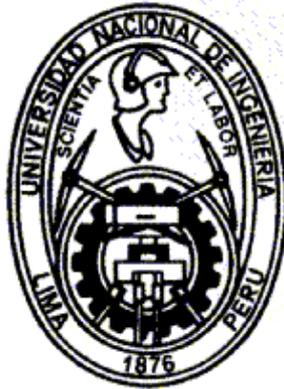


**UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE INGENIERÍA ECONÓMICA Y CCSS
Escuela Profesional de Ingeniería Estadística**



**UNA PROPUESTA METODOLÓGICA PARA LA EVALUACIÓN
DE LA COMPRENSIÓN LECTORA EN ESTUDIANTES DE
QUINTO DE SECUNDARIA EN EL SISTEMA EDUCATIVO
PERUANO**

INFORME DE SUFICIENCIA

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
ESTADÍSTICO**

**POR LA MODALIDAD DE ACTUALIZACIÓN DE
CONOCIMIENTOS**

ELABORADO POR:

OMAR ANTONIO CHINCHARO DEL CORAL

LIMA – PERÚ

2011

Resumen

Las evaluaciones de rendimiento son instrumentos importantes en el ámbito educativo, que buscan evaluar el nivel de habilidad o logro de un estudiante luego de un proceso de instrucción. Con este proyecto de investigación se pretende evaluar el rendimiento en comprensión lectora del estudiante de quinto de secundaria, mediante la estimación de variables no observables o variables latentes: la habilidad del estudiante; y los parámetros del modelo: la dificultad y discriminación del ítem. Se modela la probabilidad de que el estudiante responda correctamente al ítem, mediante el uso de modelos TRI dicotómicos de dos parámetros con enfoque bayesiano. Los datos utilizados provienen del cuadernillo 1 de una prueba de comprensión lectora de las evaluaciones realizadas por el Ministerio de Educación a nivel nacional y que está publicado en Internet. Con la investigación, a través de las estimaciones se muestra una jerarquización de ítems según la dificultad y el índice de discriminación los cuales determinan los niveles de logros que tienen los estudiantes de secundaria. Adicionalmente se muestra que el estudiante de la Costa tiene mayor capacidad que el de la Sierra y Selva, y que el estudiante de género femenino tiene mayor capacidad que el estudiante de género masculino.

Palabras clave: TRI (Teoría de Respuesta al Ítem), Variables latentes, modelos TRI dicotómicos de dos parámetros, enfoque bayesiano.

A mis padres

Agradecimientos

Agradezco a mi asesor el Lic. Luís Huamanchumo de la Cuba, quien me guió en el aspecto metodológico de la investigación. A todos los profesores del curso de titulación quienes fueron los que me dieron los lineamientos en la elaboración de la investigación. Y a los catedráticos de la Escuela Profesional de Ingeniería Estadística (EPIES) quienes fueron los que me brindaron todo el soporte teórico / práctico en estadística durante mi desarrollo profesional.

INDICE GENERAL

Introducción.....	10
Objetivo de la investigación.....	13
Justificación de la investigación	14
Capítulo I:Antecedentes	16
Capítulo II: El problema de la investigación	22
Capítulo III: Modelos TRI Dicotómicos de dos parámetros	24
3.1 Formulación del Modelo TRI dicotómico de dos parámetros	27
3.2 Función de Verosimilitud del modelo TRI.....	30
3.3 Caracterización y propiedades del modelo TRI.....	31
3.4 Estimación de variables latentes y parámetros del modelo.....	34
3.4.1 Método de estimación bayesiana para el modelo TRI	34
3.4.1.1 MCMC usando adaptative rejection sampling (ARS)	37
Capítulo IV: Hipótesis y Metodología de la investigación	38
4.1 Tipo de investigación.....	38
4.2 Hipótesis de investigación.....	39
4.3 Instrumento de medición.....	39
4.4 Tipos de variables para el análisis	40

4.5 Diseño muestral y Muestra.....	40
4.5.1 Población objetivo	40
4.5.2 Unidades de muestreo	40
4.5.3 Unidad de análisis.....	41
4.5.4 Diseño muestral	41
4.6 Recolección de datos y base de datos.....	41
4.7 Análisis estadístico.....	42
Capitulo V: Aplicación y resultados de la investigación.....	43
5.1 Análisis de la prueba bajo el enfoque TCT.....	44
5.1.1 Proporción de aciertos e índice de discriminación	45
5.1.2 Análisis de la puntuación total.....	46
5.1.3 Unidimensionalidad de la prueba	49
5.1.4 Confiabilidad de la prueba.....	51
5.2 Análisis de la prueba bajo el enfoque TRI.....	51
5.2.1 Estimación de parámetros y la variable latente del modelo	51
5.2.1.1 Estimación de parámetros del modelo	52
5.2.1.2 Estimación de la variable latente del modelo	57
5.2.2 Comparaciones de la habilidad por regiones naturales.....	61
5.2.3 Comparaciones de la habilidad promedio de por genero	65
5.2.4 Relaciones entre los parámetros y la variable latente	666
Conclusiones de la investigación	73
Recomendaciones para investigaciones futuras.....	76
Bibliografía	78
Anexo A.....	81
A.1 Suficiencia estadística	81

A.1.1 Teorema de Factorización de Neyman	81
A.2 Teoría Clásica de los Test	83
A.2.1 Proporción de aciertos	83
A.2.2 Índice de discriminación.....	84
A.2.3 Unidimensionalidad de la prueba.....	85
A.2.4 Confiabilidad de la prueba	86
A.3 Método de jackknife	86
A.4 Distribución de ítems	87
Anexo B.....	889
B.1 Sintaxis análisis de ítems bajo el enfoque TCT	889
B.2 Sintaxis Correlación Tetracórica	91
B.3 Script Método Jackknife.....	92
B.3 Código Modelo TRI dicotómico de dos parámetros	93

INDICE DE CUADROS

Cuadro N° 5.1: Distribución de estudiantes evaluados por región natural	43
Cuadro N° 5.2: Distribución de estudiantes evaluados por sexo	445
Cuadro N° 5.3: Análisis de ítems bajo TCT.....	456
Cuadro N° 5.4: Descriptivos del puntaje obtenido en la prueba.....	478
Cuadro N° 5.5: Descriptivos del puntaje por género	49
Cuadro N° 5.6: Varianza total explicada	501
Cuadro N° 5.7. Dificultad de ítems bajo el modelo TRI.....	52
Cuadro N° 5.8: Normalidad para la dificultad del ítem	53
Cuadro N° 5.9: discriminación de ítems bajo el modelo TRI.....	55
Cuadro N° 5.10: Normalidad para la discriminación del ítem.....	56
Cuadro N° 5.11: Estadísticos de las habilidades estimadas	58
Cuadro N° 5.12: Normalidad para la habilidad bajo el modelo TRI.....	59
Cuadro N° 5.13. Errores Jackknife por estratos de inferencia	60
Cuadro N° 5.14: Descriptivos de habilidad por regiones naturales.....	61
Cuadro N° 5.15: Prueba de homogeneidad de varianzas	63
Cuadro N° 5.16: Tabla ANOVA.....	63
Cuadro N° 5.17: Prueba HSD de Tukey.....	64
Cuadro N° 5.18: Descriptivos de la habilidad según género.....	65
Cuadro N° 5.19: Prueba T y de Levene	66
Cuadro N° A.4: Distribución de ítems según capacidades.....	87

INDICE DE GRAFICOS

Gráfico N°3.1 CCI bajo diferentes niveles de dificultad.....	31
Gráfico N°3.2: CCI bajo diferentes índices de discriminación	33
Gráfico N° 5.1: Distribución de puntajes de la prueba.....	46
Gráfico N° 5.2: Puntajes promedios por regiones naturales	47
Gráfico N° 5.3: Puntajes promedios por genero.....	48
Gráfico N° 5.4: Diagrama de cajas de la dificultad de items	53
Gráfico N° 5.5: Diagrama de cajas de la discriminación de items	54
Gráfico N° 5.6: Índice de discriminación Vs dificultad del item.....	57
Gráfico N° 5.7: Distribución de la habilidad de los estudiantes	60
Gráfico N° 5.8: Habilidad de los estudiantes por regiones naturales	62
Gráfico N° 5.9: Curvas características de items	67
Gráfico N° 5.10: Mapa de items a nivel nacional.....	68
Gráfico N° 5.11: Diagrama de dispersión puntuación vs habilidad.....	71
Gráfico N° 5.12. Diagrama de dispersión de la dificultad TC vs TRI	71
Gráfico N° 5.13: Índice de discriminación TCT vs TRI	72

Introducción

El Ministerio de Educación (MED), mediante la Dirección Nacional de Educación Secundaria, Superior y Tecnología (DINESST) presenta al país el Diseño Curricular Básico (DCB) de Educación Secundaria que orienta a las instituciones educativas estatales y no estatales. El DCB considera al estudiante como el centro de atención de la educación, tiene como máxima aspiración desarrollar en el estudiante capacidades, conocimientos, valores y actitudes que le permitan una educación integral para lograr su autorrealización (DINESST, 2004). Dicho de otro modo la aplicación adecuada del DCB de Educación Secundaria, proporciona estabilidad al trabajo docente y busca conseguir una educación con calidad y equidad.

Una forma de cuantificar que el DCB de Educación Secundaria se esté aplicando adecuadamente y medir los conocimientos adquiridos por el educando es a través de las evaluaciones. La evaluación del aprendizaje es un proceso pedagógico continuo, sistemático y participativo que forma parte del proceso de enseñanza. Inherente a la enseñanza y al aprendizaje las evaluaciones permiten observar, recoger, analizar e interpretar información relevante acerca de las necesidades, posibilidades, dificultades y aprendizaje del estudiante. A la vez la evaluación permite la certificación de la capacidad

del estudiante para el desempeño de determinadas actividades y tareas en el escenario local, regional, nacional e internacional.

Se considera como muy importante las evaluaciones de aprovechamiento o rendimiento, que son todas aquellas que buscan evaluar el nivel de habilidad o logro de un estudiante después de un proceso de instrucción (Aiken, 1996). La medición del rendimiento académico puede ser comprendida como una cantidad que estima lo que un estudiante ha aprendido como consecuencia de un proceso de instrucción o formación.

Con la investigación se busca evaluar la capacidad en comprensión lectora de los estudiantes mediante la estimación de variables ocultas con la implementación del modelo TRI dicotómico de dos parámetros o Modelo de Birnbaum. La metodología utilizada en esta investigación, se realiza desde un enfoque clásico y desde una perspectiva TRI (Teoría de Respuesta al Ítem). En el enfoque clásico se analizan las proporciones de acierto, la discriminación clásica de los ítems y la confiabilidad de la prueba. En la perspectiva TRI se estiman las variables latentes del modelo a través de métodos bayesianos bajo el esquema: adaptative rejection sampling (ARS).

Posterior a la introducción se muestra el objetivo y la justificación de la investigación. En el capítulo I: Antecedentes, se muestra los inicios de los modelos TRI hasta las últimas investigaciones, en el capítulo II se detalla de manera precisa el problema de la investigación tomando en cuenta investigaciones previas y la importancia de la comprensión lectora. En el capítulo III se describe el modelo, sus supuestos, sus propiedades y el método de estimación utilizado. En el capítulo IV se muestra la metodología de

desarrollo de la investigación. En el capítulo V se muestra el desarrollo de la aplicación y sus resultados a nivel estudiantes y de la prueba, con la investigación se muestra una jerarquización de ítem según la dificultad y discriminación los cuales sirven para determinar que el estudiante de educación secundaria de la Costa tiene mayores capacidades que los de la Sierra y la selva y que los estudiantes de género femenino tienen mejores capacidades en comprensión lectora que los estudiantes de género masculino. Finalmente se dan las conclusiones y las recomendaciones con la implementación del modelo estadístico para la evaluación de la capacidad lectora del estudiante de quinto de secundaria.

Objetivo de la investigación

Estudiar la capacidad en comprensión lectora de los estudiantes de quinto de secundaria a nivel nacional, por regiones naturales y género; considerando niveles de desempeño. Estos niveles son determinados mediante la implementación del modelo TRI dicotómico de dos parámetros con la estimación de las habilidades de los estudiantes (variable latente) y la dificultad real del Ítem (parámetro del modelo).

Justificación de la investigación

Con el modelo propuesto en la investigación se estiman variables latentes u ocultas cuando un estudiante se enfrenta a una prueba, las cuales no son estimados por los modelos clásicos. Mediante estas variables latentes se determinan niveles de desempeño y se analizan las capacidades reales que tienen los estudiantes en comprensión lectora debido a que se aíslan los factores no observables. Los resultados de esta investigación pueden ser implementados con mejoras efectivas en los métodos de evaluación de la capacidad lectora.

Capítulo I

Antecedentes

Un punto de partida para la TRI es la Teoría Clásica de los Test (TCT) que fue formulada en 1904 por Charles Spearman (Muñiz, 1996). Se le conoce también como Modelo Lineal Clásico, en la cual el puntaje observado en una medición (X), es igual al puntaje verdadero (V), más el error (e). Adicionalmente, los supuestos básicos de este modelo son:

- El puntaje verdadero es el valor esperado de las puntuaciones observadas (Muñiz, 1996).
- No existe relación entre el puntaje verdadero y el error del modelo, es decir son independientes (Muñiz, 1996).
- La correlación entre los errores de dos pruebas son iguales a 0, es decir los errores son independientes (Muñiz, 1996).

La TCT supone que las diferencias entre las respuestas de los evaluados se deben solamente a la variación del puntaje obtenido y las demás fuentes potenciales de variabilidad producida por los instrumentos de medición o las condiciones externas e internas de los evaluados se mantienen constantes. Las

puntuaciones de dos instrumentos de medición que midan el mismo rasgo latente son distintas, esto se debe a que cada instrumento tiene su conjunto de ítems y cada ítem tiene diferentes propiedades. Una de las limitaciones más importante de la TCT es que no pueden separarse las características del evaluado de las características del instrumento de medición. En la TCT el rasgo latente es expresado por el puntaje verdadero. Si la prueba ha sido difícil entonces el evaluado obtendrá baja puntuación y si ha sido fácil obtendrá alta puntuación. La dificultad del ítem en este contexto se define como la proporción de aciertos que tienen los evaluados.

La TCT y la TRI consideran que cada evaluado lleva asociado un parámetro individual, que en la TRI se denomina habilidad del evaluado (θ) y en la TCT se denomina puntaje verdadero (V), en ambos es una variable latente no observable. La diferencia principal en la TCT y la TRI, es que la TCT el rasgo latente es de tipo lineal mientras que en la TRI las relaciones pueden ser de tipo exponencial, tales como el modelo de poisson, de la ojiva normal o los modelos logísticos de 1, 2 o 3 parámetros.

El origen de los modelos TRI se da en la década de los años 50 y 60. En 1952, Federic M. Lord presenta su tesis doctoral: "*A theory of test scores*", en la cual presenta formalmente el modelo ojiva normal de dos parámetros. A este aporte se le une el modelo de la distancia latente y el modelo lineal formulados por Lazarsfeld en los años de 1950 y 1959 respectivamente. En la década siguiente se da a lugar a las investigaciones realizadas por el matemático, estadístico y psicometrista danés George Rasch; quien en 1952 tuvo que

analizar datos de seguimiento basado en la lectura de Registros. Rasch para esta investigación introduce el modelo de poisson multiplicativo, la cual busca modelar el número de errores leídos en un texto j por el i – esimo individuo (Fisher, G, 2007). Binbaum en 1968 desarrolla modelos logísticos de dos y tres parámetros. Las investigaciones de Lord, Rasch y Birnbaum constituyen los pilares de la TRI.

La teoría de respuesta al ítem (TRI) se centra fundamentalmente en las propiedades de los ítems y busca brindar una fundamentación probabilística al problema de medir constructos latentes (no observables), considerando al ítem como unidad básica de medición. La puntuación de una prueba en un modelo clásico se da como la sumatoria de respuesta a ítems individuales, mientras que la TRI utiliza el patrón de respuesta (Nunnally y Berstein, 1995). La TRI se encarga de modelar matemáticamente la relación entre las respuestas y la habilidad que tienen los estudiantes que son sometidos a una prueba, es decir que ocurre cuando un estudiante con una habilidad determinada se enfrenta a un ítem específico. El modelo TRI involucra variables latentes que explican la habilidad del estudiante y un conjunto de parámetros asociados al ítem de la prueba como la dificultad y la discriminación.

Como proponen Bartholomew y Knoop (1999), existen diferentes tipos de modelos de variables latentes. Dos de los mas importantes modelos son la teoría clásica de los test (TCT) (Lord y Baker, 1968) y la teoría de respuesta al ítem (TRI) (Baker, 1992). Hay dos objetivos principales de la TRI, uno de ellos

es evaluar las aptitudes del examinado y el segundo es estudiar la eficacia de las diferentes pruebas de rendimiento aplicado a los examinados.

