics. In *Compendium for research in mathematics education*, ed. Jinfa Cai. National Council of Teachers of Mathematics pp. 490–525.
- Rosseel, Yves. 2012. "lavaan: An R Package for Structural Equation Modeling." 48(1):1–36.
- Rosseel, Yves and Wen Wei Loh. 2022. "A structural after measurement approach to structural equation modeling.".   
**URL:** <http://doi.apa.org/getdoi.cfm?doi=10.1037/met0000503>
- Zieffler, Andrew, Beng Chang, Kristine Holleque, Joan Garfield, Danielle Dupuis and Shirley Alt. 2008. "What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature." 16(2):26.  
**URL:** <https://www.tandfonline.com/doi/full/10.1080/10691898.2008.11889566>