Suma Psicológica, Vol. 12, No. 1:
Mayo de 2005, 23-46,
ISSN 0121-4381

EL MODELO DE AJUSTE LINEAL COMO MÉTODO DE CONSISTENCIA INTERNA APLICADO A UNA ESCALA DE EVALUACIÓN DOCENTE

Luis Enrique Prieto Patiño*
Fundación Universitaria Konrad Lorenz, Colombia

ABSTRACT

This article of methodological court, presents the application of the Linear Adjustment Model for the obtaining of the internal consistency of a scale of educational assessment. For the present work a category of the educational assessment of a university institution of the city of Bogotá, the test was used it was applied 1.161 students of different programs, in the development of a process of institutional educational assessment. The analysis of internal consistency for the method of linear adjustment was satisfactory and additionally the coefficient Alpha of Cronbach was calculated (0.92), what indicates a high internal consistency, keeping in mind the type of measure instrument.

Key words: *Linear Adjustment Model, Internal Consistency, Alpha of Cronbach, Assessment, TCT.*

RESUMEN

Este artículo de corte metodológico, presenta la aplicación del modelo de ajuste lineal para la obtención de la consistencia interna de una escala de evaluación docente. Para el presente trabajo se utilizó una categoría de la evaluación docente de una institución universitaria de la ciudad de Bogotá, el instrumento fue aplicado a 1.161 estudiantes de diferentes facultades, en el desarrollo de un proceso de evaluación docente institucional. El análisis de consistencia interna por el método de ajuste lineal fue satisfactorio y adicionalmente se calculó el coeficiente Alfa de Cronbach (0.92), lo que indica una alta consistencia interna, teniendo en cuenta el tipo de instrumento de medida.

***Palabras clave**: Modelo de Ajuste Lineal, Consistencia Interna, Alfa de Cronbach, Evaluación, TCT.*

---

* Docente, Facultad de Psicología, Fundación Universitaria Konrad Lorenz, Bogotá, D. C., Colombia. Luiseprieto1@hotmail.com.

INTRODUCCIÓN

Para la realización de este ejercicio metodológico se tomará como base el trabajo de investigación titulado "Consistencia interna de la prueba de evaluación de desempeño de los docentes" realizado por los investigadores Rodríguez y Borbón en el año 2002, en el cual realizan la propuesta de aplicación del modelo de ajuste a la recta como método o técnica para establecer la consistencia interna de una prueba.

El presente trabajo efectúa un ejercicio metodológico de aplicación del modelo de ajuste a la recta a una de las categorías de la evaluación docente de una institución universitaria de la ciudad de Bogotá, la cual está conformada por 9 ítemes. Dada la extensión del trabajo, se limitó el análisis a solamente una categoría de la prueba en total.

Para iniciar con el análisis se revisarán algunos conceptos relacionados con la consistencia interna, con el modelo de ajuste a la recta para terminar con el ejercicio de aplicación.

ESTIMACIONES DE CONFIABILIDAD Y CONSISTENCIA INTERNA

A pesar de que los resultados en una prueba pueden variar dependiendo de aspectos como el tiempo, las condiciones de aprendizaje, las condiciones de aplicación, entre otros, el investigador espera que los mismos tengan cierta consistencia que dé cuenta de la confiabilidad de la prueba. Para la determinación de qué tan consistentes son los resultados de una prueba, existen varios métodos, los cuales dependen en gran medida del atributo a medir y de sus características, de la naturaleza de los reactivos de la prueba, de la escala de medición en la que se encuentren las puntuaciones, del tipo de prueba (si es de velocidad o de poder), y de si la prueba se lleva a cabo con referencia a un criterio o no. (Cohen & Swerdlik, 2001). Entre tales métodos se puede considerar: el de prueba y posprueba, el de formas alternas o equivalentes, y los métodos que evalúan la estructura interna de una prueba.

Aunque estos métodos son muy conocidos, algunos presentan dificultades para su aplicación, ya sea por falta de tiempo, por costos, o por la poca posibilidad de contar con pruebas realmente paralelas. El problema del test retest, por ejemplo, radica en el costo, los sesgos generados por el aprendizaje, las condiciones de aplicación y la imposibilidad en algunos casos de obtener la segunda medida. Por ello las estimaciones de confiabilidad con métodos de una sola aplicación cobran mayor vigencia, y más si se desea estimar la exactitud de la medida obtenida a través de la prueba y no la estabilidad en los resultados del instrumento.

Estos métodos de única aplicación suelen denominarse de consistencia interna "La consistencia entre reactivos es un término que se refiere al grado de correlación entre todos los reactivos en una escala" (Cohen &

Swerdlik, 2001, pág. 163). Ésta se calcula usualmente a partir de una sola aplicación de una forma única de una prueba. Su utilidad radica también, en que permite evaluar la homogeneidad de una prueba. Lo cual hace referencia a que los reactivos miden un solo rasgo o que son unifactoriales. (Cohen & Swerdlik, 2001).

Dado que buscan realizar estimaciones del grado de relación entre los reactivos que componen una prueba psicológica, además de proporcionar evidencia sobre el nivel de unidimensionalidad de la prueba no siendo ésta una muy buena estimación tal como lo plantea Satiesteban (1990). La consistencia interna también hace referencia a los procedimientos utilizados para estimar la fiabilidad que buscan estimar "... hasta qué punto los sujetos tienen un rendimiento consistente en diversas partes del test..." (Martínez, 1995, pág. 86).

Los procedimientos de consistencia interna deben aplicarse de forma obligatoria a cualquier tipo de instrumento dado que son los únicos que proporcionan evidencia respecto a interrelación entre los reactivos y esto es evidencia del posible muestreo de reactivos. (Nunnaly & Bernstein, 1995).

Dentro de los procedimientos más utilizados se encuentran el elaborado por Kuderd y Richarson (KR-20), cuya fórmula es la más conocida y la cual se aplica a los instrumentos calificados de forma dicotómica como sucede en las pruebas de conocimiento, el otro procedimiento es el denominado coeficiente Alfa de Cronbach, que aplica a cualquier tipo de instrumento psicológico incluyendo los calificados de forma dicotómica. (Anastasi & Urbina, 1998). Este índice "refleja el grado en el que covarían los ítemes que constituyen el test, es, por tanto, un indicador de la consistencia interna del test" (Muñiz, 1998, pág. 54).

Según Aiken (1996), el coeficiente alfa "es una fórmula general para calcular la confiabilidad de una prueba que consiste en reactivos a los que se asignan dos o más valores estimados de calificación a las respuestas" (pág. 90). Otra interpretación que se le puede dar al alfa es como la correlación promedio entre una y otra prueba psicológica de la misma longitud, tomadas de un mismo dominio.

EL MODELO DE AJUSTE LINEAL

De acuerdo con lo expuesto por Anastasi y Urbina (1998), existen diferentes métodos para encontrar la consistencia interna, tales como:

Métodos contrastados: consiste en seleccionar grupos extremos tomando como referente el puntaje total de la prueba. En este caso, para cada reactivo se compara la ejecución del grupo de criterio superior contra el desempeño del grupo de criterio inferior. Los ítemes que no muestran una proporción significativamente mayor de "puntajes esperados" en el grupo superior con relación al inferior se eliminan o revisan, al considerarse inválidos.

Otros procedimientos se apoyan en correlaciones, eligiendo únicamente los ítemes que presentan correlaciones significativas con la prueba. De esta forma, se considera que una prueba tiene consistencia interna empleando este método, dado que cada reactivo discrimina a los que responden en la misma dirección. Sin embargo, es importante tener en cuenta la posibilidad de determinar la consistencia interna de pruebas a partir de estrategias propias de la estadística, la lógica y el cálculo; es por esto que se sugiere como método para el cálculo de la consistencia interna el ajuste a la recta, el cual sin apartarse de la filosofía de los métodos anteriormente citados centra su función en una serie de pruebas que permiten controlar el error, basados en el modelo lineal simple.