Históricamente estos modelos han sido desarrollados en el ámbito de la Psicometría, pero actualmente la TRI tiene un creciente desarrollo en las áreas de Estadística, Computación y Probabilidad debido a la importante cantidad de problemas que se plantean con el desarrollo y aplicaciones de estos modelos. La importancia del estudio de modelos TRI subyace también en sus múltiples aplicaciones crecientes en campos tan diversos como la Educación, la Psicología, la Política, la Economía, la Sociología, el Marketing, la Ingeniería Genética entre otras.

Una gran ventaja de estos modelos es que dan la posibilidad de tener mediciones invariantes respecto de los instrumentos utilizados y las personas evaluadas (Muñiz, 1977). A parte de la invarianza, hay dos requerimientos importantes en los ítems que constituyen una prueba que son: la unidimensionalidad y la independencia local. La unidimensionalidad se refiere que cuando los ítems del instrumento de medición miden un único rasgo latente o constructo. Wright y Linacre (1998) señalan que en la práctica ningún instrumento de medición puede ser perfectamente unidimensional. Lo importante es que un instrumento de medición represente con sus puntuaciones un solo factor dominante. Con esto se quiere lograr que la mayor cantidad de varianza observada en las respuestas a los ítems sea explicada por un solo rasgo latente (Embreston y Reise, 2000). Es muy importante tener un instrumento unidimensional, dado que es un requisito indispensable para

generar buenas estimaciones (Wright y Master, 1982; Wright y Stone 1998). Una técnica estadística multivariada para evaluar la dimensionalidad de un conjunto de ítems es el análisis factorial (AF) (Muñiz, 1997). En diferentes instrumentos de medición tenemos respuestas dicotómicas o politómicas, en este tipo de datos no podemos utilizar la matriz de correlación de Pearson en el análisis factorial para ello tenemos que considerar la matriz de correlaciones tetracóricas (si la data es dicotómica), policóricas (si la data es nominal) y la Spearman (si la data es ordinal). Así un conjunto de ítems será unidimensional si el primer factor explica entre el 20% y 40% de la varianza, utilizando correlaciones tetracóricas (Zwick, 1985). Existen otros métodos más recientes para detectar la unidimensionalidad como DETECT (Zhang y Stout, 1907), el análisis factorial no lineal propuesto McDonald e implementado en el software NOHARM (Fraser y McDonald, 1988), el método propuesto por Muthén e implementado en el programa LISCOMP (Muthén, 1987), y el análisis factorial de información completa propuesto por Muraki e implementado en los programas TESTFACT (Wilson, Wood, y Gibbons, 1993) y POLYFACT (Muraki, 1993).

La independencia local se da cuando las respuestas del evaluado a cualquier par de ítems son independientes. De esta forma de responder a un conjunto de ítems es igual al producto de probabilidades a cada ítem individual.

Los modelos TRI se pueden clasificar en:

1.- Por el número de parámetros:

- De un parámetro, más conocido como Modelo de Rasch la cual tiene como variable latente la habilidad del evaluado y el parámetro de la dificultad del ítem.

- De dos parámetros, conocido como el modelo de Birnbaum., que tienen también como variable latente a la habilidad del estudiante, pero tiene dos parámetros que es la dificultad y discriminación del ítem.
- De tres parámetros, estos modelos presentan tres parámetros como son la dificultad del ítem, la discriminación del ítem y el factor de adivinación o azar por parte del evaluado al responder a un ítem.

2.- Por el tipo de función utilizada

- Normal, son aquellos que suponen una distribución normal
- Logístico, aquellos que suponen una distribución logística

3.- Por el número de dimensiones

- Unidimensional, cuando se evalúa un solo rasgo latente
- Multidimensional, cuando se evalúa un conjunto de rasgos latentes

4.- Por el formato de calificación del ítem

- Dicotómicos, están los modelos de un parámetro o de Rasch, modelos de dos parámetros o Birnbaum y el modelo de tres parámetros.
- Politémicos, dentro de estas están el modelo de créditos parciales, el modelo de créditos parciales generalizado, el modelo rating scale, el modelo de respuesta graduada o de Samejima, el modelo de respuesta graduada modificada o de Muraki y el modelo de respuesta nominal o de Bock.

La explosión de estas investigaciones se da en principales revistas científicas como, por ejemplo, el *Journal of Educational Measurement* (1977, Vol. 14, Nº. 2), *Applied Psychological Measurement* (1982, Vol. 6, Nº. 4) o la *International Journal of Educational Research* (1989, Vol. 13, Nº. 2), entre otras. Asimismo se publicaron manuales y monografías de Andrich (1988), Baker (1985),

Hambleton y Swaminathan (1985), Hambleton, Swaminathan y Rogers (1991), Hulin, Drasgow y Parsons (1983), Lord (1980), Weiss (1983) y el de Fisher y Molenaar (1995), sobre los últimos desarrollos con el modelo de Rasch, y el de Vander Linden y Hambleton (1997) sobre los últimos avances en general, y el de Embretson y Hershberger (1999) sobre las nuevas reglas de medición.

Actualmente los modelos TRI han servido para el desarrollo de otras investigaciones y planteamientos de nuevos modelos como:

- Los modelos TRI con enlace asimétrico o “A Skew Ítem Response Model” (Bazan, J; D. Branco M. y Bolfarine H. (2006)). Este modelo es utilizado cuando las puntuaciones presentan distribuciones asimétricas.
- El modelo de ojiva normal asimétrico, en la cual este modelo supone que la probabilidad de responder correctamente a un ítem sigue una distribución skew normal.
- Los modelos LPE y RLPE, el modelo LPE o Modelo Logístico Positivo Exponencial fue propuesto por Samejima (2000) pero ha sido implementado por Bolfarine y Bazan (2007) junto al modelo RPLE o reflejo del modelo logístico positivo exponencial propuestos por ellos.
- Los modelos TRI bajo restricciones de rapidez, estos modelos se utilizan cuando el instrumento de medición tiene restricciones de tiempo. La posibilidad de construir un modelo que tome en cuenta estas restricciones fue considerada por Yanamoto y Everson (1997) .

Capítulo II

El problema de la investigación

En estudios de investigación a nivel internacional desarrollados por la UNESCO muestran que el Perú en comprensión lectora se ubica debajo del promedio del total de países participantes (Flores M., E.F. (2010)). Debido a los resultados mostrados en estas investigaciones, es que a partir del año 1996 el Ministerio de Educación (MED), evalúa a los estudiantes peruanos mediante evaluaciones muestrales y censales.

En la evaluación nacional 2004, en la prueba de comunicación, se mostró que el 9.8% de los estudiantes de quinto de secundaria demuestran un dominio suficiente en el desempeño evaluado para el grado, el 45.3% demuestran un dominio incipiente del desempeño en comprensión lectora en el grado evaluado, el 14.8% tienen capacidad del grado anterior (cuarto de secundaria) y el 30.1% solo pueden resolver una que otra pregunta de la prueba.

Los resultados de las evaluaciones nacionales e internacionales muestran que la mayoría de la población estudiantil del Perú y de varios países

latinoamericanos no alcanzan los estándares requeridos en competencias básicas de aprendizaje.

La comunicación es uno de los elementos más importantes para el desarrollo personal y social. En el Perú que está caracterizado por la diversidad cultural, la comunicación se convierte, además, en un requisito indispensable para la convivencia armoniosa, basada en el reconocimiento y valoración de las diferencias. De allí que el desarrollo de las capacidades comunicativas es una tarea primordial de la educación. Igualmente podemos decir que la comprensión lectora es la base para el desarrollo de las capacidades de los estudiantes en esta área pedagógica.

Debido a la importancia por la comprensión lectora y a los resultados de investigaciones previas, con la investigación se busca evaluar el rendimiento del estudiante de quinto de secundaria de las instituciones educativas estatales mediante la estimación de variables latentes o variables no observables presentes cuando el estudiante se enfrenta a una prueba. Nos referimos a la estimación de la habilidad y discriminación del ítem de la prueba, mediante la utilización de modelos de TRI dicotómicos de dos parámetros o modelos de Birnbaum. Estos factores y parámetros sirven de variables de entrada en la determinación de los niveles de desempeño, con estos niveles se analizarán la capacidad en comprensión lectora que tiene el estudiante en comprensión lectora.

Con los métodos clásicos no podemos establecer niveles de desempeño dado que no se puede determinar la habilidad y dificultad real que tiene el estudiante al rendir una prueba de rendimiento.

Capítulo III

Modelos TRI Dicotómicos de dos parámetros o Modelos de Birnbaum

El origen de los modelos TRI puede localizarse en los años 1950, pero es en la década de los 60 donde surge la explosión de estos modelos con las investigaciones del psicometrista, matemático y estadístico George Rasch.

El objetivo de la TRI es dar fundamentación probabilística al problema de medición de variables latentes o variables no observables. Se considera al ítem como unidad básica del instrumento de medición. Estos modelos son funciones matemáticas en la cual se relacionan las probabilidades de una respuesta a un ítem con la habilidad que tiene el evaluado.

En investigaciones de origen cuantitativo generalmente se emplean instrumentos de medición las cuales generan bases de datos dicotómicas, donde los evaluados responden a preguntas del instrumento de medición. Asociadas a estas respuestas existen variables latentes que son características de los individuos θ , y los parámetros del modelo como la dificultad de ítem β y discriminación del ítem α .

Definimos un modelo latente como el par ordenado (\mathbf{Y}, \mathbf{U}) , donde la secuencia aleatoria $Y=(Y_1, Y_2, Y_3, \dots)$ y el vector aleatorio $U=(U_1, U_2, U_3, \dots, U_n)$ contienen variables de entrada y latentes respectivamente. Una variable es considerada latente cuando están presentes, pero no son observables. En general los modelos de variables latentes se formulan con el propósito de conocer las principales características de \mathbf{U} en base al conocimiento de \mathbf{Y} .

Un modelo latente (Y, U) se dice que satisface:

- La independencia condicional, si la distribución condicional de cualquier subconjunto de Y dado U , puede expresarse con el producto de las distribuciones marginales condicionales de Y dado U . En este caso diremos que el modelo es condicionalmente independiente.
- La monotocidad, si:

$$1 - P(Y_i \leq y | U = u) = P(Y_i > y | U = u)$$

Son funciones no decrecientes de $u \in \mathfrak{R}^n$, para cualquier $y \in \mathfrak{R}$, e $i \in N^+$. En este caso diremos que el modelo es monótono.

- La unidimensionalidad, si $U=(U_1, U_2, U_3, \dots, U_n)$ para $(n=1)$ se dice que es unidimensional.

Un modelo latente (Y, U) se denomina condicionalmente monótono si satisface la independencia local y la monotocidad. Adicionalmente si el modelo es unidimensional, se dice que el modelo es condicionalmente monótono unidimensional.

Los postulados básicos de la TRI son los siguientes:

- El resultado de un evaluado en un ítem puede ser explicado por un conjunto de factores llamados rasgos latentes.

- La relación entre la respuesta de un evaluado a un ítem y el rasgo latente puede describirse como una función monótona creciente que se llama Curva Característica del Ítem (CCI), con esta función se muestra que a medida que aumenta la habilidad aumenta la probabilidad de responder correctamente al ítem.
- La invariancia de los parámetros del modelo TRI, las estimaciones de las habilidades de los evaluados obtenidas con distintos ítems serían iguales y las estimaciones de los parámetros de los ítems obtenidos con distintas muestras de evaluados serían también iguales. Es decir que en la TRI los parámetros de habilidad y de los ítems son invariantes. Esta propiedad de invarianza se obtiene incluyendo información de los evaluados e ítems en el proceso de estimación de sus parámetros.

En la TRI existen una gama de modelos, entre ellos tenemos los dicotómicos y politómicos. Dentro de los dicotómicos están los modelos de un parámetro o modelos de Rasch, los modelos de dos parámetros o modelos de Birnbaum y los modelos de tres parámetros. Dentro de los politómicos tenemos al modelo de créditos parciales, el modelo de créditos parciales generalizado, el modelo rating scale, el modelo de respuesta graduada o modelo de Samejima, el modelo de respuesta nominal, y los modelos de facetas múltiples.

Otra gama son los modelos asimétricos con enlaces a familias de distribuciones.

En la investigación desarrollamos los modelos TRI dicotómicos de dos parámetros o modelos de Birnbaum.

3.1 Formulación del Modelo TRI dicotómico de dos parámetros

El modelo TRI Dicotómico de dos parámetros o modelos de Birnbaum modela la probabilidad de que el estudiante responda correctamente a un ítem de la prueba, en la cual se estiman variables latentes como la habilidad del estudiante, y los parámetros del modelo como la dificultad y discriminación del ítem.

El modelo TRI de dos parámetros queda definido de la siguiente forma:

$$(Y_{ij}|\theta_i, \beta_j, \alpha_j) \sim \text{Beronulli}(p_{ij}) \quad (3.1.1)$$

$$p_{ij} = F[\alpha_j(\theta_i - \beta_j)]$$

$$p_{ij} = \frac{e^{y_{ij}\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \quad (3.1.2)$$

Donde:

- Y_{ij} es la respuesta del i – esimo estudiante al j – esimo ítem de la prueba, para todo $i=1\dots n$ y $j=1\dots k$.
- α_j y β_j son la discriminación y dificultad del j – esimo ítem (parámetros del modelo) para todo $\alpha_j \in \mathfrak{R}$ y $\beta_j \in \mathfrak{R}$.
- θ_i es la habilidad del i – esimo estudiante (variable latente) para todo $\theta_i \in \mathfrak{R}$.
- p_{ij} es la probabilidad condicionado que el i – esimo estudiante responda el j – esimo ítem de la prueba.
- $F(\cdot)$ es la distribución acumulada o curva característica del ítem (CCI).

La probabilidad de responder correctamente al ítem esta dado por:

$$p(Y_{ij} = 1|\theta_i, \beta_j, \alpha_j) = \frac{e^{\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \quad (3.1.3)$$

La probabilidad de fallar o responder incorrectamente al ítem es dado por:

$$p(Y_{ij} = 0|\theta_i, \beta_j, \alpha_j) = \frac{1}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \quad (3.1.4)$$

El Modelo TRI dicotómico de dos parámetros es una generalización del Modelo de Rasch, este modelo asume la existencia de discriminación de cada ítem. La discriminación del ítem es el cambio de la probabilidad de responder correctamente al ítem por cada variación del nivel de habilidad del estudiante, dicho de otro modo es la pendiente de la curva característica del ítem bajo cierto nivel de dificultad fijo.

Supuestos del Modelo TRI de dos parámetros o Modelo de Birnbaum

- 1) Del modelo planteado en la ecuación (3.1.2) observamos que es unidimensional, dado que hay la presencia de una única variable latente. Es decir el modelo es unidimensional.
- 2) De la curva característica del ítem observamos que es una función no decreciente de θ . El modelo es monótono creciente.
- 3) Para valores fijos de la variable latente θ todas las respuestas son estocásticamente independientes, así la probabilidad de responder a los K ítems de la prueba es el producto de las k probabilidades. Dicho de otra manera:

$$P(Y_i|\theta_i, \alpha_j, \beta_j) = P(Y_{i1}|\theta_i, \alpha_1, \beta_1)P(Y_{i2}|\theta_i, \alpha_2, \beta_2) \dots P(Y_{ik}|\theta_i, \alpha_k, \beta_k)$$

$$P(Y_i|\theta_i, \alpha_j, \beta_j) = \prod_{j=1}^k P(Y_i|\theta_i, \alpha_j, \beta_j)$$

$$P(Y_i|\theta_i, \alpha_j, \beta_j) = \prod_{j=1}^k \frac{e^{\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}}$$

3.2 Función de Verosimilitud del modelo TRI dicotómico de dos parámetros

Teniendo en cuenta el supuesto de independencia de los ítems o independencia local, se plantea la función de verosimilitud del modelo:

Sea $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ y $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$:

$$L(Y_i|\theta_i, \alpha, \beta) = \prod_{j=1}^k [F_Y(\theta_i, \alpha_j, \beta_j)]^{y_{ij}} [1 - F_Y(\theta_i, \alpha_j, \beta_j)]^{1-y_{ij}}$$

Reemplazando con la función logística:

$$L(Y_i|\theta_i, \alpha, \beta) = \prod_{j=1}^k \left[\frac{e^{\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \right]^{y_{ij}} \left[\frac{1}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \right]^{1-y_{ij}}$$

$$L(Y_i|\theta_i, \alpha, \beta) = \prod_{j=1}^k \left[\frac{e^{y_{ij}\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}} \right]$$

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\sum_{j=1}^k y_{ij}\alpha_j(\theta_i - \beta_j)}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]}$$

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\theta_i \sum_{j=1}^k y_{ij}\alpha_j - \sum_{j=1}^k y_{ij}\alpha_j\beta_j}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]} \quad (3.2.1)$$

Finalmente la función de verosimilitud para todos los estudiantes evaluados queda expresada de la siguiente forma:

Sean $Y = (Y_1, Y_2, \dots, Y_n)$ y $Y_i = (y_{i1}, y_{i2}, \dots, y_{ik})$ en la cual $Y_{(n \times k)}$ es una matriz $n \times k$ que contiene las respuestas de los n estudiantes evaluados con lo k ítems de la prueba, $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ las habilidades de los estudiantes (variable latente),

$\beta = (\beta_1, \beta_2, \dots, \beta_k)$ dificultad de los ítems (parámetro) y $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$

discriminación de los ítems (parámetro)

$$L(Y|\theta, \beta, \alpha) = \prod_{i=1}^n L(Y_i|\theta_i, \alpha, \beta)$$

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]} \quad (3.2.2)$$

Hacemos:

$$r_i = \sum_{j=1}^k y_{ij} \alpha_j \quad (3.2.3)$$

En la data r_i es la puntuación obtenida ponderada por la discriminación del j -ésimo ítem que obtiene el i -ésimo estudiante evaluado.

$$c_j = \sum_{i=1}^n y_{ij} \quad (3.2.4)$$

En la data c_j es la puntuación obtenida en el j -ésimo ítem que obtienen todos los estudiantes evaluados.