De esta manera, lo que se pretende con este modelo es tratar de ajustar una ecuación al conjunto de datos, para que de esta manera se pueda proporcionar un modelo teórico, no disponible de otra forma.

Por lo anterior, al obtener n mediciones $y1$, $y2$,.., yn de una variable respuesta Y (puntaje total), las cuales son producto de un conjunto de datos, $x1$, $x2$.., xn que representan los valores de k variables de predicción (puntaje de los ítemes). El interés consiste en determinar una función matemática elemental (un polinomio), que describa el comportamiento de la variable respuesta, dados los valores de las variables de predicción.

Teniendo en cuenta lo anterior, el modelo que más se aproxima con estas consideraciones es el lineal simple, conocido ampliamente como regresión lineal.

En general, la regresión tiene dos significados, que para el caso que nos ocupa, se emplean simultáneamente:

El primer significado asume la regresión como una distribución conjunta de probabilidad de dos variables aleatorias (puntaje total, puntaje de los ítemes).

Sea X el puntaje de los ítemes y Y el puntaje total. Es obvio que para un valor dado de x es imposible predecir exactamente el puntaje total de alguien en particular. Sin embargo, es posible predecir el puntaje total promedio de todas las personas para los que el puntaje de x es el mismo. En otras palabras, para cada valor de x existe una distribución de puntaje total y lo que se busca es la media de esa distribución, dado x. La gráfica de la media condicional $E(Y/x)$ como una función de x recibe el nombre de curva de regresión de Y sobre X.

El segundo significado es más empírico, surge de la necesidad de ajustar alguna función a un conjunto de datos (Canavos, 1988). En el primer caso, se trata de predecir el puntaje total de una persona en particular, dado el puntaje en cada uno de los ítemes.

Dado un conjunto de datos, puede asumirse una forma funcional para la curva de regresión y así, tratar de ajustar ésta a los datos. En estas circunstancias, la variable respuesta es una variable aleatoria cuyos valores se observan mediante la selección de los valores de las variables de predicción en un intervalo de interés.

En consecuencia, las variables de predicción no se consideran como variables aleatorias, sino que éstas son un conjunto de valores fijos que representan los puntos de observación para la variable respuesta.

Para ello, el procedimiento más útil, sugerido cuando se tiene sólo una variable de predicción es graficar la variable respuesta, contra la variable de predicción. Si esta gráfica refleja una tendencia lineal, deberá suponerse un modelo de regresión lineal. Pero si se evidencia alguna curva, se debe suponer un modelo cuadrático o de mayor grado para ajustarse a los datos.

En este sentido, es importante tener en cuenta, como lo propone Canavos (1988), el análisis de los valores **residuales**, entendidos como el análisis de tres deficiencias comunes:

1. la ecuación de regresión puede no ser lineal en las variables de predicción;
2. la varianza del error (s^2) puede no ser constante y una o más de las variables de predicción que ejercen influencia importante pueden no estar incluidas en el modelo. Igualmente, es necesario considerar los puntajes discrepantes (aberrantes), que son aquellos cuyos valores se encuentran alejados del comportamiento general del resto de los puntajes.

En esencia, un análisis de residuales significa (en principio) realizar un análisis de sus gráficas de los residuos. Si se ha definido la ecuación de regresión en forma correcta y no existe ninguna deficiencia, entonces una gráfica de los residuos contra cualquiera de los valores estimados yi a los correspondientes valores de cada variable de predicción en la ecuación no mostrará ningún patrón, es decir, no existirá ninguna relación entre los residuos y los valores ajustados o entre los residuos y los valores de las variables de predicción.

En otras palabras, la gráfica de los residuos mostrará una absoluta independencia o aleatoriedad, entendiendo que dichos valores residuales no corresponden a errores sistemáticos en la prueba, por el contrario, los residuos corresponderán a errores aleatorios provenientes de los sujetos.

Las preguntas que formula el modelo son:

General:

¿Cuál es nivel de consistencia interna de las puntuaciones de las variables predictora y de respuesta?

Subpreguntas:

¿Cuál es la tendencia y en consecuencia el modelo que mejor describe la relación existente entre la variable respuesta y las variables de predicción en la estimación de la consistencia interna?

¿Existe independencia y aleatoriedad en los residuales de la variable de la respuesta estimada con relación a las variables de predicción?

¿Existe varianza constante en el error de la estimación de los promedios de los puntajes de las variables pre-dictoras y de la variable de respuesta?

¿Cuál es la función de distribución acumulada de los residuales?

¿Existen *outliers* en el contexto de la distribución normal estandarizada?

¿Son significativos al 0.01 los coeficientes de correlación?

En concordancia con esto, las hipótesis estadísticas para los estudios fueron:

H1: La varianza de los residuales es inconstante (H1: s^2 = no es constante al 0.05 y 0.01).

H2: Los residuales se distribuyen de manera diferente a una función acumulativa de una normal (H2: $DN > Dn$).

Ho1: La varianza de los residuales es constante (Ho1: s^2 = constante).

Ho2: Los residuales se distribuyen dentro de una función acumulativa de una normal (Ho2: $DN < Dn$).

APLICACIÓN DEL MODELO DE AJUSTE LINEAL

Los pasos para el desarrollo del modelo de ajuste a la recta se resumen en: a) Efectuar corrección de *Yi,* b) realizar los diagramas de dispersión, c) ajuste a la recta, d) realizar el examen de residuales, e) analizar la independencia de las variables y los residuales, f) realizar la prueba de normalidad de los errores, g) determinar la presencia/ausencia de *outliers* y h) estimar los coeficientes de correlación.

DIAGRAMAS DE DISPERSIÓN Y AJUSTE A LA RECTA

Debido al interés del estudio se utiliza un modelo de regresión para poder determinar el ajuste a la recta de las variables en este caso los reactivos de la categoría. Se asume una forma funcional para la curva de regresión, entonces se utiliza como procedimiento la gráfica de la variable respuesta (y) contra la variable de predicción (x). Como se observa en las figuras 1 a 9, las gráficas revelan una tendencia lineal, para todos los ítemes, entonces se puede suponer un modelo de regresión lineal para cada una de las variables en este caso, los 9 ítemes que componen la categoría.

En esta parte se puede concluir que los 9 reactivos de la categoría superan la prueba de ajuste a la recta, por tanto, se puede seguir con el análisis.

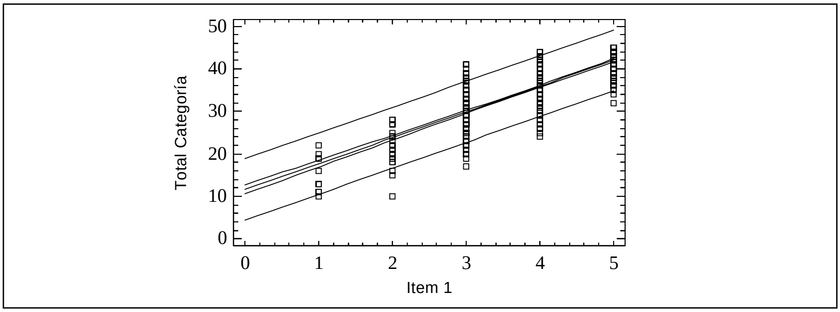

Figura 1. Ajuste a la recta del ítem 1 contra la categoría en total.

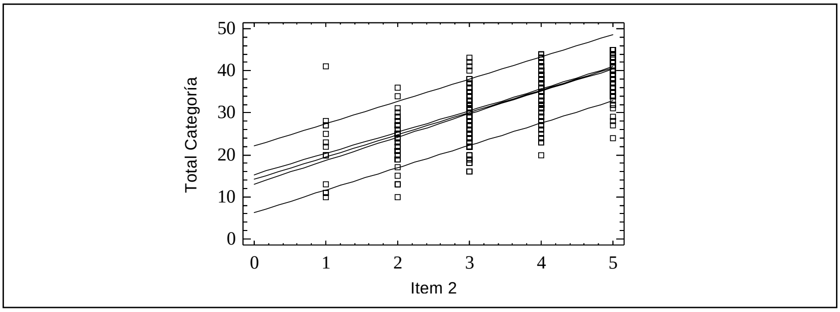

Figura 2. Ajuste a la recta del ítem 2 contra la categoría en total.

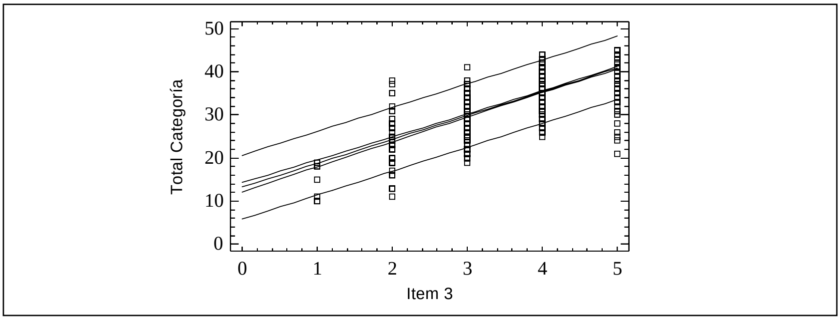

Figura 3. Ajuste a la recta del ítem 3 contra la categoría en total.

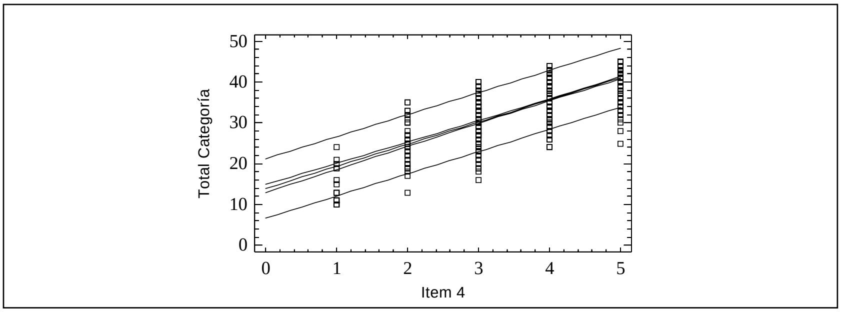

Figura 4. Ajuste a la recta del ítem 4 contra la categoría en total.

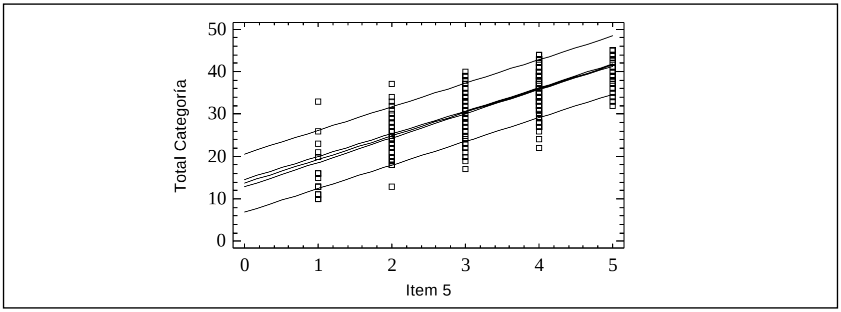

Figura 5. Ajuste a la recta del ítem 5 contra la categoría en total.

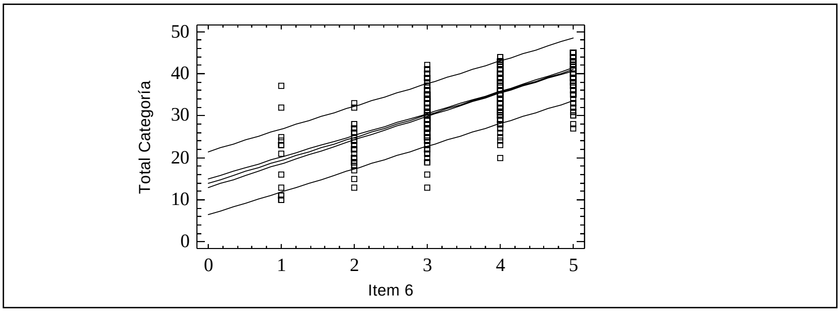

Figura 6. Ajuste a la recta del ítem 6 contra la categoría en total.

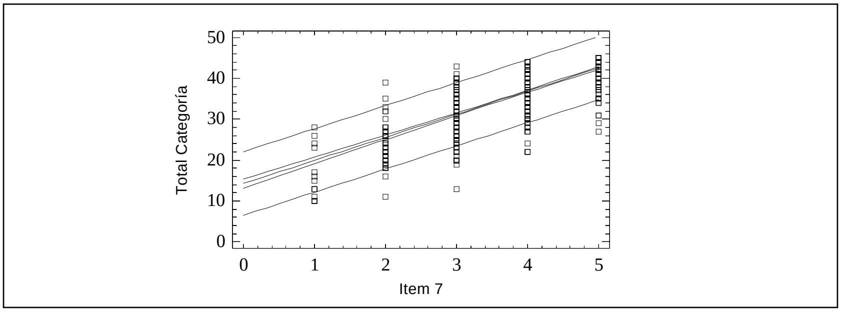

Figura 7. Ajuste a la recta del ítem 7 contra la categoría en total.

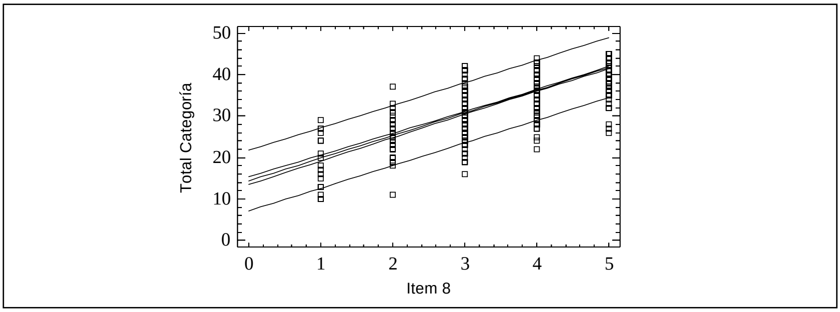

Figura 8. Ajuste a la recta del ítem 8 contra la categoría en total.

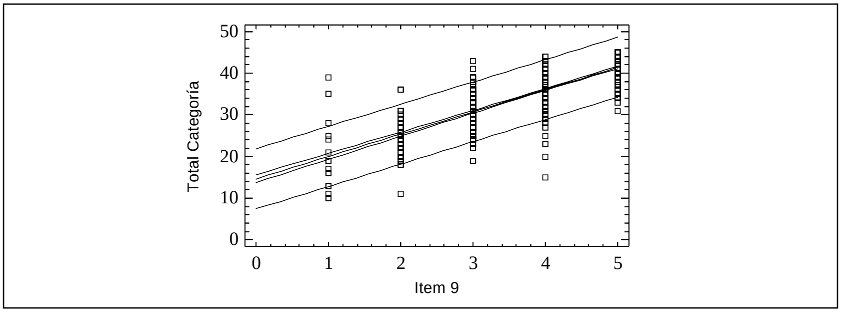

Figura 9. Ajuste a la recta del ítem 9 contra la categoría en total.