Reemplazando las ecuaciones (3.2.3) y (3.2.4) en la ecuación (3.2.2) obtenemos:

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\sum_{i=1}^n r_i \theta_i - \sum_{j=1}^k c_j \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]} \quad (3.2.5)$$

Tomando logaritmos neperianos:

$$l(Y_i|\theta_i, \alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \beta_j - \sum_{i=1}^n \sum_{j=1}^k \ln[1 + e^{\alpha_j (\theta_i - \beta_j)}] \quad (3.2.6)$$

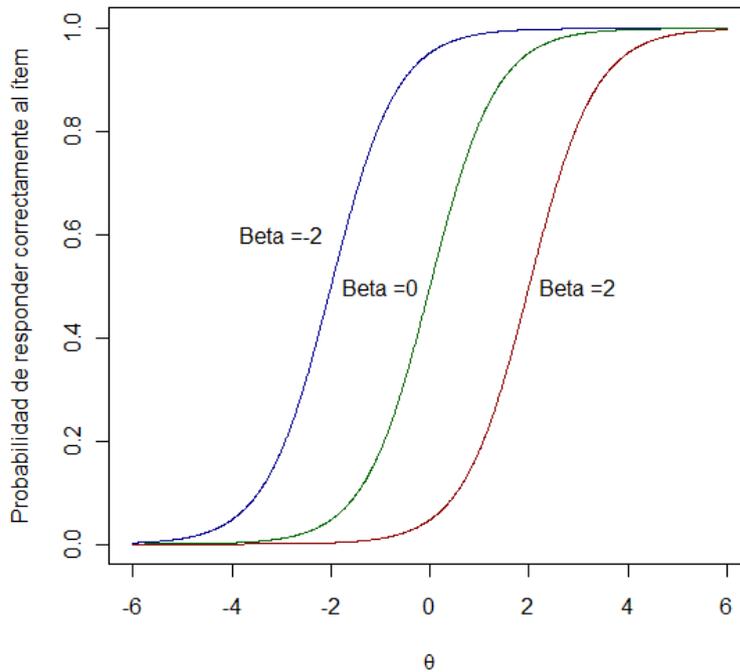
$$l(Y_i|\theta_i, \alpha, \beta) = \sum_{i=1}^n r_i \theta_i - \sum_{j=1}^k c_j \alpha_j \beta_j - \sum_{i=1}^n \sum_{j=1}^k \ln[1 + e^{\alpha_j (\theta_i - \beta_j)}] \quad (3.2.7)$$

3.3 Caracterización y propiedades del modelo TRI dicotómico de dos parámetros

De la ecuación 3 graficamos la curva característica del ítem:

Grafico N°3.1

Curva Característica del Ítem bajo diferentes niveles de dificultad



FUENTE: Chincaro Del Coral, Omar Antonio (2010)
ELABORACION: Propia

Del grafico 3.1 se observa las siguientes propiedades:

- Bajo cierta discriminación fija, observamos que las curvas características son crecientes. Es decir si $\theta_1 < \theta_2$ entonces $P(Y_{ij} = 1|\theta_1, \alpha, \beta_j) < P(Y_{ij} = 1|\theta_2, \alpha, \beta_j)$.
- Bajo cierto nivel de habilidad (θ) y discriminación (α), si los β aumentan la probabilidad de responder correctamente disminuye. Por lo

tanto a mayor β el ítem es considerado mas difícil caso contrario el ítem es considerado fácil.

- La curva característica del ítem refleja una función de probabilidad , es decir se cumple que:

$$P(Y_{ij} = 1|\theta_i, \alpha_j, \beta_j) + P(Y_{ij} = 0|\theta_i, \alpha_j, \beta_j) = 1.$$

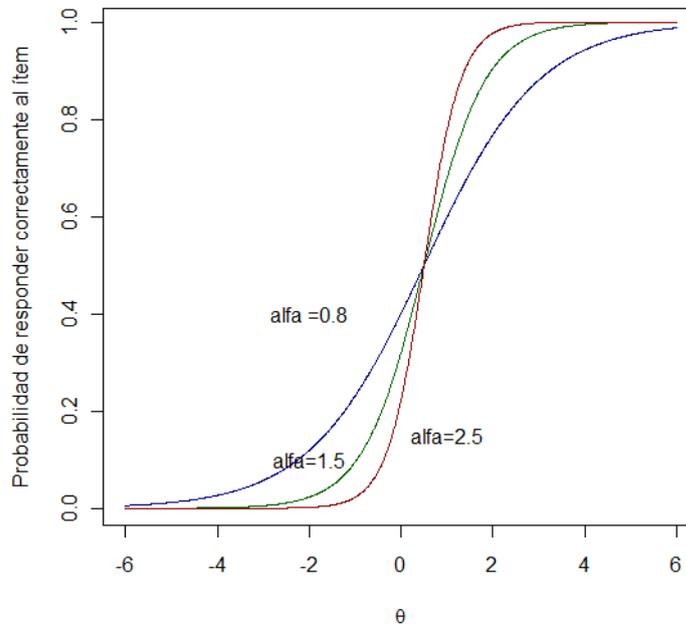
$$0 \leq P(Y_{ij} = 1|\theta_i, \alpha, \beta_j) \leq 1$$

$$\lim_{\theta \rightarrow -\infty} P(Y_{ij} = 1|\theta_i, \alpha, \beta_j) = 0 \text{ y } \lim_{\theta \rightarrow \infty} P(Y_{ij} = 1|\theta_i, \alpha, \beta_j) = 1$$

- De las ecuaciones 7, 8 y 9 se demuestra que r_i y c_j son estadísticas suficientes de θ_i y β_j respectivamente (ver anexo A).
- De la ecuación 3, si $\alpha < 0$ la curva característica del ítem es decreciente con pendiente negativa. Si $\alpha = 0$, para cualquier nivel de habilidad la probabilidad de responder correctamente al ítem siempre es 0.5. El ítem funciona bien siempre y cuando la discriminación es positiva ($\alpha > 0$).

Grafico N°3.2

Curva Característica del Ítem bajo diferentes índices de discriminación y $\beta = 0.5$



FUENTE: Chincaro Del Coral, Omar Antonio (2010)
ELABORACION: Propia

Del grafico 3.2 se observan dos escenarios:

- En el caso de que la habilidad (θ) sea menor a la dificultad (β): observamos que bajo cierto nivel de habilidad fijo, cuando la discriminación del ítem aumenta la probabilidad de responder correctamente al ítem disminuye y son menores a 0.5, es decir el ítem se hace más difícil.
- En el caso de que la habilidad (θ) sea mayor a la dificultad (β): observamos que bajo cierto nivel de habilidad fijo cuando la discriminación del ítem aumenta la probabilidad de responder correctamente al ítem aumenta y son mayores a 0.5, es decir el ítem se hace más fácil.

3.4 Estimación de variables latentes y parámetros del modelo

Para estimar las variables latentes y los parámetros del modelo existen diferentes métodos de estimación que son vistos desde una perspectiva clásica y bayesiana.

Los métodos de estimación clásica toman como punto de partida la función de verosimilitud del modelo. Desde una perspectiva clásica tenemos:

- Métodos de Máxima Verosimilitud Conjunta (Baker 1992 y Wright 1998).
- Métodos de Máxima Verosimilitud Marginal (Fischer, 2007), (Baker, 1992), (Kim, 2001).
- Métodos de Máxima Verosimilitud Condicional (Fischer, 2007), (Baker, 1992), (Kim, 2001).

En la investigación desarrollamos el método de estimación bayesiana

3.4.1 Método de estimación bayesiana para el modelo TRI dicotómico de dos parámetros o modelo de Birnbaum

Para la estimación bayesiana del modelo debemos especificar las distribuciones a prioris de las habilidades (θ) de los evaluados y los parámetros de los ítems: la dificultad (β) y la discriminación (α).

Considerando la densidad a priori conjunta entonces la densidad a posteriori es proporcional a:

$$P(\theta, \beta, \alpha|Y) \propto L(Y|(\theta, \beta, \alpha))P(\theta, \beta, \alpha) \quad (3.4.1)$$

Teniendo en cuenta a la función de verosimilitud del modelo como:

$$l(Y_i|\theta_i, \alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \beta_j - \sum_{i=1}^n \sum_{j=1}^k \ln[1 + e^{\alpha_j(\theta_i - \beta_j)}]$$

Asumiendo que las distribuciones a priori de las variables latentes y los parámetros del modelo tienen una distribución normal: $\theta_i \sim N(0, 1)$, $\beta_j \sim N(0, \sigma_\beta^2)$ y $\alpha_j \sim LN(0, \sigma_\alpha^2)$.

Teniendo en cuenta la distribución conjunta de las variables latentes y parámetros del modelo como:

$$P(\theta, \beta, \alpha) = \prod_{i=1}^n P(\theta_i) \prod_{j=1}^k P(\beta_j) \prod_{j=1}^k P(\alpha_j) \quad (3.4.2)$$

Como $\theta_i \sim N(0, 1)$ entonces:

$$P(\theta_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta_i^2}{2}}$$

$$\prod_{i=1}^n P(\theta_i) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{\sum_{i=1}^n \theta_i^2}{2}} \quad (3.4.3)$$

Como $\beta_j \sim N(0, \sigma_\beta^2)$ entonces:

$$P(\beta_j) = \frac{1}{\sigma_\beta \sqrt{2\pi}} e^{-\frac{\beta_j^2}{2\sigma_\beta^2}}$$

$$\prod_{j=1}^k P(\beta_j) = \frac{1}{(2\pi)^{\frac{k}{2}} \sigma_\beta^k} e^{-\frac{\sum_{j=1}^k \beta_j^2}{2\sigma_\beta^2}} \quad (3.4.4)$$

Como $\alpha_j \sim LN(0, \sigma_\alpha^2)$ entonces:

$$P(\alpha_j) = \frac{1}{\sigma_\alpha \alpha_j \sqrt{2\pi}} e^{-\frac{(\ln \alpha_j)^2}{2\sigma_\alpha^2}}$$

$$\prod_{j=1}^k P(\alpha_j) = \frac{1}{(2\pi)^{\frac{k}{2}} \sigma_\alpha^k \prod_{j=1}^k \alpha_j} e^{-\frac{\sum_{j=1}^k (\ln \alpha_j)^2}{2\sigma_\alpha^2}} \quad (3.4.5)$$

Reemplazando las ecuaciones (3.3.3), (3.3.4), (3.3.5) en (3.3.2) y resolviendo:

$$P(\theta, \beta, \alpha) = \frac{1}{(2\pi)^{\frac{n+2k}{2}} (\sigma_\beta \sigma_\alpha)^k \prod_{j=1}^k \alpha_j} e^{-\frac{1}{2} \left[\sum_{i=1}^n \theta_i^2 + \sum_{j=1}^k \left(\frac{\beta_j}{\sigma_\beta}\right)^2 + \sum_{j=1}^k \left(\frac{\ln \alpha_j}{\sigma_\alpha}\right)^2 \right]}$$

$$P(\theta, \beta, \alpha) \propto \frac{1}{\prod_{j=1}^k \alpha_j} e^{-\frac{1}{2} \left[\sum_{i=1}^n \theta_i^2 + \sum_{j=1}^k \left(\frac{\beta_j}{\sigma_\beta}\right)^2 + \sum_{j=1}^k \left(\frac{\ln \alpha_j}{\sigma_\alpha}\right)^2 \right]} \quad (3.4.6)$$

Reemplazamos las ecuaciones (3.2.5) y (3.4.6) en (3.4.1) y obtenemos la densidad a priori conjunta:

$$P(\theta, \beta, \alpha | Y) \propto \frac{e^{\sum_{i=1}^n r_i \theta_i - \sum_{j=1}^k c_j \alpha_j \beta_j - \frac{1}{2} \left[\sum_{i=1}^n \theta_i^2 + \frac{\sum_{j=1}^k \beta_j^2}{\sigma_\beta^2} + \frac{\sum_{j=1}^k (\ln \alpha_j)^2}{\sigma_\alpha^2} \right]}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}] \prod_{j=1}^k \alpha_j} \quad (3.4.7)$$

Para todo $r_i = \sum_{j=1}^k y_{ij} \alpha_j$ y $c_j = \sum_{i=1}^n y_{ij}$

De la ecuación (3.4.7) obtenemos la distribución condicional completa de θ :

$$P(\theta_i | \beta, \alpha, Y) \propto \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{j=1}^k \sum_{i=1}^n y_{ij} \alpha_j \beta_j - \frac{1}{2} \sum_{i=1}^n \theta_i^2}}{\prod_{i=1}^n [1 + e^{\alpha_j (\theta_i - \beta_j)}]} \quad (3.4.8)$$

De la ecuación (3.4.7) obtenemos la distribución condicional completa de β :

$$P(\beta_j | \theta, \alpha, Y) \propto \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{j=1}^k \sum_{i=1}^n y_{ij} \alpha_j \beta_j - \frac{1}{2\sigma_\beta^2} \sum_{j=1}^k \beta_j^2}}{\prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]} \quad (3.4.9)$$

De la ecuación (3.4.7) obtenemos la distribución condicional completa de α :

$$P(\alpha_j | \theta, \beta, Y) \propto \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{j=1}^k \sum_{i=1}^n y_{ij} \alpha_j \beta_j - \frac{1}{2\sigma_\alpha^2} \sum_{j=1}^k (\ln \alpha_j)^2}}{\prod_{j=1}^k \alpha_j} \quad (3.4.10)$$

De las distribuciones condicionales completas observamos que no son estándares es por ello que es difícil implementar un muestreador Gibbs. Para

estimar las variables latentes y los parámetros daremos uso a los métodos MCMC.

3.4.1.1 MCMC usando adaptative rejection sampling (ARS)

El modelo TRI dicotómico de dos parámetros o modelo de Birnbaum puede ser ajustado usando métodos MCMC. El logaritmo de las densidades condicionales para cada uno de los parámetros es cóncavos (Sahu, 2001), el número exacto de ellas es utilizando rejection sampling (Wilks y Wald, 1982). El ARS trabaja dinámicamente dos envolventes (una superior y una inferior) para la distribución a ser muestreada de evaluaciones sucesivas de la densidad en los puntos rechazados. El algoritmo se detiene cuando el punto propuesto ha sido aceptado. Este algoritmo viene implementado por defecto en el software WinBugs para el modelo TRI dicotómico de dos parámetros. Ver el código del programa utilizado en el Anexo B.

Capítulo IV

Hipótesis y Metodología de la investigación

4.1 Tipo de investigación

La investigación que desarrollamos es de tipo exploratorio, descriptivo y explicativo. Es exploratorio porque investigamos y analizamos las habilidades que tienen los estudiantes y la discriminación de ítem (variables latentes) al momento de enfrentarse a una prueba de comprensión lectora. Es descriptivo, por que a través de una prueba analizamos el comportamiento de los estudiantes y ver que tan difícil son los ítems de la prueba. Y es explicativo por que a través de un modelo estadístico se explica la probabilidad de responder correctamente a un ítem de la prueba en función a la habilidad que tienen los estudiantes, y a los parámetros del modelo como la dificultad y discriminación de los ítems del instrumento de medición.