EXAMEN DE RESIDUALES

Teniendo en cuenta las figuras 10 a la 18, se puede observar el comportamiento de los residuales para cada una de las variables, en este caso, los ítemes de la categoría. Debido a que no se puede utilizar la prueba de Durbin - Watson puesto que no son datos recogidos a través del tiempo, se puede observar el comportamiento aleatorio de los residuales al presentarse en forma de serrucho la gráfica y mostrando más de 12 picos necesarios, para demostrar que los errores corresponden más a errores de los sujetos que a errores del instrumento.

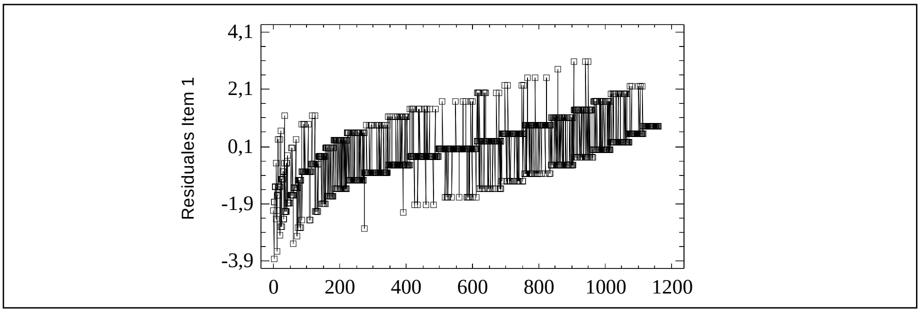

Figura 10. Análisis de residuales del ítem 1 contra la categoría en total.

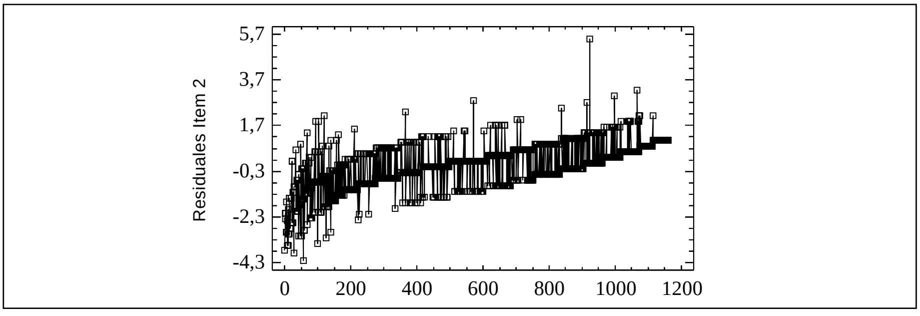

Figura 11. Análisis de residuales del ítem 2 contra la categoría en total.

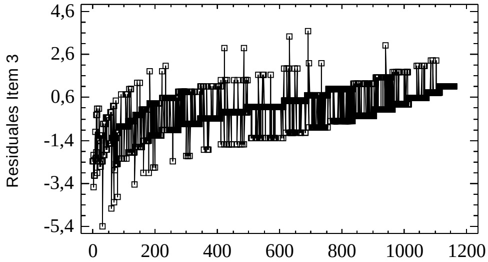

Figura 12. Análisis de residuales del ítem 3 contra la categoría en total.

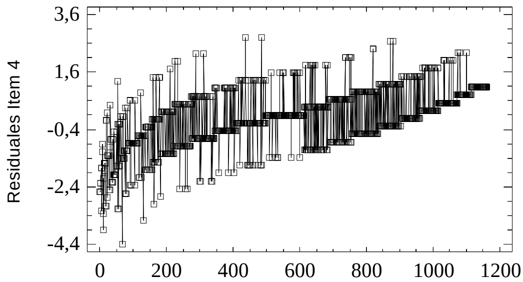

Figura 13. Análisis de residuales del ítem 4 contra la categoría en total.

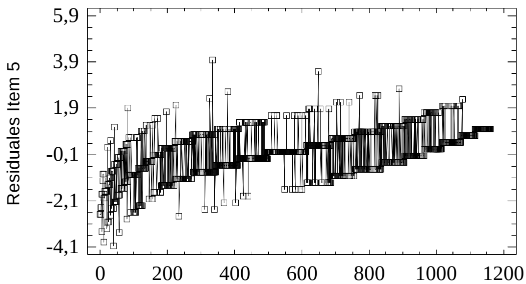

Figura 14. Análisis de residuales del ítem 5 contra la categoría en total.

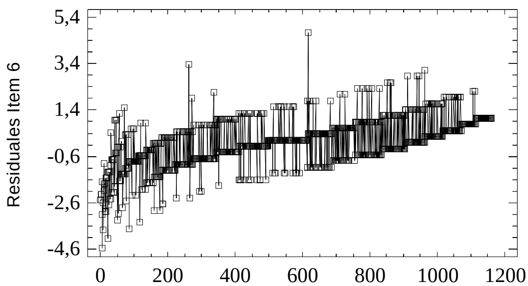

Figura 15. Análisis de residuales del ítem 6 contra la categoría en total.

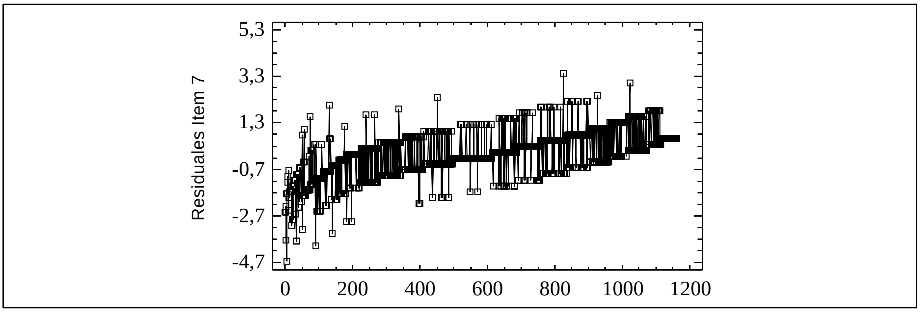

Figura 16. Análisis de residuales del ítem 7 contra la categoría en total.

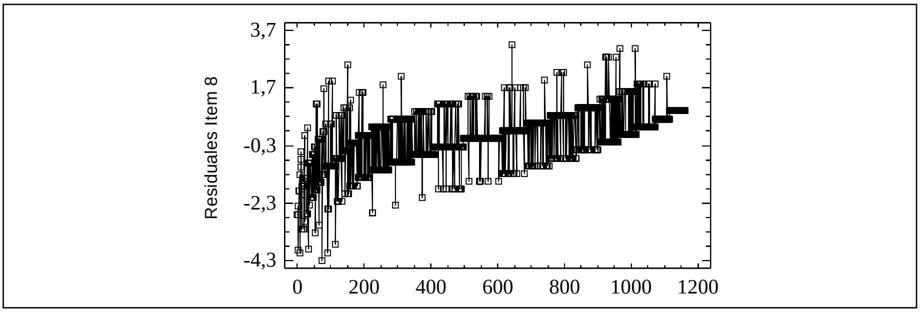

Figura 17. Análisis de residuales del ítem 8 contra la categoría en total.

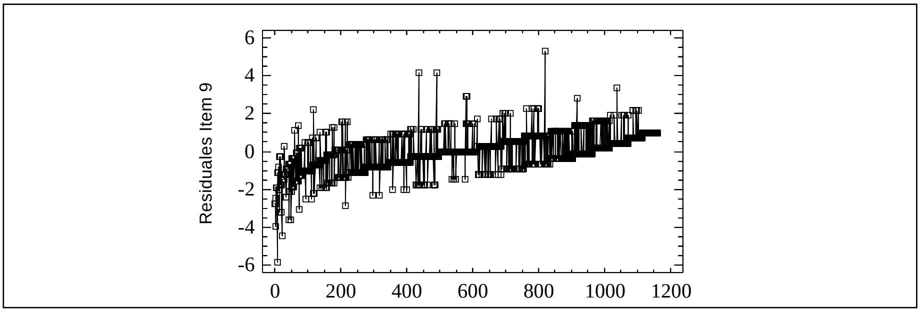

Figura 18. Análisis de residuales del ítem 9 contra la categoría en total.

INDEPENDENCIA DE LAS VARIABLES CON LOS RESIDUALES

Al cruzar los residuales con cada uno de los ítemes se encontró que existe independencia entre la variable y los residuales, lo que muestra la aleatoriedad de los residuales. (Véanse figuras 19 a la 27). Este análisis es satisfactorio dado que lo que se busca, es que lo que mide la variable no esté relacionado con los errores de medida.

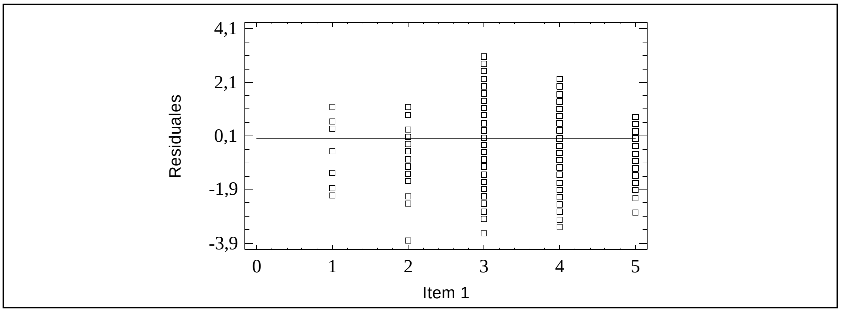

Figura 19. Tendencia de los residuales respecto al ítem 1.

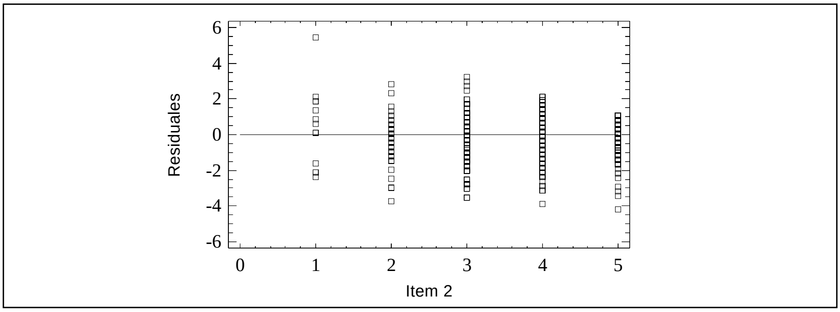

Figura 20. Tendencia de los residuales respecto al ítem 2.

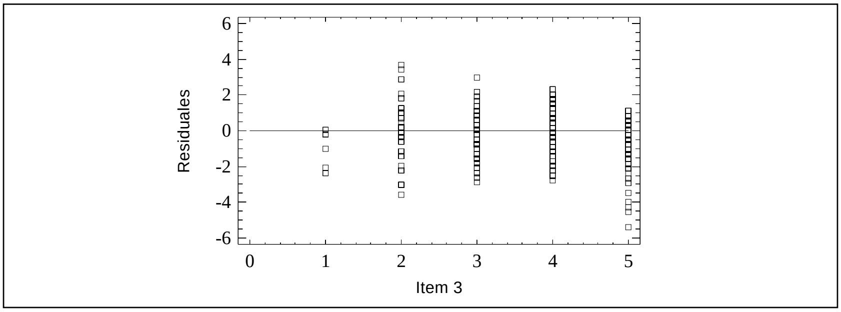

Figura 21. Tendencia de los residuales respecto al ítem 3.

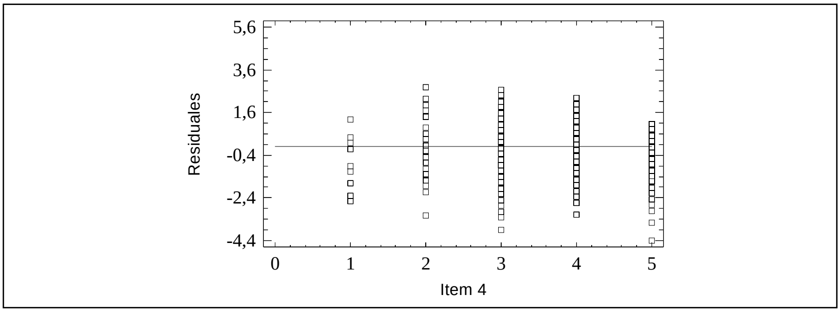

Figura 22. Tendencia de los residuales respecto al ítem 4.

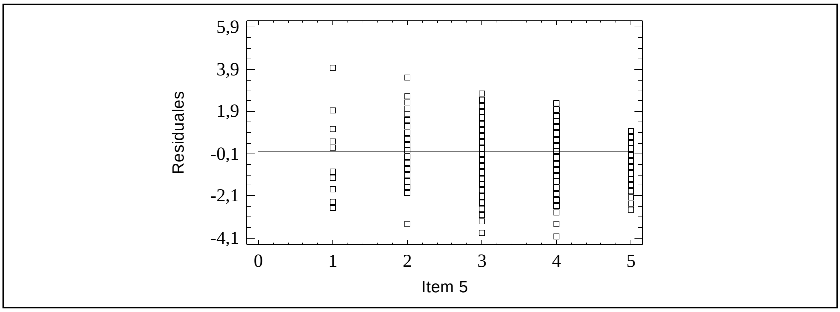

Figura 23. Tendencia de los residuales respecto al ítem 5.

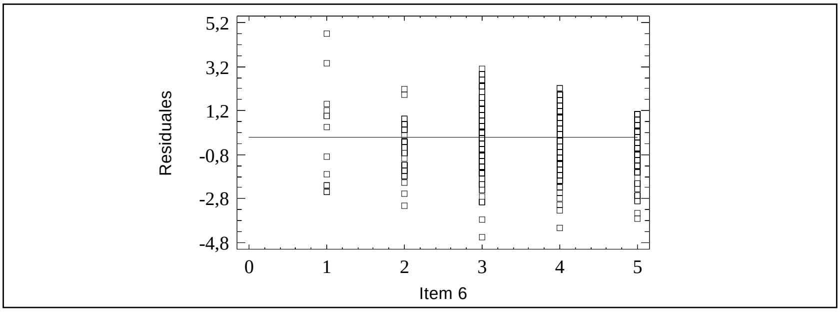

Figura 24. Tendencia de los residuales respecto al ítem 6.

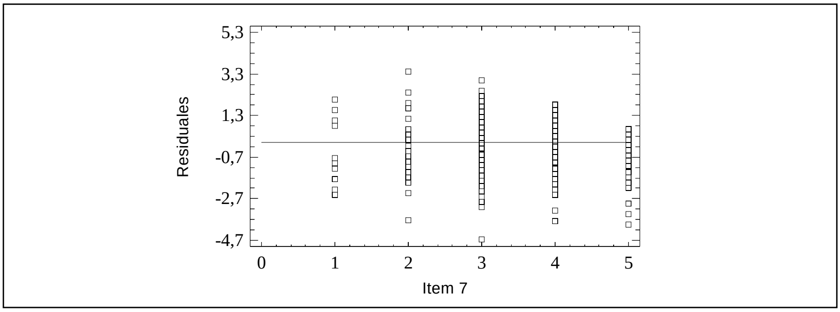

Figura 25. Tendencia de los residuales respecto al ítem 7.