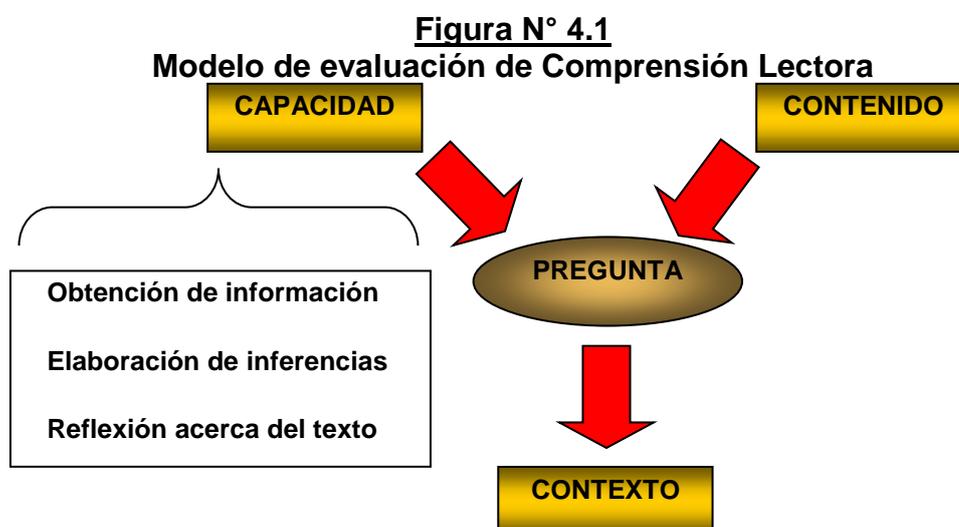
4.2 Hipótesis de investigación

Vinculado a las respuestas de una prueba existen variables latentes: la habilidad que tiene el estudiante; y los parámetros del modelo: dificultad y discriminación del ítem, las cuales van a determinar las capacidades que tiene el estudiante de quinto de secundaria.

4.3 Instrumento de medición

Para el desarrollo de la investigación se considera como instrumento de medición el cuadernillo 1 de la prueba de comprensión lectora utilizada por el Ministerio de Educación (MED) en el proyecto “Movilización Nacional por la Comprensión Lectora” (ver Anexo C). La prueba consta de 5 textos y 25 preguntas que busca evaluar las capacidades de obtener información explícita del texto, hacer inferencias y reflexionar acerca del texto leído en la prueba. El texto 1, texto 2 y texto 3 engloba 5 ítems cada texto, el texto 4 consta de 6 ítems y el texto 5 tiene 4 ítems.

Con los ítems de la prueba se busca conseguir el siguiente modelo de evaluación:



FUENTE: Ministerio de educación, 2004

4.4 Tipos de variables para el análisis

Cada ítem de la prueba consta de 4 alternativas, de las cuales 1 es correcta y 3 incorrectas (ver Anexo C las preguntas de la prueba). Si la respuesta es correcta se le asigna valor 1, caso contrario valor 0. Con esto el tipo de variables ingresadas al modelo son dicotómicas.

4.5 Diseño muestral y Muestra

En la investigación se considero a todos los estudiantes de quinto de secundaria evaluados con el cuadernillo 1 de la prueba de comprensión lectora, es decir con este cuadernillo se evaluaron a 1223 estudiantes.

La prueba utilizada en la investigación ha sido parte del proyecto: Movilización Nacional por la Comprensión Lectoral, en lo cual se evaluaron a los estudiantes de los 5 grados de educación secundaria de las instituciones educativas estatales de todo el Perú (DIGBER, 2009).

4.5.1 Población objetivo

Está formada por los estudiantes de 5to secundaria de las instituciones educativas estatales existentes a nivel nacional.

4.5.2 Unidades de muestreo

En la investigación las unidades de muestreo son: las instituciones educativas estatales que albergan a los estudiantes de quinto de secundaria de todas las regiones del Perú.

4.5.3 Unidad de análisis

La unidad de análisis es un estudiante de quinto de secundaria de una institución educativa estatal seleccionada.

4.5.4 Diseño muestral

El tipo de muestreo utilizado es un muestreo complejo de varias etapas. En la primera etapa se seleccionaron instituciones educativas proporcionales a la cantidad de estudiantes. En la segunda etapa, dentro de cada institución educativa se escogieron secciones o aulas a evaluar. Y en la última etapa se seleccionaron sistemáticamente con arranque aleatorio a los estudiantes a evaluarse.

4.6 Recolección de datos y base de datos

Para formar la base de datos del proyecto de investigación se utilizaron fichas ópticas, en la cual los estudiantes evaluados marcan sus respuestas. Para minimizar errores al ingreso de los datos, en la ficha óptica ya está predefinido la región, la provincia, el distrito, nombre de institución educativa, el código de la institución educativa y el grado a ser evaluado.

Con lectura de las fichas ópticas se obtienen la base de datos de respuestas de los estudiantes evaluados. Posteriormente se realiza el control de calidad y validación de la data con las pruebas (las pruebas contienen todas las respuestas de los estudiantes) y las listas de asistencias. Las listas de asistencia son documentos en las cuales se registran nombre y apellidos de los estudiantes a ser evaluados, el tipo de cuadernillo que rindieron, la institución

educativa, la región, la provincia, el distrito, la sección evaluada y el total de pruebas tomadas en esa sección (DIGBER, 2009).

Finalmente se califica la prueba: si la respuesta es correcta se codifica como 1 y si es incorrecto se codifica como 0, obteniéndose de esta manera una base de datos dicotómica para su posterior análisis.

4.7 Análisis estadístico

Se realiza bajo el enfoque clásico y TRI, en el enfoque clásico se analizarán las proporciones de acierto de cada pregunta de la prueba, la discriminación del ítem desde un punto clásico y la confiabilidad de la prueba utilizada. Considerando la metodología TRI, se aplicará el modelo propuesto en la investigación estimándose las variables latentes, ver esquema de modelamiento estadístico en el anexo A.

Se realizan inferencias a nivel nacional, por regiones naturales y género de los estudiantes evaluados. En la Costa se consideran los siguientes departamentos: Tumbes, Piura, La Libertad, Lambayeque, Ancash, Lima, Callao, Ica, Arequipa, Moquegua y Tacna. Para la Sierra se tienen los siguientes departamentos: Cajamarca, Huanuco, Pasco, Junín, Huancavelica, Ayacucho, Apurímac, Cusco y Puno. Y en la selva: Loreto, Amazonas, San Martín, Ucayali y Madre de Dios.

Capítulo V

Aplicación y resultados de la investigación

La aplicación de la investigación está orientada en el ámbito educacional, en la cual con una prueba de comprensión lectora se evalúan a estudiantes de quinto de secundaria de instituciones educativas estatales a nivel nacional. Mediante las respuestas de los estudiantes a la prueba y con la implementación del modelo TRI dicotómico de dos parámetros, se estiman variables latentes (habilidad de los estudiantes) y los parámetros del modelo implementado (dificultad y discriminación de los ítems de la prueba) a través métodos de estimación bayesiana.

Cuadro N° 5.1

Distribución de estudiantes evaluados por región natural

Región	Tamaño de la Muestra	%
COSTA	621	50.80%
SIERRA	406	33.20%
SELVA	196	16.00%
TOTAL	1223	100.00%

FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.2

Distribución de estudiantes evaluados por género

Género	Tamaño de la Muestra	%
FEMENINO	729	59.61%
MASCULINO	494	40.39%
TOTAL	1223	100.00%

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro 5.1 Y 5.2 observamos la distribución de los estudiantes evaluados por estratos de inferencias: por región natural y género.

5.1 Análisis de la prueba bajo el enfoque TCT

Según Muñiz (1998), para analizar los ítems y la prueba se consideran criterios a nivel ítems y a nivel de la prueba. A nivel de ítems se considera proporciones de aciertos e índice de discriminación. Y a nivel de la prueba se tiene en cuenta el análisis de la puntuación total, la unidimensionalidad y confiabilidad de la prueba (ver la metodología de la TCT en el anexo A).

A nivel de ítems

5.1.1 Proporción de aciertos e índice de discriminación

Cuadro N° 5.3

Análisis de ítems bajo la TCT

Ítem	% aciertos	Discriminación	Estadístico de prueba	p-valor
P01	96.566	0.196	7.001	0.000
P02	84.779	0.372	14.004	0.000
P03	38.834	0.339	12.572	0.000
P04	58.197	0.449	17.524	0.000
P05	73.732	0.372	14.012	0.000
P06	57.787	0.462	18.165	0.000
P07	45.172	0.333	12.354	0.000
P08	50.940	0.366	13.738	0.000
P09	33.716	0.421	16.193	0.000
P10	75.020	0.416	15.956	0.000
P11	50.947	0.376	14.128	0.000
P12	49.549	0.514	20.902	0.000
P13	48.182	0.355	13.196	0.000
P14	46.053	0.314	11.527	0.000
P15	62.850	0.404	15.355	0.000
P16	75.687	0.382	14.297	0.000
P17	70.183	0.470	18.448	0.000
P18	68.729	0.445	17.295	0.000
P19	38.747	0.355	13.195	0.000
P20	59.850	0.442	17.075	0.000
P21	43.295	0.308	11.258	0.000
P22	61.199	0.426	16.299	0.000
P23	29.250	0.428	16.392	0.000
P24	68.828	0.421	16.064	0.000
P25	12.231	-0.005	-0.187	0.574

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro 5.3, observamos que los ítems más fáciles de la prueba han sido los ítems 1, 2, 16, 10, 5 y 17 dado que sus probabilidades de aciertos son

superiores a 0.7 y los más difíciles son los ítems 23 y 25 dado que sus probabilidades de acierto son inferiores a 0.3.

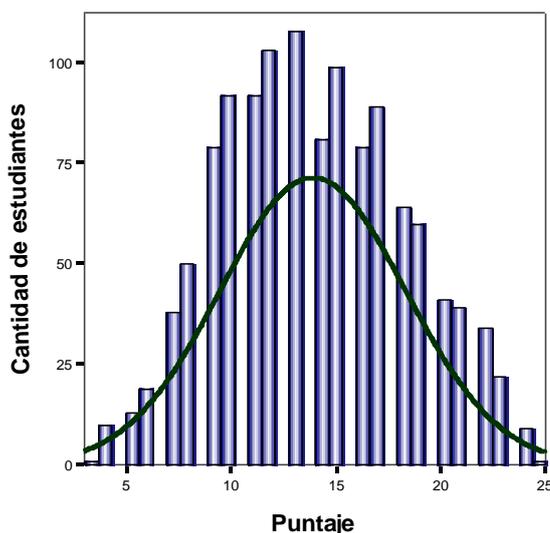
Adicionalmente por el índice de discriminación (correlación biseral puntual) y el estadístico de prueba concluimos que un nivel de significancia del 5% existe evidencia estadística suficiente para afirmar que la puntuación que obtienen los estudiantes en el ítem 25 no correlaciona con la puntuación total de la prueba, es decir que el ítem 25 de la prueba no discrimina significativamente. En la literatura psicométrica si la discriminación es menor a 0.25 se dice que el ítem no discrimina.

A nivel de la prueba:

5.1.2 Análisis de la puntuación total

Gráfico N° 5.1

Distribución de puntajes de la prueba



FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.4

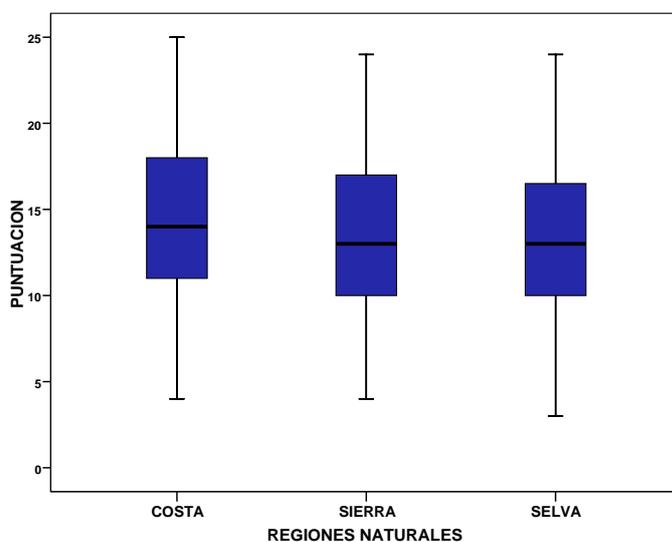
Estadísticos descriptivos del puntaje por regiones naturales

ESTADISTICOS	COSTA	SIERRA	SELVA	TOTAL
Mínimo	4	3	3	3
Media	14.38	13.37	13.43	13.89
Mediana	14	13	13	14
Moda	13	10	13	13
Máximo	25	24	24	25
Desviación	4.38	4.49	4.27	4.43
Asimetría	0.06	0.25	0.22	0.14
Curtosis	-0.55	-0.67	-0.46	-0.61
TOTAL	621	406	196	1223

FUENTE: Ministerio de Educación
ELABORACION: Propia

Gráfico N° 5.2

Gráfico de puntajes promedios por regiones naturales



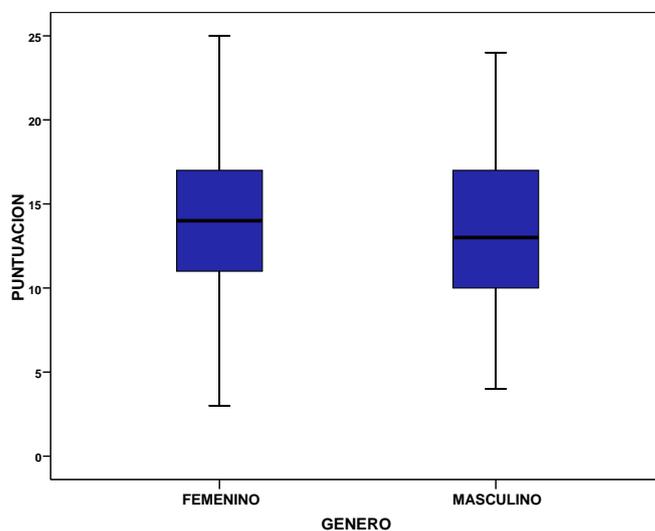
FUENTE: Ministerio de Educación
ELABORACION: Propia

Del gráfico N°5.1, del cuadro N°5.4 y gráfico N°5.2 observamos que la puntuación promedio a nivel nacional de la prueba es de 13.89 puntos. En la Costa se observa una leve mejoría en la puntuación promedio con respecto a la

nacional y las puntuaciones en las regiones Sierra y Selva muestran puntuaciones menores a la nacional. La mediana es de 14 puntos y en las regiones naturales son similares, y la moda es de 13 puntos tanto a nivel nacional como para la Costa y la Selva, pero para la Sierra la moda es de 10 puntos. Se intuye que la distribución de la puntuación obtenida por los estudiantes es casi simétrica en la Costa, Selva y Nacional debido a que la media, la moda, la mediana son casi similares; salvo en la Sierra se muestra una distribución algo asimétrica. Observamos que la distribución de los puntuaciones obtenidas por los estudiantes en la prueba de comprensión lectora es platicurtica dado que muestra un coeficiente de curtosis menor a 0 a nivel nacional y en todas las regiones naturales.

Gráfico N° 5.3

Gráfico de puntajes promedios por género



FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.5

Estadísticos descriptivos del puntaje por género

ESTADISTICOS	FEMENINO	MASCULINO	TOTAL
Mínimo	3	4	3
Media	14.13	13.54	13.89
Mediana	14	13	14
Moda	13	12	13
Máximo	25	24	25
Desviación	4.53	4.25	4.43
Asimetría	0.08	0.22	0.14
Curtosis	-0.66	-0.5	-0.61
TOTAL	729	494	1223

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del gráfico N°5.3 y el cuadro N°5.5 se muestra la distribución y los estadísticos descriptivos del puntaje obtenido por los estudiantes en la prueba de comprensión lectora por género. Se observa que los estudiantes evaluados de género femenino presentan una leve mejoría en la puntuación promedio con respecto a la puntuación promedio nacional. Las puntuaciones obtenidas por los estudiantes de género masculino muestran puntuaciones menores a la nacional. Adicionalmente, las puntuaciones obtenidas por los estudiantes de género femenino son mayores a los de género masculino.

5.1.3 Unidimensionalidad de la prueba

Con la unidimensionalidad se busca que la prueba de comprensión lectora explique solo un rasgo latente que es la habilidad que tiene el estudiante quinto de secundaria para resolver la prueba. Para ello se utiliza la técnica estadística multivariada análisis factorial, considerando como matriz de correlación del modelo a la matriz de correlación tetracórica (ver código en Anexo B) dado que

se correlacionan ítems dicotómicos. Para poder decir que la prueba es unidimensional el primer factor tienen que explicar la varianza entre 20% y 40% (Zwick, 1985). A continuación, se muestran los resultados del análisis factorial:

Cuadro N° 5.6

Varianza total explicada

COMPONENTE	AUTOVALORES INICIALES			SUMA DE LAS SATURACIONES AL CUADRADO DE LA EXTRACCION		
	TOTAL	% VARIANZA	% ACUMULADO	TOTAL	% VARIANZA	% ACUMULADO
1	6.37	25.47	25.47	6.37	25.47	25.47
2	1.46	5.82	31.29	1.46	5.82	31.29
3	1.28	5.13	36.42	1.28	5.13	36.42
4	1.24	4.96	41.37	1.24	4.96	41.37
5	1.14	4.55	45.92	1.14	4.55	45.92
6	1.09	4.36	50.29	1.09	4.36	50.29
7	1.04	4.14	54.43	1.04	4.14	54.43
8	0.98	3.93	58.36			
9	0.94	3.76	62.12			
10	0.90	3.60	65.72			
11	0.83	3.32	69.03			
12	0.80	3.22	72.25			
13	0.79	3.15	75.39			
14	0.77	3.08	78.48			
15	0.68	2.72	81.20			
16	0.67	2.66	83.86			
17	0.66	2.65	86.51			
18	0.62	2.48	88.99			
19	0.58	2.32	91.31			
20	0.51	2.05	93.36			
21	0.47	1.90	95.26			
22	0.44	1.74	97.00			
23	0.38	1.52	98.53			
24	0.30	1.19	99.71			
25	0.07	0.29	100.00			

Método de extracción: Análisis de componentes principales.