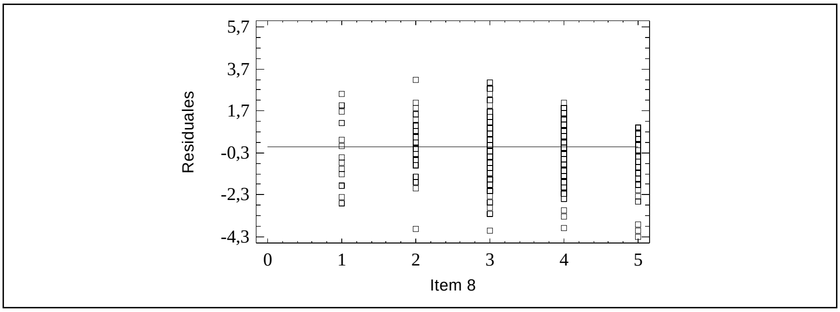

Figura 26. Tendencia de los residuales respecto al ítem 8.

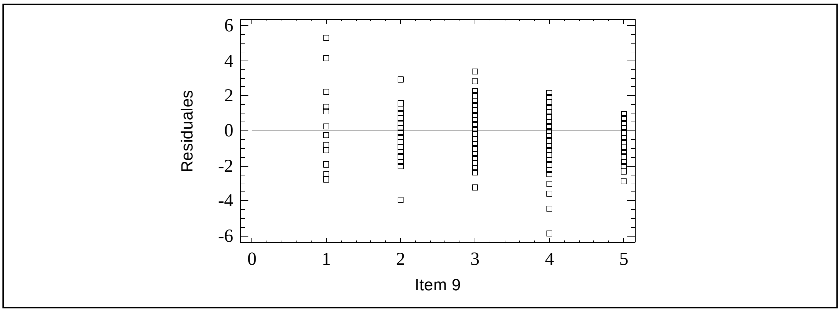

Figura 27. Tendencia de los residuales respecto al ítem 9.

PRUEBA DE HOMOCEDASTICIDAD

En las tablas 1 a 9, se puede observar el análisis de homocedasticidad para cada uno de los ítemes encontrándose que en todos los casos los ítemes presentan una varianza constante respecto al error.

Tabla 1. Prueba de homocedasticidad para el ítem 1 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 49.839438	N2 = 10 SCR1 = 20.0502768
SCR1/8 = 6.22992975	SCR1/8 = 2.5062846

$$\frac{SCR1/8}{SCR2/8} = \frac{6.22992975}{2.5062846} = 2.48572319$$

Teniendo en cuenta que:
Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 5% es rechazar Ho si Fc > ft,α,8,8
Pero Fc = 2.48572319 < ft 0.05,8,8
2.48572319 < 3.44

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 1 de la categoría es constante.

Tabla 2. Prueba de homocedasticidad para el ítem 2 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 84.3099086	N2 = 10 SCR1 = 23.4610489
SCR1/8 = 10.5387386	SCR1/8 = 2.93263111

$$\frac{SCR1/8}{SCR2/8\ 2.93263111} = \frac{10.5387386}{} = 3.59361207$$

Teniendo en cuenta que:
Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 1 % es rechazar Ho si Fc > ft,α,8,8
Pero Fc = 3.59361207 < ft 0.01,8,8
3.59361207 < 6.03

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 2 de la categoría es constante.

Tabla 3. Prueba de homocedasticidad para el ítem 3 contra el puntaje total.

Primer grupo				Tercer grupo
n1 = 10 SCR1 = 75.6418897				N2 = 10 SCR1 = 23.9512766
SCR1/8 = 9.45523621				SCR1/8 = 2.99390958
SCR1/8 / SCR2/8	=	9.45523621 / 2.99390958	=	3.1581569

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 5% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 3.1581569 < ft 0.05,8,8
3.1581569 < 3.44

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 3 de la categoría es constante.

Tabla 4. Prueba de homocedasticidad para el ítem 4 contra el puntaje total.

Primer grupo				Tercer grupo
n1 = 10 SCR1 = 65.1888041				N2 = 10 SCR1 = 23.5702327
SCR1/8 = 8.14860051				SCR1/8 = 2.94627908
SCR1/ / SCR2/8	=	8 8.14860051 / 2.94627908	=	2.76572595

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 5% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 2.76572595 < ft 0.05,8,8
2.76572595 < 3.44

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 4 de la categoría es constante.

Tabla 5. Prueba de homocedasticidad para el ítem 5 contra el puntaje total.

Primer grupo				Tercer grupo
n1 = 10 SCR1 = 58.6405714				N2 = 10 SCR1 = 22.7671885
SCR1/8 = 7.33007142				SCR1/8 = 2.84589857
SCR1/ / SCR2/8	=	8 7.33007142 / 2.84589857	=	2.57566152

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 5% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 2.57566152 < ft 0.05,8,8
2.57566152 < 3.44

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 5 de la categoría es constante.

Tabla 6. Prueba de homocedasticidad para el ítem 6 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 81.2152545	N2 = 10 SCR1 = 23.7233926
SCR1/8 = 10.1519068	SCR1/8 = 2.96542407

$$\frac{SCR1/}{SCR2/8} = \frac{8\ 10.1519068}{2.96542407} = 3.4234497$$

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 5% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 3.4234497 < ft 0.05,8,8
3.4234497 < 3.44

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 6 de la categoría es constante.

Tabla 7. Prueba de homocedasticidad para el ítem 7 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 70.8234767	N2 = 10 SCR1 = 16.5853187
SCR1/8 = 8.85293458	SCR1/8 = 2.07316483

$$\frac{SCR1/}{SCR2/8} = \frac{8\ 70.8234767}{16.5853187} = 4.27025119$$

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 1% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 4.27025119 < ft 0.01,8,8
4.27025119 < 6.03

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 7 de la categoría es constante.

Tabla 8. Prueba de homocedasticidad para el ítem 8 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 73.6729316	N2 = 10 SCR1 = 21.0328307
SCR1/8 = 9.20911645	SCR1/8 = 2.62910384

$$\frac{SCR1/8}{SCR2/8} = \frac{73.6729316}{21.0328307} = 3.50275874$$

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 1% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 3.50275874 < ft 0.01,8,8
3.50275874 < 6.03

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 8 de la categoría es constante.

Tabla 9. Prueba de homocedasticidad para el ítem 9 contra el puntaje total.

Primer grupo	Tercer grupo
n1 = 10 SCR1 = 87.6590409	N2 = 10 SCR1 = 22.8293968
SCR1/8 = 10.9573801	SCR1/8 = 2.8536746

$$\frac{SCR1/8}{SCR2/8} = \frac{87.6590409}{22.8293968} = 3.83974406$$

Teniendo en cuenta que:

Ho: σ2 = constante Vs. H1: σ2 = no es constante.
La R de D sobre α = 1% es rechazar Ho si Fc > ft,α,8,8

Pero Fc = 3.83974406 < ft 0.01,8,8
3.83974406 < 6.03

No se puede rechazar Ho, es decir σ2 = constante. Se puede concluir que la varianza del error en la estimación de los promedios del reactivo 9 de la categoría es constante.

PRUEBA DE NORMALIDAD DE LOS ERRORES

En el siguiente histograma (figura 28) se observa una distribución de los residuales para el reactivo 1 entre - 3.80914 y 3.06604, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0487155.

Según el valor teórico para esta misma prueba con 1 - α = 0.95, n = 21 Dn = 0.0669189. Entonces:

DN	<	Dn
0.0487155	<	0.0669189

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 1 se distribuyen dentro de una distribución normal acumulativa.

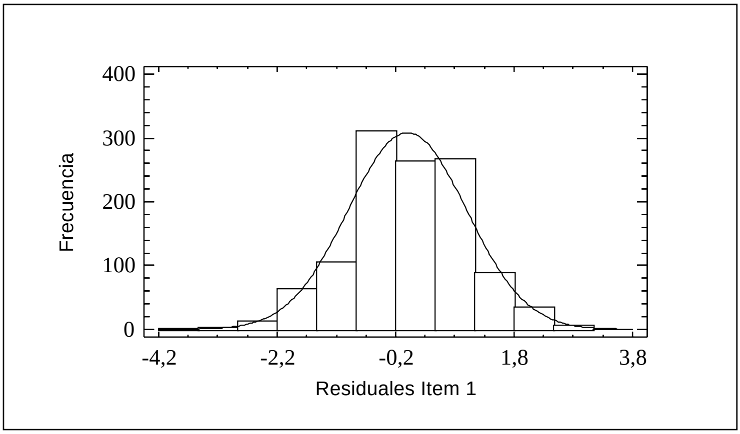

Figura 28. Histograma para los residuales del ítem 1 contra la categoría total.

En el histograma anexo (figura 29) se observa una distribución de los residuales para el reactivo 2 entre - 4.21145 y 5.46436, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0591175.

Según el valor teórico para esta misma prueba con 1 - α = 0.95, n = 21 Dn = 0.0744909. Entonces:

DN	<	Dn
0.0591175	<	0.0744909

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 2 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 30) se observa una distribución de los residuales para el reactivo 3 entre - 5.37868 y 3.68999, es decir, entre -δ d y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0500772.

Según el valor teórico para esta misma prueba con 1 - α= 0.95, n = 21 Dn = 0.0673027. Entonces:

DN	<	Dn
0.0500772	<	0.0673027

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 3 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 31) se observa una distribución de los residuales para el reactivo 4 entre - 4.39543 y 2.78453, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0360437

Según el valor teórico para esta misma prueba con 1 - α = 0.95, n = 21 Dn = 0.0570024. Entonces:

DN	<	Dn
0.0360437	<	0.0570024

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 4 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 32) se observa una distribución de los residuales para el reactivo 5 entre - 4.03264 y 3.96812, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0375298.

Según el valor teórico para esta misma prueba con 1 - α = 0.95, n = 21 Dn = 0.062098. Entonces:

DN	<	Dn
0.0375298	<	0.062098

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 5 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 33) se observa una distribución de los residuales para el reactivo 6 entre - 4.56157 y 4.70898, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0488513

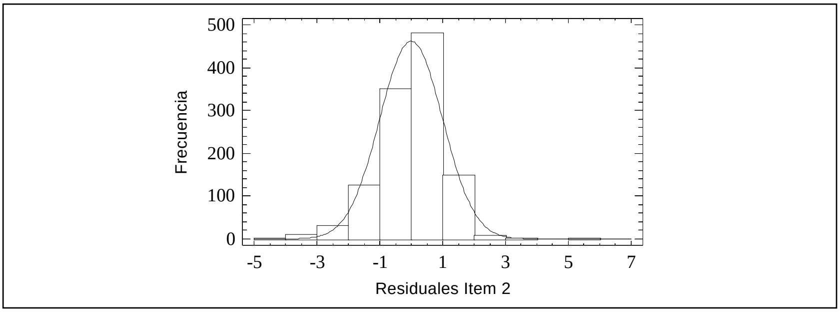

Figura 29. Histograma para los residuales del ítem 2 contra la categoría total.

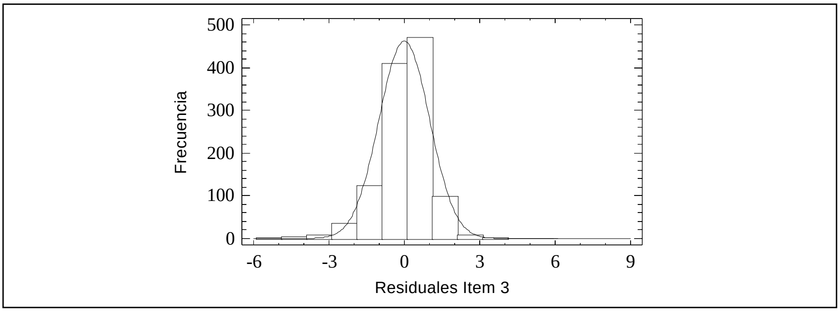

Figura 30. Histograma para los residuales del ítem 3 contra la categoría total.

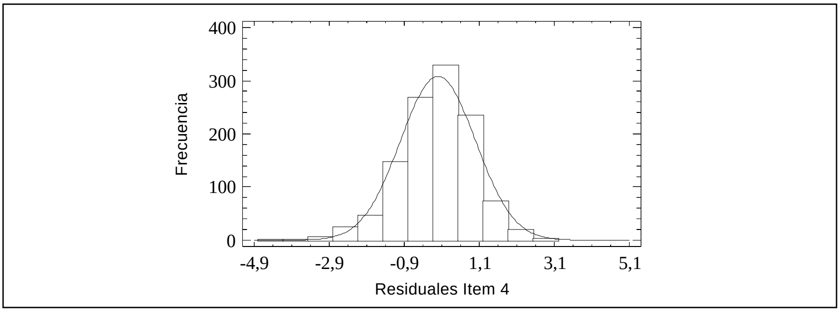

Figura 31. Histograma para los residuales del ítem 4 contra la categoría total.

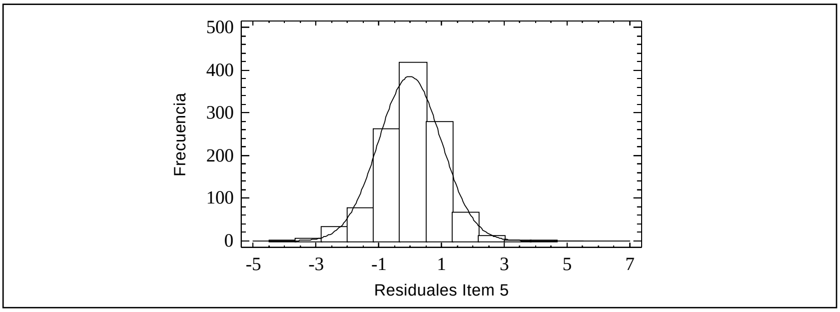

Figura 32. Histograma para los residuales del ítem 5 contra la categoría total.

Según el valor teórico para esta misma prueba con $1 - \alpha = 0.95$, $n = 21$ $Dn = 0.0505643$. Entonces:

DN	<	Dn
0.0488513	<	0.0505643

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 6 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 34) se observa una distribución de los residuales para el reactivo 7 entre - 4.66475 y 3.43301, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0418391.

Según el valor teórico para esta misma prueba con $1 - \alpha = 0.95$, $n = 21$ $Dn = 0.0682248$. Entonces:

DN	<	Dn
0.0418391	<	0.0682248

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 7 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 35) se observa una distribución de los residuales para el reactivo 8 entre - 4.28601 y 3.18356, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.561891.

Según el valor teórico para esta misma prueba con $1 - \alpha = 0.95$, $n = 21$ $Dn = 0.0793591$. Entonces:

DN	<	Dn
0.0561891	<	0.0793591

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 8 se distribuyen dentro de una distribución normal acumulativa.

En el histograma anexo (figura 36) se observa una distribución de los residuales para el reactivo 9 entre - 5.84186 y 5.27683, es decir, entre -3 δ y 3 δ. Según los resultados obtenidos para la prueba Kolmogorov - Smirnov, tenemos que el K - S calculado es igual a (DN) = 0.0497389.

Según el valor teórico para esta misma prueba con $1 - \alpha = 0.95$, $n = 21$ $Dn = 0.0575106$. Entonces:

DN	<	Dn
0.0497389	<	0.0575106

Luego se acepta la hipótesis nula, lo cual indica que los errores o residuales entre g total contra el reactivo 9 se distribuyen dentro de una distribución normal acumulativa.

PRESENCIA DE *OUTLIERS*

Observando las gráficas de ajuste a la recta de todas las variables de análisis en este caso los 9 reactivos de la categoría (véanse figuras 1 al 9) se puede concluir que no existen entre los valores de la muestra, valores aislados que sesguen la estimación de los parámetros.

Se confirman estos resultados con la estandarización de los residuales, verificando que estén en el rango de - 3 δ y 3 δ, rango en que se mueve una distribución normal estandarizada.