FUENTE: Ministerio de Educación
ELABORACION: Propia

En el cuadro N^o5.6 observamos la varianza total muestral explicada por los factores. Observamos que el primer factor 25.47% de la varianza y los demás factores explican la varianza muestral en menos del 6%. Como el primer factor explica entre el 20% y 40% de la varianza usando la matriz de correlación tetracórica (Zwick, 1985) concluimos que la prueba explica un solo rasgo latente es decir es unidimensional.

5.1.4 Confiabilidad de la prueba

De la prueba de comprensión lectora obtenemos una confiabilidad de 0.752 de lo que concluimos que el instrumento de medición es confiable.

De análisis de TCT se muestra que en la prueba de comprensión lectora que la prueba no ha sido tan fácil tan difícil dado que la distribución de la puntuación es casi simétrica, la prueba es confiable y unidimensional. Adicionalmente se observa que a nivel de ítems los ítems 1, 2, 16, 10, 5 y 17 fáciles, los ítems 3 y 25 son difíciles y el ítem 25 no discrimina significativamente.

De acuerdo a lo mostrado con el análisis de la TCT concluimos que se puede implementar el modelo TRI dicotómico de dos parámetros o modelo de Birnbaum.

5.2 Análisis de la prueba bajo el enfoque TRI

Se analiza la prueba de comprensión lectora bajo el enfoque TRI con la implementación del modelo TRI dicotómico de dos parámetros o modelo de Birnbaum.

5.2.1 Estimación de parámetros y la variable latente del modelo

Se estiman los parámetros del modelo: la discriminación y dificultad del ítem, y la variable latente: la habilidad de los estudiantes de quinto de secundaria de

las instituciones educativas estatales evaluadas con la prueba de comprensión lectora. Se estiman los parámetros del modelo y la variable latente mediante métodos bayesianos bajo el esquema: *adaptive rejection sampling* (ARS). Para la estimación de los parámetros y la variable latente se hace uso del software WinBugs (ver código en Anexo B).

5.2.1.1 Estimación de parámetros del modelo: dificultad y discriminación

Cuadro N° 5.7

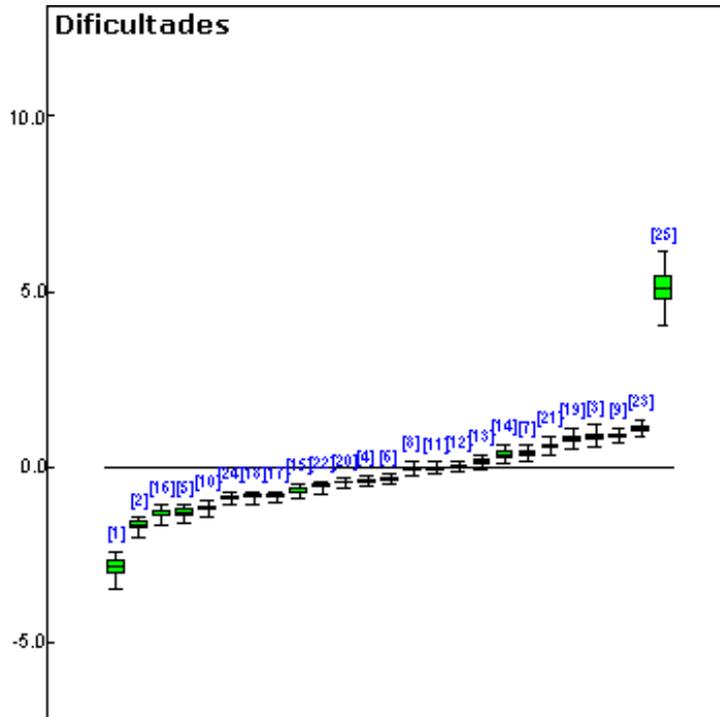
Dificultad de ítems bajo el modelo TRI

ITEM	DIFICULTAD	DESVIACION	Q1	MEDIANA	Q3	CAPACIDAD
1	-2.892	0.293	-3.076	-2.865	-2.684	Obtiene información explícita del texto
2	-1.684	0.135	-1.763	-1.67	-1.594	Obtiene información explícita del texto
3	0.832	0.148	0.724	0.821	0.919	Reflexiona acerca del texto
4	-0.413	0.074	-0.461	-0.41	-0.363	Obtiene información explícita del texto
5	-1.335	0.141	-1.419	-1.323	-1.234	Hace inferencias sobre información del texto
6	-0.378	0.072	-0.426	-0.377	-0.334	Obtiene información explícita del texto
7	0.367	0.122	0.281	0.365	0.452	Hace inferencias sobre información del texto
8	-0.067	0.1	-0.13	-0.066	0.001	Hace inferencias sobre información del texto
9	0.875	0.108	0.799	0.867	0.944	Hace inferencias sobre información del texto
10	-1.186	0.104	-1.257	-1.184	-1.118	Hace inferencias sobre información del texto
11	-0.066	0.093	-0.131	-0.07	0	Obtiene información explícita del texto
12	0.002	0.061	-0.041	-0.001	0.045	Hace inferencias sobre información del texto
13	0.122	0.097	0.054	0.123	0.183	Hace inferencias sobre información del texto
14	0.341	0.14	0.244	0.339	0.432	Hace inferencias sobre información del texto
15	-0.707	0.093	-0.766	-0.701	-0.645	Hace inferencias sobre información del texto
16	-1.355	0.131	-1.44	-1.34	-1.263	Obtiene información explícita del texto
17	-0.856	0.081	-0.908	-0.852	-0.8	Obtiene información explícita del texto
18	-0.876	0.086	-0.932	-0.871	-0.819	Obtiene información explícita del texto
19	0.778	0.131	0.687	0.772	0.856	Hace inferencias sobre información del texto
20	-0.478	0.074	-0.528	-0.477	-0.429	Reflexiona acerca del texto
21	0.561	0.14	0.463	0.553	0.644	Reflexiona acerca del texto
22	-0.566	0.085	-0.618	-0.565	-0.51	Hace inferencias sobre información del texto
23	1.063	0.112	0.983	1.057	1.129	Obtiene información explícita del texto
24	-0.922	0.099	-0.987	-0.912	-0.854	Hace inferencias sobre información del texto
25	5.07	0.536	4.717	5.05	5.416	Reflexiona acerca del texto

FUENTE: Ministerio de Educación
ELABORACION: Propia

Gráfico N° 5.4

Diagrama de cajas de la dificultad de ítems bajo el modelo TRI



FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.8

Prueba de normalidad para la dificultad de ítems bajo el modelo TRI

		Dificultad
Parámetros normales	Media	-0.15
	Desviación típica	1.43
Diferencias más extremas	Absoluta	0.16
	Positiva	0.16
	Negativa	-0.12
Z de Kolmogorov-Smirnov		0.79
Sig. asintót. (bilateral)		0.56

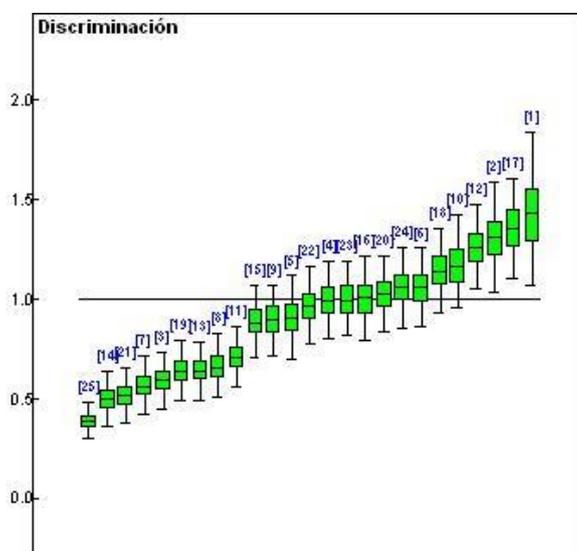
FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.8, cuadro N°5.9 y el gráfico N°5.4 se muestra las estimaciones bayesianas de las dificultades de los ítems implementando el modelo TRI

dicotómico de dos parámetros de la prueba de comprensión lectora que rindieron los estudiantes de quinto de secundaria de instituciones educativas estatales a nivel nacional. Del cual se muestra que los ítems 1, 2, 16, 5 y 10 son los ítems más fáciles de la prueba y los ítems 23 y 25 los más difíciles. Es decir que para los estudiantes evaluados no es muy difícil obtener información implícita en los textos que lee y pueden realizar ciertas inferencias de lo que lee, pero le son muy difíciles en textos relativamente grandes obtener información implícita y reflexionar acerca de lo leído en el texto. Además se concluye que la dificultad de ítems mediante la implementación del modelo es una variable aleatoria y que a un nivel de significancia del 5% existe evidencia estadística suficiente para afirmar que sigue una distribución normal.

Gráfico N° 5.5

Diagrama de cajas de la discriminación de ítems bajo el modelo TRI



FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.9

Discriminación de ítems bajo el modelo TRI

ITEM	DISCRIMINACION	DESVIACION	Q1	MEDIANA	Q3
1	1.43	0.20	1.29	1.42	1.56
2	1.31	0.13	1.22	1.30	1.39
3	0.59	0.07	0.54	0.60	0.65
4	0.99	0.10	0.93	0.99	1.06
5	0.91	0.10	0.84	0.91	0.97
6	1.06	0.10	0.99	1.05	1.13
7	0.57	0.08	0.52	0.56	0.62
8	0.66	0.08	0.61	0.65	0.71
9	0.90	0.09	0.84	0.90	0.96
10	1.17	0.12	1.08	1.16	1.25
11	0.71	0.08	0.66	0.71	0.76
12	1.26	0.11	1.19	1.25	1.33
13	0.64	0.08	0.59	0.64	0.69
14	0.50	0.07	0.45	0.50	0.55
15	0.89	0.09	0.83	0.88	0.95
16	1.01	0.11	0.93	1.00	1.07
17	1.36	0.13	1.26	1.36	1.45
18	1.14	0.11	1.07	1.14	1.21
19	0.64	0.08	0.59	0.64	0.69
20	1.02	0.10	0.96	1.02	1.09
21	0.52	0.07	0.47	0.52	0.57
22	0.97	0.10	0.90	0.97	1.03
23	0.99	0.10	0.93	0.99	1.06
24	1.06	0.10	0.99	1.06	1.12
25	0.39	0.04	0.36	0.38	0.41

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.9 y el grafico N°5.5 se muestra las estimaciones del índice de discriminación de los ítems implementando el modelo TRI dicotómico de dos parámetros o modelo de Birnbaum de la prueba de comprensión lectora que rindieron los estudiantes de quinto de secundaria de instituciones educativas estatales a nivel nacional. Del cual se muestra que los ítems 1, 17, 2 y 12 tiene mayores índices de discriminación, los cuales me ayudan a diferenciar mejor a

los estudiantes evaluados en función a sus capacidades en comprensión lectora dado que la probabilidad de responder correctamente al ítem aumenta. Los ítems 25 y 14 muestran bajos índices de discriminación, los cuales no ayudan mucho a diferencias a los estudiantes dado que las probabilidades de responder correctamente a los ítems disminuyen y va depender solo del nivel de habilidad y dificultad del ítem.

Cuadro N° 5.10

Prueba de normalidad para la discriminación de ítems bajo el modelo TRI

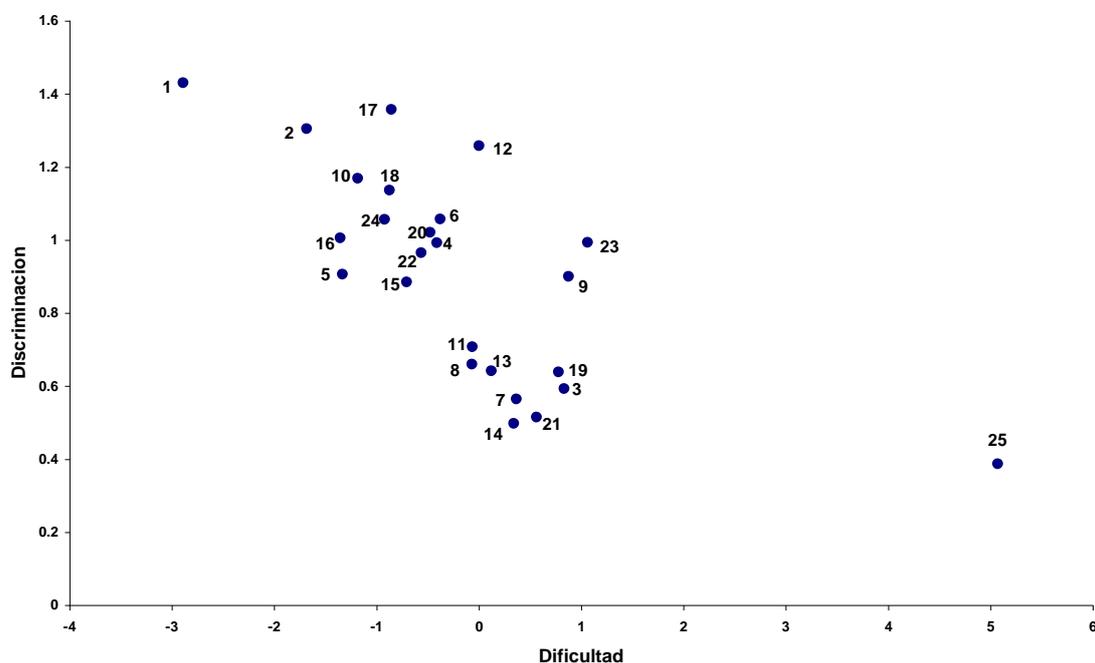
		Discriminación
Parámetros normales	Media	0.91
	Desviación típica	0.29
	Absoluta	0.12
Diferencias más extremas	Positiva	0.12
	Negativa	-0.11
Z de Kolmogorov-Smirnov		0.61
Sig. asintót. (bilateral)		0.85

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.10 se concluye que el índice de discriminación de los ítems mediante la implementación del modelo es una variable aleatoria y que a un nivel de significancia del 5% existe evidencia estadística suficiente para afirmar que sigue una distribución normal a pesar de que se ha establecido como distribución a priori una distribución log normal. Adicionalmente los índices de discriminación se pueden expresar en intervalos de confianza.

Gráfico N° 5.6

Diagrama de dispersión del índice discriminación de vs la dificultad de ítems bajo el modelo TRI



FUENTE: Ministerio de Educación
ELABORACION: Propia

En el gráfico N°5.6 se muestra el diagrama de dispersión de la discriminación vs la dificultad, observamos la relación entre los parámetros del modelo TRI dicotómico de dos parámetros o modelo de Birnbaum. Los ítems 1, 17, 2, 10 y 18 presentan discriminaciones altas y bajas dificultades los cuales van a ser útiles para diferenciar a estudiantes de quinto de secundaria con altas capacidades en comprensión lectora. El ítem 25 tiene el índice de discriminación bajo y la dificultad alta, este ítem no ayuda a discriminar a los estudiantes evaluados.

5.2.1.2 Estimación de la variable latente del modelo: la habilidad

Los métodos de estimación clásica no estiman las habilidades cuando el estudiante no ha respondido la totalidad de las preguntas, para esos caso se

aplican técnicas de imputación las cuales generan sesgo. En el método de máxima verosimilitud conjunta cuando el estudiante responde correctamente o incorrectamente se sobrestiman o subestiman las habilidades. Para la estimación de las habilidades en los métodos de máxima verosimilitud marginal y condicional hay que aplicar métodos bayesianos. Adicionalmente los métodos bayesianos son más precisos que los métodos clásicos. (Bazan J y Chincaro O, 2010).

Por estas razones las estimaciones de las habilidades de los estudiantes se realizo mediante métodos bayesianos.

Cuadro N° 5.11

Estadísticos de las habilidades estimadas bajo el modelo TRI

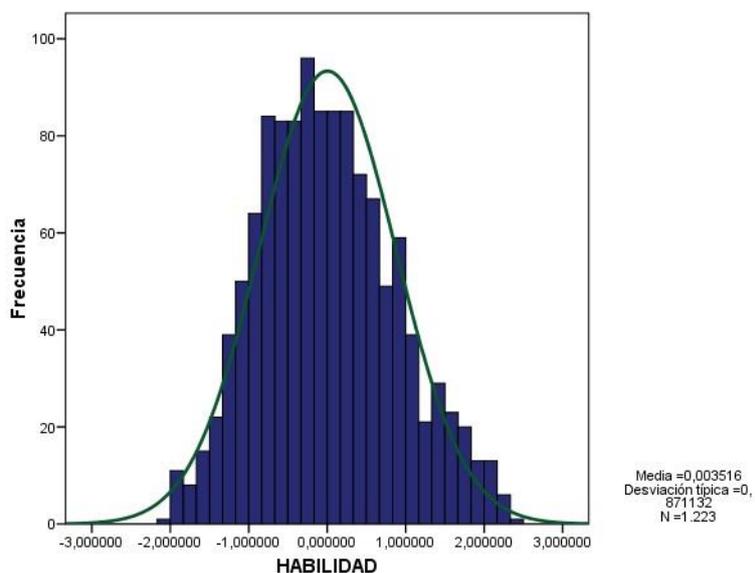
<u>ESTADISTICO</u>	<u>HABILIDAD</u>
Mínimo	-2.084
Media	0.004
Máximo	2.424
Desviación	0.871
Asimetría	0.286
Curtosis	-0.337

FUENTE: Ministerio de Educación
ELABORACION: Propia

En el cuadro N°5.11 se muestra los estadísticos descriptivos de la habilidad estimada con la implementación del modelo TRI dicotómico de dos parámetros o modelo de Birnbaum. Se observa que la habilidad es una variable latente aleatoria el cual siguen una distribución de probabilidad leptocurtica (el coeficiente de curtosis es menor a 0) y asimétrica positiva (el coeficiente de asimetría mayor a 0) con media de 0.0035 logits y desviación estándar de 0.8711 logits.

Gráfico N° 5.7

Distribución de la habilidad de los estudiantes bajo el modelo TRI



FUENTE: Ministerio de Educación
ELABORACION: Propia

Cuadro N° 5.12

Prueba de normalidad para la habilidad bajo el modelo TRI

		Habilidad
Parámetros normales	Media	0.004
	Desviación típica	0.871
	Absoluta	0.034
Diferencias más extremas	Positiva	0.034
	Negativa	-0.024
Z de Kolmogorov-Smirnov		1.204
Sig. asintót. (bilateral)		0.110

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del grafico N°5.7 y el cuadro N°5.12 se concluye que a un nivel de significancia del 5% existe evidencia estadística suficiente para afirmar que la habilidades de

los estudiantes de quinto de secundaria que rindieron la prueba de comprensión lectora siguen una distribución normal.

Como es un muestreo complejo en varias etapas, para poder realizar comparaciones entre estratos se calcula el error de muestreo vía técnica de remuestreo con la metodología jackknife. En el anexo A se muestra la metodología de remuestreo y el anexo B se muestra el código en R utilizado. En el cuadro N°5.13 se muestra los errores vía jackknife a nivel nacional y por estratos de inferencia:

Cuadro N° 5.13
Errores jackknife nacional y por estratos de inferencia

ESTRATO	ERRORES
COSTA	0.035
SIERRA	0.044
SELVA	0.059
FEMENINO	0.033
MASCULINO	0.037
NACIONAL	0.025

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.13 se muestra que las inferencias son validas dado que los errores son aceptables para cada estrato de inferencia.

5.2.2 Comparaciones de la habilidad promedio por regiones naturales

Debido a las diferencias existentes en las puntuaciones obtenidas por los estudiantes de quinto de secundaria de las instituciones educativas estatales en la prueba de comprensión lectora se busca mostrar diferencias en las habilidades promedios en las regiones naturales: costa, sierra y selva.

Cuadro N° 5.14

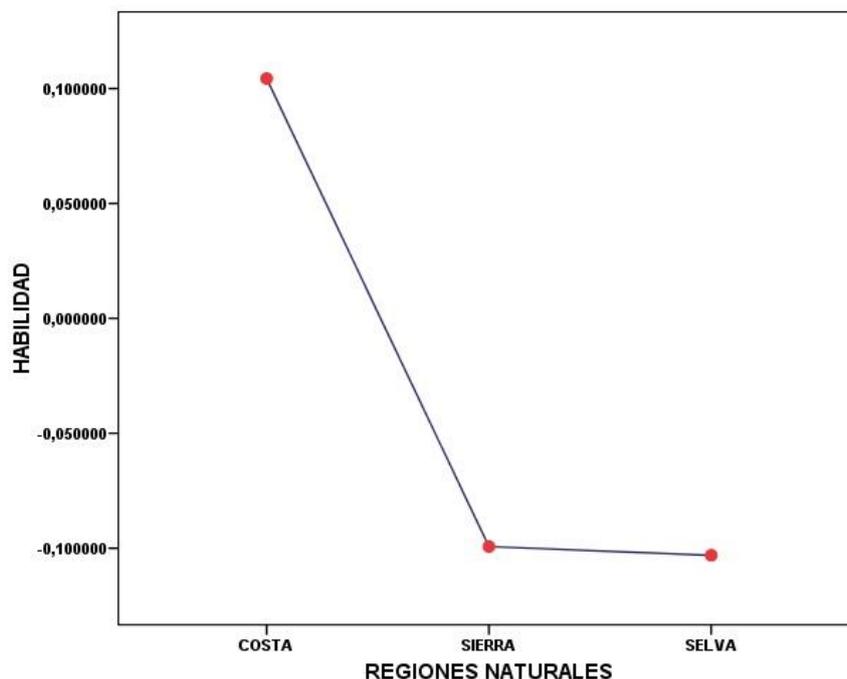
Estadísticos descriptivos de la habilidad de los estudiantes por regiones naturales

Región	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
COSTA	-2.084	0.104	2.424	0.865	0.203	-0.288
SIERRA	-1.946	-0.099	2.252	0.886	0.395	-0.395
SELVA	-1.925	-0.103	2.144	0.826	0.362	-0.093
TOTAL	-2.084	0.004	2.424	0.871	0.286	-0.337

FUENTE: Ministerio de Educación
ELABORACION: Propia

Gráfico N° 5.8

Gráfico de la habilidad de los estudiantes por regiones naturales



FUENTE: Ministerio de Educación
ELABORACION: Propia

En el cuadro N°5.14 y el gráfico N°5.8 se muestra los estadísticos descriptivos de la habilidad de los estudiantes de quinto de secundaria de los colegios estatales que rindieron la prueba de comprensión lectora según regiones naturales, en la cual se observa que los estudiantes de la Costa tienen habilidades promedios superior a la habilidad promedio nacional y a la vez estas son mayores que los de la Sierra y la Selva. Las habilidades de los estudiantes de la Sierra son ligeramente mayores a los de la Selva.

Debido a las diferencias mostradas en las regiones naturales, es de interés determinar si dichas diferencias son significativas para ello se realiza un análisis de varianza:

Análisis de varianza de las habilidades de los estudiantes

Cuadro N° 5.15

Prueba de homogeneidad de varianzas

Estadístico de Levene	gl1	gl2	Sig.
1.626	2	1220	0.197

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.15, mediante el estadístico de Levene se realiza la prueba de homogeneidad de varianzas, como el p_value es mayor que 0.05 concluimos que a un nivel de significancia del 5% existe evidencia estadística suficiente para afirmar que no existen diferencias significativas entre las varianzas de las habilidades de los estudiantes de quinto de secundaria que rindieron la prueba de comprensión lectora en cada una de las regiones naturales.

Cuadro N° 5.16

Tabla ANOVA

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	12.834	2	6.417	8.561	0.000
Intra-grupos	914.506	1220	0.750		
TOTAL	927.340	1222			

FUENTE: Ministerio de Educación
ELABORACION: Propia

En el cuadro N°5.16 se muestra que el p_value es menor que 0.05 de lo cual se concluye que a un nivel de significancia del 5% existe evidencia estadística suficiente para determinar que hay diferencias significativas en las habilidades promedio de los estudiantes evaluados en cada una de las regiones naturales.

Debido a presencia de diferencias significativas en las habilidades promedios de los estudiantes evaluados en cada una de las regiones naturales, es de interés ver entre que regiones ocurren las diferencias.

Cuadro N° 5.17

Prueba HSD de Tukey para comparaciones múltiples

(I) Región Natural	(J) Región Natural	Diferencias de medias (I-J)	Error Típico	Sig.
COSTA	SIERRA	0.204	0.055	0.001
	SELVA	0.207	0.071	0.010
SIERRA	COSTA	-0.204	0.055	0.001
	SELVA	0.004	0.075	0.999
SELVA	COSTA	-0.207	0.071	0.010
	SIERRA	-0.004	0.075	0.999

FUENTE: Ministerio de Educación
ELABORACION: Propia

Del cuadro N°5.17 de comparaciones múltiples se muestra que a un nivel de significancia del 5% existe diferencia significativa entre las habilidades promedios de los estudiantes de la Costa y las otras regiones, teniendo mayores habilidades promedios los estudiantes de la Costa. Por otro lado, los estudiantes de la Sierra tienen habilidades ligeramente mayores a los de la selva pero dicha diferencia no es significativa.

Se forman los siguientes grupos homogéneos: el primer grupo está formado por los estudiantes de quinto de secundaria que estudian en la Sierra y la Selva, los cuales muestran habilidades promedias menores a la habilidad promedio nacional. Y el segundo grupo que está formado por los estudiantes de quinto de secundaria que estudian en la Costa, los cuales muestran habilidades promedias mayores a la habilidad promedio nacional.

5.2.3 Comparaciones de la habilidad promedio de por genero

Debido a las diferencias existentes en las puntuaciones obtenidas por los estudiantes de quinto de secundaria de las instituciones educativas estatales en la prueba de comprensión lectora se busca mostrar diferencias en las habilidades promedios por género: femenino y masculino.

Cuadro N° 5.18

Estadísticos descriptivos de la habilidad de los estudiantes según género

Genero	N	Media	Desviación
FEMENINO	729	0.048	0.894
MASCULINO	494	-0.062	0.832

FUENTE: Ministerio de Educación
ELABORACION: Propia

En el cuadro N°5.18 se muestra los estadísticos descriptivos de la habilidad de los estudiantes de quinto de secundaria de los colegios estatales que rindieron la prueba de comprensión lectora según genero, en la cual se observa que las mujeres en comprensión lectora tienen habilidades promedios superiores los hombres.

Debido a las diferencias mostradas según género, es de interés determinar si dichas diferencias son significativas para ello se realiza la prueba T:

Cuadro N° 5.19

Prueba T y de Levene para muestras independientes

	Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias				
	F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia
Se han asumido varianzas iguales	3.572	0.059	2.178	1221	0.030	0.110	0.051
No se han asumido varianzas iguales			2.208	1107.265	0.027	0.110	0.050

FUENTE: Ministerio de Educación
ELABORACION: Propia

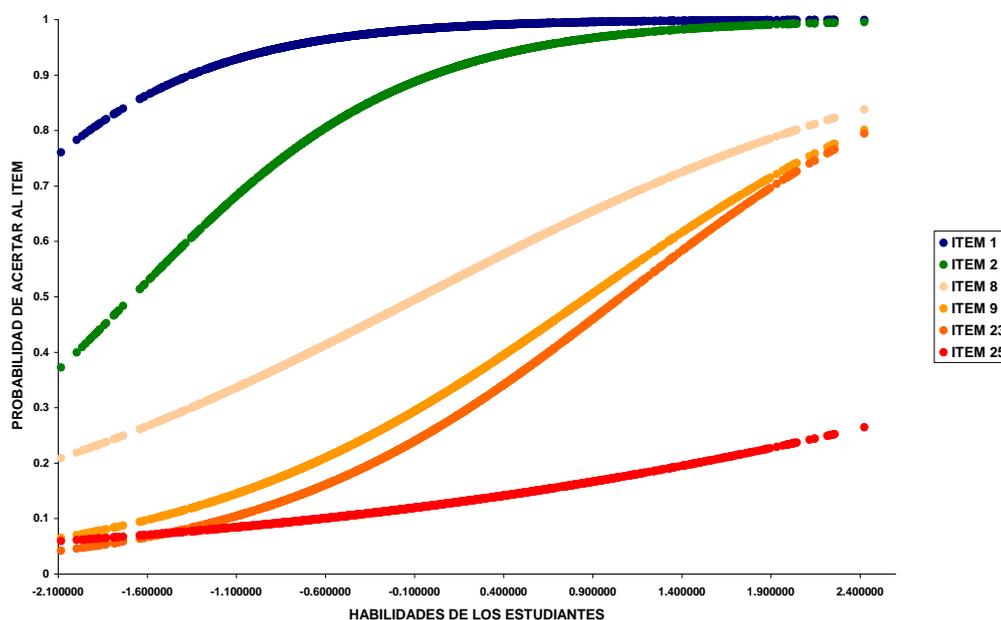
En el cuadro N°5.19 el p_value del estadístico de Levene es mayor a 0.05 concluimos que a un nivel de significancia del 5% no hay diferencias significativas las varianzas de las habilidades de los estudiantes según género. Adicionalmente por la prueba de T asumiendo varianzas iguales, se concluye que a un nivel de significancia del 5% existe evidencia estadística suficiente para determinar que hay diferencias significativas entre las habilidades promedios de los estudiantes evaluados según género.

5.2.4 Relaciones entre los parámetros del modelo y la variable latente

Una forma de mostrar la relación existente de los parámetros del modelo (dificultad y discriminación del ítem) y la variable latente (habilidad del estudiante) es a través de la curva característica del ítem y del mapa de ítems.

Gráfico N° 5.9

Curva características de ítems

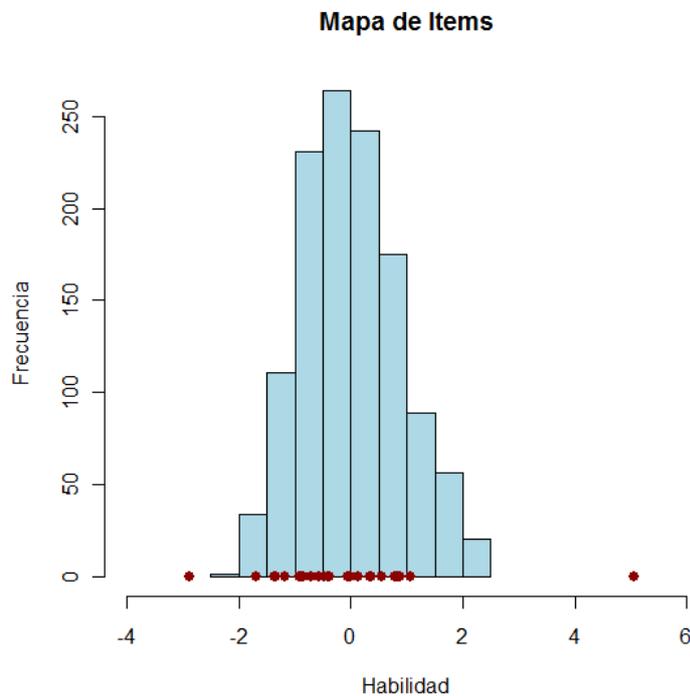


FUENTE: Ministerio de Educación
ELABORACION: Propia

Del Gráfico N°5.9 se muestra el comportamiento de los ítems desde el más fácil hasta el más difícil. Para los estudiantes de quinto de secundaria los ítems 1 y 2 han sido los ítems más fáciles y son los que presenta mayor índice de discriminación. El ítem 25 ha sido el más difícil y el que tiene menor índice de discriminación.

Gráfico N° 5.10

Mapa de ítems a nivel nacional



FUENTE: Ministerio de Educación
ELABORACION: Propia

Del gráfico N°5.10, se muestra el mapa de ítems en el cual se busca graficar las habilidades que tienen los estudiantes (las barras en celeste) y la dificultad que tienen cada ítem (los puntos en el gráfico) con la finalidad de evaluar las capacidades que tienen los estudiantes de quinto de secundaria en comprensión lectora. Las capacidades que tienen los estudiantes se analizan mediante niveles de desempeño o logro que alcanzan los estudiantes, existen diferentes metodologías para determinar dichos niveles, con la investigación se determina los niveles mediante el mapa de ítems, teniendo en cuenta la habilidad que tienen los estudiantes y la dificultad de la prueba determinándose tres niveles de desempeño: nivel previo, nivel básico y nivel suficiente

Nivel previo: En este nivel los ítems presenta dificultades menores a 0 logits. Los estudiantes de quinto de secundaria tienen las habilidades suficientes en comprensión lectora para resolver los ítems fáciles de la prueba; es decir tienen las capacidades de obtener información rápidamente que se encuentra en partes evidentes del texto, interpreta el significado de palabras o símbolos a partir del contexto de la lectura y pueden realizar inferencias mínimas según el texto leído.

Nivel básico: En este nivel los ítems presentan dificultades de 0 y menores a 1 logit. Los estudiantes resuelven los ítems relativamente difíciles de la prueba; es decir tienen las capacidades de obtener información explícita que se encuentra en partes no tan evidentes del texto, interpretan el mensaje del texto en textos no muy complejos y realizan inferencias más complejas.