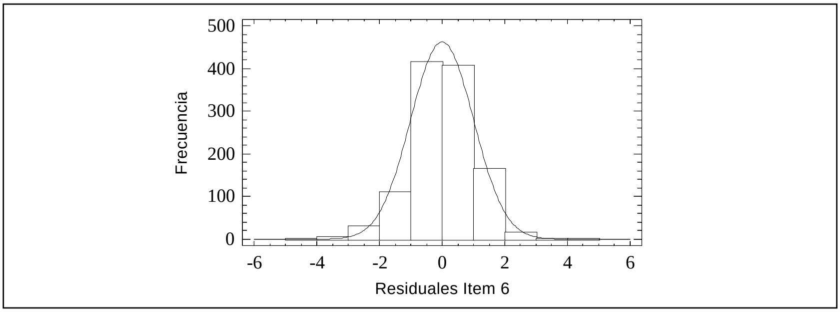

Figura 33. Histograma para los residuales del ítem 6 contra la categoría total.

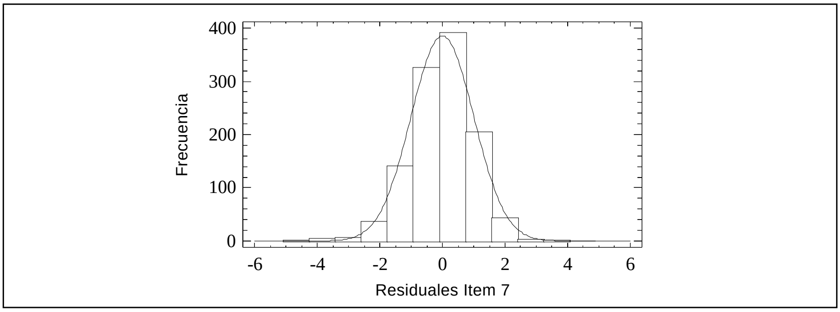

Figura 34. Histograma para los residuales del ítem 7 contra la categoría total.

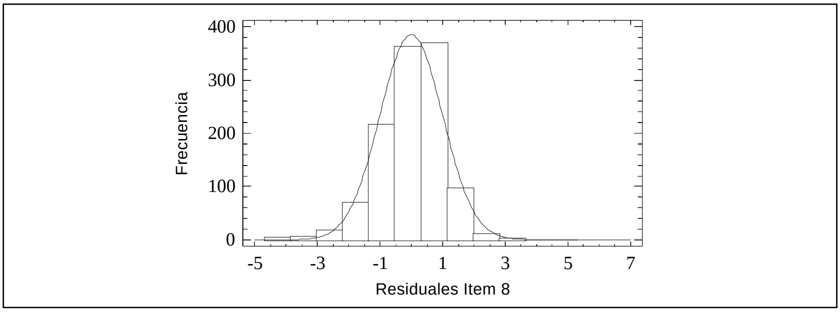

Figura 35. Histograma para los residuales del ítem 8 contra la categoría total.

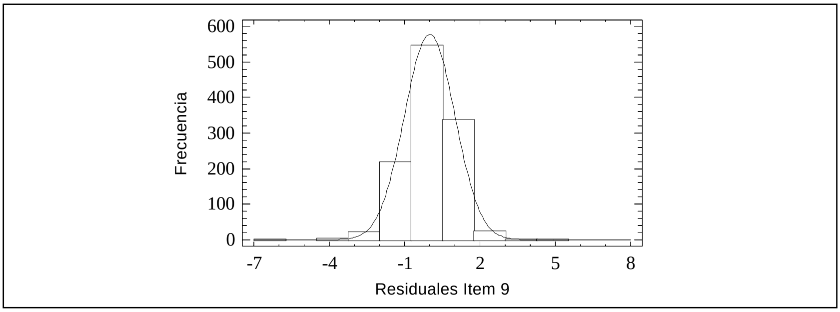

Figura 36. Histograma para los residuales del ítem 9 contra la categoría total.

COEFICIENTES DE CORRELACIÓN

Efectuando las respectivas correlaciones, se encuentra que éstas son positivas altas y fuertes, al nivel de significancia de 0.01 (véase tabla 10). Por otro lado, se observan coeficientes de determinación moderados, lo cual implica que una misma proporción, la variación presente en las observaciones puede explicarse por la presencia lineal de x en la ecuación de regresión.

Tabla 10. Coeficientes de correlación, nivel de significancia y coeficiente de determinación del puntaje total contra las variables, en este caso, los 9 reactivos.

	Valor de coeficiente de correlación	Nivel de significancia	Coeficiente de determinación %
Puntaje total contra la variable 1	0.796284	0.01	63.4068
Puntaje total contra la variable 2	0.747739	0.01	55.9113
Puntaje total contra la variable 3	0.783704	0.01	61.4192
Puntaje total contra la variable 4	0.791193	0.01	62.5987
Puntaje total contra la variable 5	0.815062	0.01	66.4326
Puntaje total contra la variable 6	0.776041	0.01	60.224
Puntaje total contra la variable 7	0.75677	0.01	57.2701
Puntaje total contra la variable 8	0.790665	0.01	62.5151
Puntaje total contra la variable 9	0.794984	0.01	63.2

DISCUSIÓN

De acuerdo con los resultados presentados anteriormente, se puede concluir que:

Existe independencia entre cualquier par de residuales. Esto implica que cualquier variabilidad en la variable de respuesta que no pueda explicarse mediante el empleo de la ecuación de regresión, se debe a un error aleatorio.

Existe independencia entre las variables en este caso los reactivos de la escala de la prueba con respecto al error. Esto permite determinar que los datos se comportaron de acuerdo con el modelo lineal simple.

Los errores se encuentran distribuidos normalmente, es decir que son homocedásticos.

Existe independencia entre los residuales y cada una de las variables (reactivos) por consiguiente, no existe relación entre el error y lo que mide cada uno de los reactivos de acuerdo a como lo plantea la teoría Clásica de los Test (T.C.T).

Se observa varianza constante. Dado que las variables de predicción no son aleatorias, la varianza de Yi también es d 2 para toda i y de esta forma es independiente del punto de observación y en consecuencia, la variación de los cálculos tendrá una variación constante.

Se observa ausencia de *outliers*. Es decir, no existen puntajes que sesgarán los promedios.

Como consecuencia de los anteriores análisis, se observa consistencia interna entre las variables y el puntaje total de la escala evaluada. En cada uno de los procedimientos realizados se encontró un comportamiento adecuado de cada uno de los reactivos lo que permite concluir que la prueba sí tiene una estructura interna apropiada que se puede contrastar con el coeficiente Alfa que fue igual a 0.92, que indica una buena consistencia interna.

REFERENCIAS

Aiken, L. (1996). *Test Psicológicos y evaluación.* (8ª Ed.). México, D.F., México: Prentice Hall.

Anastasi, A. & Urbina, S. (1998). *Test psicológicos.* (7ª Ed.). México, D.F., México: Prentice Hall.

Canavos, G.C. (1998). *Probabilidad y estadística. Aplicaciones y métodos.* México, D.F., México: McGraw-Hill.

Cohen, R. & Swerdlik, M. (2001). *Pruebas y evaluación psicológicas.* (4ª Ed.) México: McGraw-Hill.

Martínez, R. (1995). *Psicometría: teoría de los test psicológicos y educativos* Madrid, España: Síntesis.

Muñiz, J. (1998). *Teoría clásica de los test.* México, D.F., México: Pirámide.

Nunnally, J. & Bernstein, I. (1995). *Teoría psicométrica.* (3ª Ed.) México, D.F., México: McGraw-Hill.

Santiesteban, R. (1990). Psicometría teoría y práctica en la construcción de test. Madrid, España: Norma.

Recibido el 30 de noviembre de 2004 y aceptado el 3 de febrero de 2005