Nivel suficiente: En este nivel las dificultades de los ítems son mayores o iguales a 1 logit. Los estudiantes tratan de resolver los ítems más difíciles de la prueba; es decir tienen las capacidades para obtener información explícita que se encuentra en partes incrustadas del texto y cuya recuperación requiere varios pasos, deducen el propósito con el cual el autor escribió el texto, hacen inferencias más complejas en las que se deben interpretar metáforas, relacionan varios textos a la vez y tienen las capacidades para reflexionar acerca del texto.

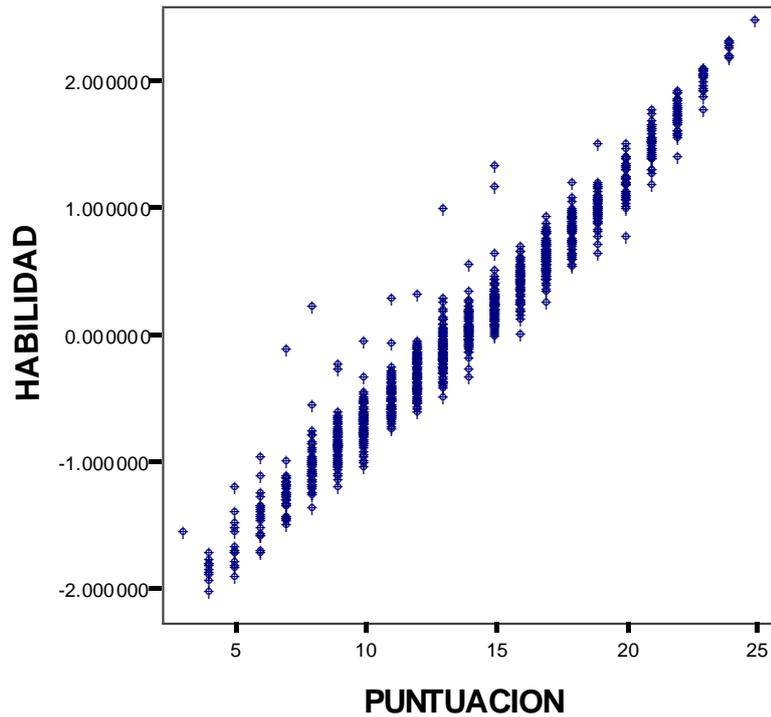
5.3 Relaciones entre la TCT y la TRI

Con la TCT se obtienen una primera mirada sobre el comportamiento de los estudiantes al momento de rendir una prueba de rendimiento, pero con el modelo TRI planteado en la investigación se pueden analizar variable latentes

presentes cuando un estudiante enfrenta la prueba y vez establecer relaciones con el modelo de TCT:

Gráfico N° 5.11

Diagrama de dispersión puntuación vs habilidad

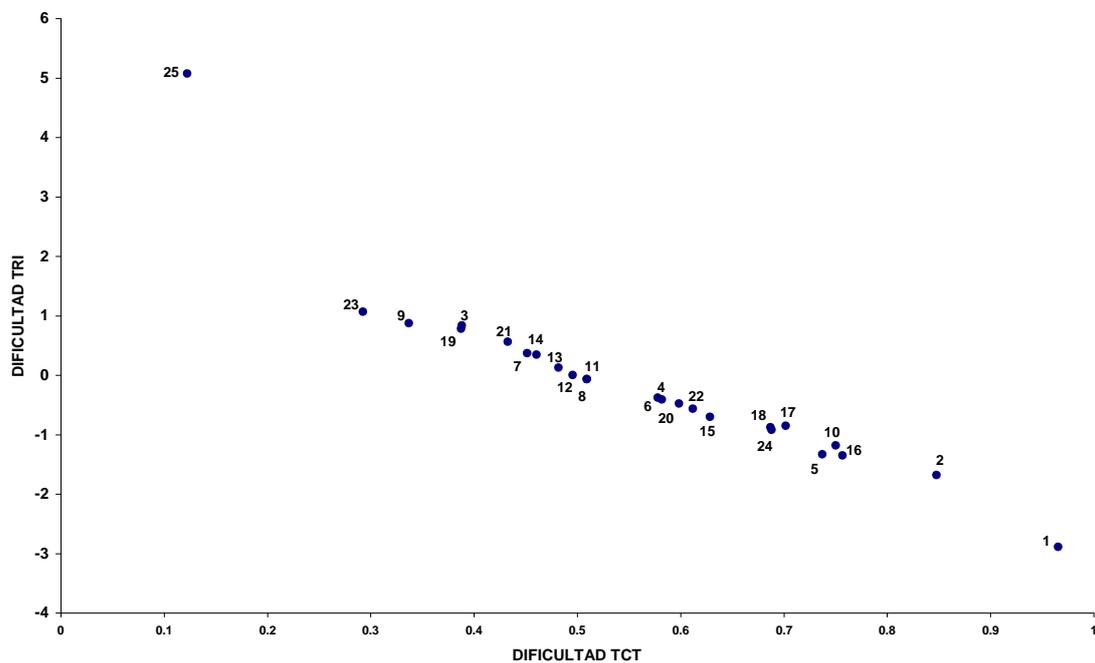


FUENTE: Ministerio de Educación
ELABORACION: Propia

Del grafico N°5.11 se observa que existe una relación directa entra la habilidad que tienen los estudiantes de quinto de secundaria y la puntuación que obtienen en la prueba de comprensión lectora, es decir cuánto más hábil sea un estudiante es más probable que tenga más puntaje en la prueba adicionalmente a esto no hay que dejar de lado el grado de dificultad que tiene cada pregunta de la prueba.

Gráfico N° 5.12

Diagrama de dispersión de la dificultad TCT vs dificultad TRI

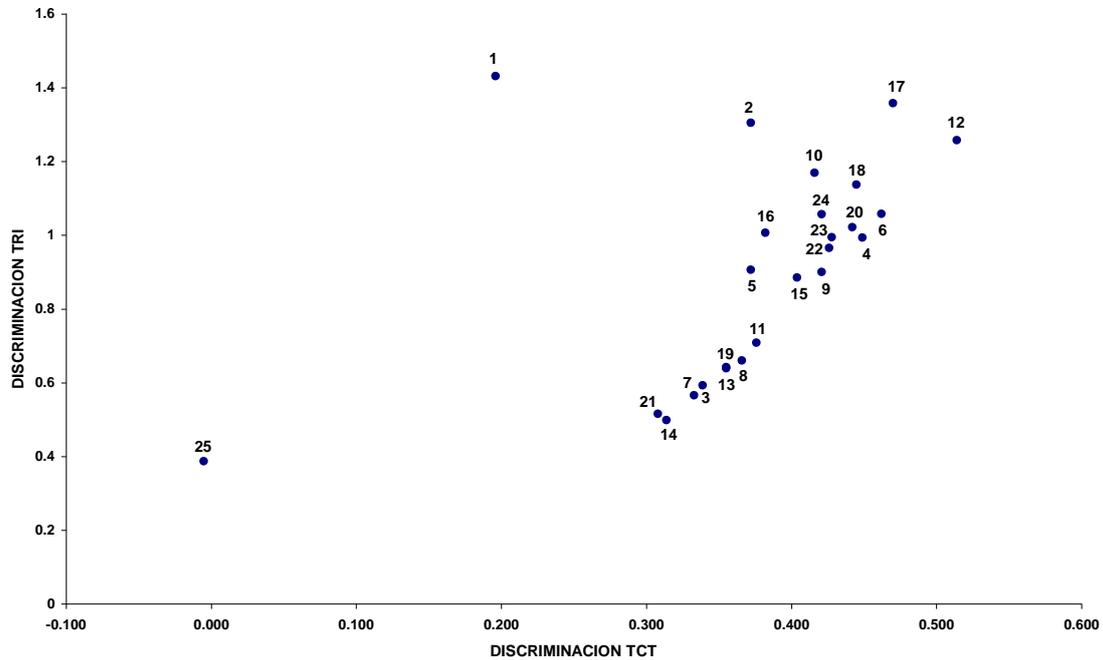


FUENTE: Ministerio de Educación
ELABORACION: Propia

Del gráfico N°5.12 se observa que existe una relación inversa entre la dificultad que tienen los ítems de la prueba de comprensión lectora bajo el enfoque TRI y la TCT, es decir a menor dificultad TCT mayor dificultad TRI o viceversa. Adicionalmente la TRI produce la misma jerarquización de los ítems que la TCT.

Gráfico N° 5.13

Diagrama de dispersión del índice de discriminación TCT vs índice de discriminación TRI



FUENTE: Ministerio de Educación
ELABORACION: Propia

En el gráfico N°5.13 se muestra el diagrama de dispersión entre la índice discriminación que produce el modelo TRI y la del modelo TCT las cuales producen diferentes índices de discriminación. El modelo TCT muestra que ítem 25 no discrimina, pero el TRI muestra un índice discriminación mínimo. Adicionalmente el índice de discriminación TRI es una variable aleatoria que sigue una distribución de probabilidad.

Conclusiones de la investigación

- Con la implementación del modelo y la estimación de la variable latente (habilidad) y los parámetros (dificultad y discriminación) se determinan 3 niveles de desempeño: **previo, básico y suficiente**. A nivel nacional aproximadamente el 52.4% de los estudiantes de quinto de secundaria pertenecen al **Nivel previo**: presentan logros de aprendizaje en comprensión lectora para el grado anterior y mínimos para el grado actual, es decir que en lugar de estar quinto deberían estar en cuarto de secundaria. El 34.1% pertenecen al **Nivel básico**; es decir, presentan logros suficientes para el grado evaluado. En este nivel, los estudiantes tienen las capacidades de obtener información explícita que se encuentra en partes no tan evidentes de los textos, elaboran inferencias no muy complejas. Finalmente, el 13.5% pertenecen al **Nivel suficiente**: presentan logros de aprendizaje suficientes para egresar del grado, dado que son capaces de obtener información explícita que se encuentra en partes incrustadas del texto y cuya recuperación requiere varios pasos, deducen el propósito con el cual el autor escribió el texto, hacen inferencias complejas en las que se deben interpretar metáforas,

relacionan varios textos a la vez y tienen las capacidades para reflexionar acerca del texto.

- Para la estimación de las variables latentes y parámetros del modelo se utilizaron métodos bayesianos dado que estos métodos no presentan restricciones a puntuaciones extremas y son más precisos que los métodos clásicos. Adicionalmente, los métodos bayesianos no presentan problemas en caso de datos incompletos. Para la implementación de estos métodos no se requieren de otras variables.
- Con la TCT se muestra que las puntuaciones que obtienen los estudiantes siguen una distribución normal, la prueba unidimensional y confiable. A la vez se pueden conocer los tipos de ítems que tiene la prueba (ítems fáciles y difíciles).
- Para analizar una prueba de rendimiento es mucho mejor analizarlo bajo un modelo TRI que con la TCT, dado que los modelos TRI brindan mucha más información que los clásicos y se pueden realizar inferencias más precisas. Con la TCT no se determinan diferencias significativas entre las puntuaciones promedio de los estudiantes tanto por regiones naturales ni por género, pero sin embargo con la implementación del modelo se determinan existen diferencias significativas entre las habilidades promedios de los estudiantes de la Costa y Sierra, teniendo mayores habilidades en Comprensión Lectora los alumnos de la Costa; y diferencias significativas entre las habilidades promedios según género siendo mayores la de las estudiantes mujeres.
- Existen diferentes softwares psicométricos para analizar pruebas de rendimiento, pero con la presente investigación se da importancia al uso

de softwares estadísticos como SPSS, WinBugs y R dando énfasis a la programación computacional. Mediante sintaxis (ver anexo A) del SPSS se determinaron la metodología TCT y la correlación tetracórica, con el WinBugs se implementó el modelo mediante la interfaz de programación (ver anexo A), finalmente con el software R se calcularon los errores muestrales mediante el módulo “bootstrap” (ver anexo A).

Recomendaciones para investigaciones futuras

- Para investigaciones futuras, en el modelo propuesto hay que tener en cuenta otros factores como las respuestas al azar del estudiante, aumentar el número de cuadernillos en la prueba y con ítems de anclaje entre los cuadernillos. Y realizar estudios de simulación para ver cuanto afecta los cambios en tamaños de muestra e ítems en las estimaciones de las variables latentes y parámetros del modelo. Adicionalmente aplicar mecanismos de validación del modelo propuesto.
- Existen peligros en utilizar softwares que aplican métodos de estimación diferente al bayesiano dado que pueden presentar sesgos e inconvenientes a la estimación. No obstante, existe mayor demanda computacional en la estimación bayesiana para muestras grandes.
- Se pueden realizar investigaciones complementarias para analizar el efecto de las prioris de la variable latente y parámetros del modelo.
- En una prueba de rendimiento en caso de obtener puntuaciones asimétricas utilizar modelos de Teoría de Respuesta al Ítem asimétricos con funciones de enlace logístico o la normal.

- Hay que aplicar políticas de mejoras en la comprensión lectora de los estudiantes y desarrollar pruebas que exijan más las capacidades de los estudiantes en esta área pedagógica.
- Con la finalidad monitorear los conocimientos adquiridos por lo estudiantes de quinto de secundaria se deben tomar 3 evaluaciones: uno al empezar el grado, otro a la mitad y la otra al final del grado. Realizar comparaciones con las tres evaluaciones y analizar el comportamiento de los niveles de desempeño.

Bibliografía

- Albert, J.H (1992). "Bayesian Estimation of Normal Ogive Item Response Curves Using Gibbs Sampling". Journal of Educational Statistics 17(3): 251 - 269.
- Albert, J (2009). Bayesian Computation with R , New York: Springer
- Anderson, J. Gerbin, D y Hunter, J. (1987). "On the assessment of unidimensional measurement: internal and external consistency, and overall consistency criteria". Journal of Marketing Research 24(4): 432 – 437.
- Attorresi, H. (1999). "Aplicaciones del modelo logístico de tres parámetros en una prueba de completar frases". Investigaciones en Psicología 4(1): 7-25.
- Baker, F.B. (1992). Ítem Response Theory. Parameter estimation techniques. New York: Marcel Dekker.
- Barry, R.J (2005). Probabilidad: Un curso a nivel intermedio. IMCA
- Bayes, C. y Bazán, J. (2010). "Inferencia Bayesiana en Modelos de respuesta Binaria usando BRMUV". I Jornada de Probabilidad y Estadística PUCP, Serie B(25).
- Bazán, J. Calderon, A. y Valdivieso, L. (2010). "Enfoque Bayesiano en Modelos de Teoría de Respuesta al Ítem". I Jornada de Probabilidad y Estadística PUCP, Serie B(27).
- Congdon, P. (2003). Applied Bayesian Modelling. John Wiley & Sons.

- Chincaro, O y Bazan, J. (2010). "Póster. Una comparación de los métodos de estimación del modelo de Rasch". I Jornada de Probabilidad y Estadística PUCP.
- Cuesta, M. (1996). Unidimensionalidad. Pirámide
- De Boeck, P y Wilson, M. (2009). Explonatory Ítem Response Models. New York: Springer.
- Fischer, G. H. (2007). "Rasch models", Handbook of Statistics in Psychometrics, 26, 515 - 586.
- Flores M., E.F. (2010). "A propósito de la Movilización Nacional por la Comprensión Lectora Placer y Comprensión: Dos caras de una misma moneda".
<http://destp.minedu.gob.pe/secundaria/nwdes/pdfs/placerycomprension.pdf>
- García, L. "Factores Asociados el rendimiento académico de estudiantes de psicología de la UNMSM". Revista de Investigación en Psicología 5(1)
- González B. C. (2003). "Factores determinantes del bajo rendimiento académico en educación secundaria". Tesis doctoral. Universidad Complutense de Madrid.
- Hambleton, R.K, Swaminathan, H. y Jane, R.H. (1991). "MMSS Fundamentals of Item Response". Sage Publications.Inc.
- Johnson V. y Albert J. (2000). Ordinal Data Modelling. New York: Springer.
- Kim, S. (2001). "An Evaluation of a Markov Chain Monte Carlo Method for the Rasch Model", Applied Psychological Measurement 25(2): 163 - 176.
- Kohan, N. (1999). Teoría de respuesta al ítem y su aplicación al test. Laboratorio de Evaluación Psicológica y Educativa.

- Lazarfeld, P. (1950). The logical and mathematical foundations of latent structure analysis. Princeton University Press.
- Lord, F.M. (1986). "Maximum likelihood and bayesian parameter estimation in the item response theory". Journal of Educational Measurement 23(2) 157-162.
- Lord, F. y Novick, M. (1968). Statistical Theory of Mental Test Scores. New York: Addison Wesley.
- Lord, F. (1980). Applications of Item Response theory to practical testing problems. Hillsdale: Erlbaum Associates.
- Muñiz, J. (1998). Teoría Clásica de los Test. Pirámide.
- Nunnally, J. y Bernstein, I. (1995). Teoría Psicométrica. Mexico: Mc Graw Hill.
- Swaminathan, H. y Gifford, J. (1982). "Bayesian Estimation in the Rasch Model", Journal of Educational Statistics 7(3): 175 - 191.
- Tamner, M.A. (1995). Tools for Statistical Inference. Springer.
- Vander Liden, W. y Hambleton, R. (1997). Handbook of Modern Item response Theory. New York: Springer.
- Wright, B. (1994). "Unidimensionality coefficient". Rasch Measurement Transactions 8(33), 385.
- Zwick, R (1985). "Assessment of the dimensionality of NEAP your 15 reading data (ETS Res. Rep No. 86 – 4)". Princeton: Educational Testing service

Anexos

Anexo A

A.1 Suficiencia estadística

Una estadística $T = g(X_1, X_2, X_3, \dots, X_n)$ es suficiente para el parámetro θ si $f_{X_1, X_2, X_3, \dots, X_n | T=t}(\cdot)$ no depende de θ para cada t valor posible de T .

A.1.1 Teorema de Factorización de Neyman

Tomando la literatura de (Hogg R.V. and Craig A.T., 1978) y (Tammer M.A. 1995):

Sea $(x_1, x_2, x_3, \dots, x_n)$ una muestra aleatoria de tamaño n que tiene una distribución $f(x, \theta)$, $\theta \in \Omega$. Sea $Y_1 = u_1(x_1, x_2, x_3, \dots, x_n)$ una estadística cuya función de probabilidad es $g_1(Y_1, \theta)$. Entonces Y_1 es una estadística suficiente para todo θ si y solamente si:

$$\frac{f(x_1, \theta)f(x_2, \theta)f(x_3, \theta) \dots f(x_n, \theta)}{g(u_1(x_1, x_2, x_3, \dots, x_n, \theta))} = H(x_1, x_2, x_3, \dots, x_n)$$

Donde $H(x_1, x_2, x_3, \dots, x_n)$ no depende de θ

$$\frac{f_{X_1, X_2, X_3, \dots, X_n}(x_1, x_2, x_3 \dots x_n, \theta)}{g(u_1(x_1, x_2, x_3, \dots, x_n, \theta))} = H(x_1, x_2, x_3, \dots, x_n)$$

Suficiencia de r_i

Partiendo del supuesto de independencia local:

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = \prod_{j=1}^k P(Y_i | \theta_i, \alpha_j, \beta_j)$$

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = \prod_{j=1}^k \frac{e^{\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}}$$

Resolviendo:

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = \frac{e^{\theta_i \sum_{j=1}^k y_{ij} \alpha_j - \sum_{j=1}^k y_{ij} \alpha_j \beta_j}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]}$$

Sea: $r_i = \sum_{j=1}^k y_{ij} \alpha_j$:

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = \frac{e^{\theta_i r_i - \sum_{j=1}^k y_{ij} \alpha_j \beta_j}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]}$$

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = \frac{e^{\theta_i r_i}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]} e^{-\sum_{j=1}^k y_{ij} \alpha_j \beta_j}$$

Sea $h(y) = e^{-\sum_{j=1}^k y_{ij} \alpha_j \beta_j}$ y $l(g(r_i | \theta_i)) = \frac{e^{\theta_i r_i}}{\prod_{j=1}^k [1 + e^{\alpha_j(\theta_i - \beta_j)}]}$

$$P(Y_i | \theta_i, \alpha_j, \beta_j) = l(g(r_i | \theta_i)) h(y)$$

Entonces por el teorema de factorización de Neyman la estadística

$r_i = \sum_{j=1}^k y_{ij} \alpha_j$ es una estadística suficiente de θ .

Suficiencia de c_j

Partiendo de la función verosimilitud del modelo.

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]}$$

Sea $c_j = \sum_{i=1}^n y_{ij}$

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i - \sum_{j=1}^k c_j \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]}$$

$$L(Y_i|\theta_i, \alpha, \beta) = \frac{e^{-\sum_{j=1}^k c_j \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]} e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i}$$

Hacemos: $l(g(c_j, \beta_j)) = \frac{e^{-\sum_{j=1}^k c_j \alpha_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^k [1 + e^{\alpha_j (\theta_i - \beta_j)}]}$ y $h(y) = e^{\sum_{i=1}^n \sum_{j=1}^k y_{ij} \alpha_j \theta_i}$

$$L(Y_i|\theta_i, \alpha, \beta) = l(g(c_j, \beta_j))h(y)$$

Por el teorema de factorización de Neyman la estadística $c_j = \sum_{i=1}^n y_{ij}$ es una estadística suficiente de β .

A.2 Teoría Clásica de los Test

A.2.1 Proporción de aciertos

Con la proporción de aciertos (p) analizamos que tanto han acertado al ítem los estudiantes, una forma de cuantificarlo es a través de la frecuencia relativa del ítem:

$$p_j = \frac{a_j}{n}$$

Donde:

p_j : Proporción de aciertos a j - esimo ítem

a_j : Cantidad de estudiantes que respondieron correctamente al ítem

n : Cantidad de estudiantes evaluado con la prueba

A.2.2 Índice de discriminación

Se dice que un ítem tiene poder discriminativo si distingue entre aquellos estudiantes que puntúan alto en la prueba y los que puntúan bajo. Se define como la correlación de la puntuación obtenida en el ítem y la puntuación total que obtiene el estudiante en la prueba. Como se correlacionan una variable dicotómica con una variable de escala de razón el tipo de correlación a utilizar es la correlación biseral puntual y define como sigue (ver en Anexo B la sintaxis utilizada):

$$r_{bpj} = \frac{\mu_{j1} - \mu_{j0}}{\sigma_x} \left[\sqrt{\frac{n_{j1}n_{j0}}{n(n-1)}} \right]$$

Donde:

r_{bpj} : Coeficiente de correlación biseral puntual del j – esimo ítem

μ_{j1} : Promedio de puntuaciones cuando el estudiante acierta al j – esimo ítem

μ_{j0} : Promedio de puntuaciones cuando el estudiante falla al j – esimo ítem

n_{j1} : Cantidad de estudiantes que aciertan al j – esimo ítem

n_{j0} : Cantidad de estudiantes que fallan al j – esimo ítem

σ_x : Desviación estándar muestral de las puntuaciones obtenidas por el estudiante en la prueba

n : Cantidad de estudiantes evaluados

Adicionalmente mediante estadística:

$$t_j = r_{bpj} \sqrt{\frac{n-2}{1-r_{bpj}^2}} \sim t_{(\alpha, n-1)}$$

Se prueba a un nivel de significancia del 5% la siguiente hipótesis:

$H_0 : r_{bpj} = 0$ (No existe correlación entre la puntuación de j esimo item y la puntuación total)

$H_0 : r_{bpj} \neq 0$ (Existe correlación entre la puntuación de j esimo item y la puntuación total)

A.2.3 Unidimensionalidad de la prueba

La unidimensionalidad de la prueba implica que un solo rasgo latente se encuentra en todos los ítems. Para ella la analizamos mediante el análisis factorial, en la cual buscamos que la mayor cantidad de varianza se vea explicado por el primer factor.

El punto de partida del análisis factorial es la matriz de correlación a utilizar, como los ítems son dicotomizados utilizamos la matriz de correlación tetracórica.

La correlación tetracórica se utiliza cuando correlacionamos dos variables dicotomizados, para una tabla de contingencia de 2 x 2 se define como sigue:

		X	
		0	1
Y	1	A	B
	0	C	D

De la tabla de contingencia si: $B \cdot C > A \cdot D$ entonces:

$$R_t = \cos\left(\frac{180}{1 + \sqrt{\frac{B \cdot C}{A \cdot D}}}\right)$$

Caso contrario:

$$R_t = \cos\left(\frac{180}{1 + \sqrt{\frac{A \cdot D}{B \cdot C}}}\right)$$

Ver en Anexo B la sintaxis utilizada para el cálculo de la correlación tetracórica. Con la matriz de correlación tetracórica se corrió el modelo factorial (ver sintaxis en Anexo B).

A.2.4 Confiabilidad de la prueba

La confiabilidad de la prueba se cuantifica mediante el alpha de cronbach el cual se define como:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{j=1}^k \sigma_j^2}{\sigma_x^2} \right)$$

Donde:

σ_j^2 : Varianza del j – esimo ítem

σ_x^2 : Varianza del puntaje total

k : Número de ítems en la prueba

A.3 Método de jackknife

Tomando la literatura de (Efron B. and Tibshirani R.J., 1993):

Se tiene una muestra aleatoria $X = (x_1, x_2, x_3, \dots, x_n)$ y un estimador $\theta = h(x)$, se busca estimar el error de θ . El método jackknife se enfoca sobre muestras en la cual se deja un elemento en un instante dado:

$$X_{(i)} = (x_1, x_2, x_3, \dots, x_{(i-1)}, x_{(i+1)}, x_{(i+2)}, \dots, x_n)$$

Para todo $i = 1, 2, 3, \dots, n$ es llamada muestra Jackknife. $\theta_{(i)} = h(x_{(i)})$ es la i -ésima replicación jackknife de θ . El error estándar mediante el método jackknife se define como:

$$S(\theta_{(i)})_{Jackknife} = \sqrt{\frac{n-1}{n} \sum_{i=1}^n (\theta_{(i)} - \theta_{(.)})^2}$$

Donde:

$$\theta_{(.)} = \frac{\sum_{i=1}^n \theta_{(i)}}{n}$$

A.4 Distribución de ítems en función a las capacidades y desempeños

Cuadro N° A.4

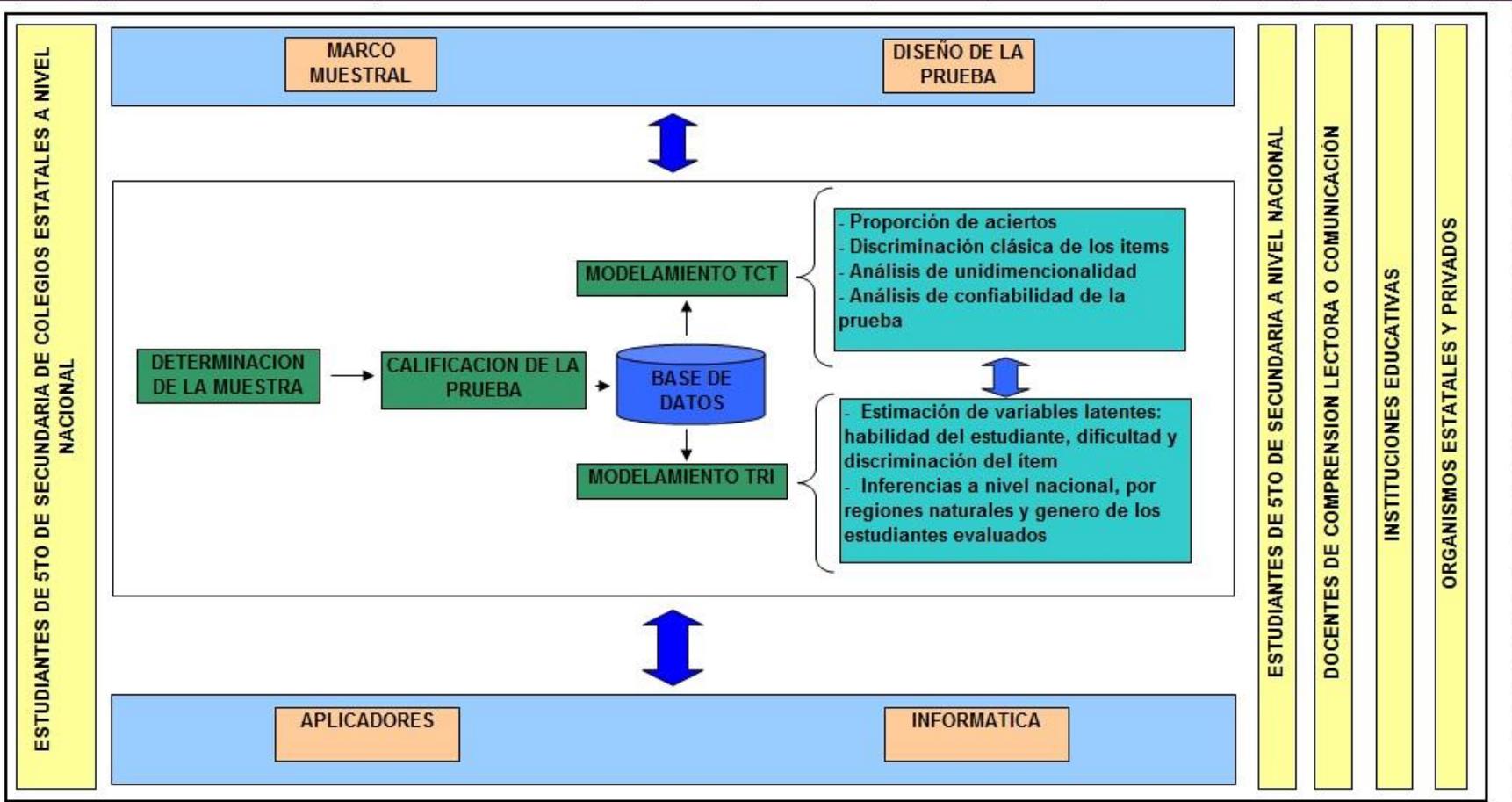
Distribución de ítems según capacidades y desempeños en la prueba de comprensión lectora

ITEM	CAPACIDAD	DESEMPEÑO
P01	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P02	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P03	Reflexiona acerca del texto	Reflexiona acerca del propósito que tuvo el autor al escribir el texto o al incluir alguna de sus partes
P04	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P05	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P06	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P07	Hace inferencias sobre información del texto	Deduce la idea central o las ideas principales del texto
P08	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P09	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P10	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P11	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P12	Hace inferencias sobre información del texto	Deduce relaciones semánticas implícitas del texto
P13	Hace inferencias sobre información del texto	Deduce relaciones semánticas implícitas del texto
P14	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P15	Hace inferencias sobre información del texto	Deduce relaciones semánticas implícitas del texto
P16	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P17	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P18	Obtiene información explícita del texto	Obtiene información ubicada en un cuadro, tabla o gráfico
P19	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P20	Reflexiona acerca del texto	Reflexiona acerca del propósito que tuvo el autor al escribir el texto o al incluir alguna de sus partes
P21	Reflexiona acerca del texto	Reflexiona acerca del propósito que tuvo el autor al escribir el texto o al incluir alguna de sus partes
P22	Hace inferencias sobre información del texto	Infiere el significado de palabras o frases del texto tomando en cuenta su contexto
P23	Obtiene información explícita del texto	Obtiene información ubicada en un párrafo del texto
P24	Hace inferencias sobre información del texto	Deduce relaciones semánticas implícitas del texto
P25	Reflexiona acerca del texto	Extrapolación de situaciones o características de personajes del texto a casos de la vida cotidiana

FUENTE: Chincaro Del Coral, Omar (2010)

ELABORACION: Propia

A.5 Esquema de modelamiento estadístico



FUENTE: Chincaro Del Coral, Omar (2010)
 ELABORACION: Propia

Anexo B

B.1 Sintaxis análisis de ítems bajo el enfoque TCT

La sintaxis del análisis de ítems bajo el enfoque TCT ha sido desarrollada en el software estadístico SPSS, la cual muestra las proporciones de acierto, el índice de discriminación mediante la correlación biseral puntual, el test estadístico para la discriminación y la confiabilidad de la prueba mediante al alfa de cronbach.

```
*RECODIFICACION DE VALORES PERDIDOS .
RECODE
  C1 TO C25 (SYSMIS=9) .
EXECUTE .

SET MXLOOPS 10000 MITERATE 10000 .
MATRIX .
GET X /FILE=* /NAMES=VARNAME/ VARIABLES=C1 TO PUNTAJE .
COMPUTE N=NROW(X) .
COMPUTE K=NCOL(X) .
COMPUTE NX1=varname(1:K).
COMPUTE ID=IDENT(N) .
COMPUTE UNOS=DIAG(ID) .
COMPUTE MEDIAX=(1/N)*T(X(:,K))*UNOS .

/*CALCULO DE LA DESVIACION ESTANDAR*/
LOOP #I=1 TO N .
  COMPUTE SUMA=T(X(#I : #I,K) - MEDIAX)*(X(#I : #I,K) - MEDIAX) .
  DO IF (#I=1) .
    COMPUTE S={SUMA} .
  ELSE .
    COMPUTE S={S;SUMA} .
  END IF .
END LOOP .

COMPUTE SIGMA=SQRT(CSUM(S)/(N-1)) .

/*CALCULO DE LA PROPORCION DE ACIERTOS Y EL INDICE DE DISCRIMINACION*/
LOOP #J=1 TO (K-1) .
  COMPUTE N0=0 .
  COMPUTE N1=0 .
```

```

LOOP #I=1 TO N .
  DO IF (X(#I:#I,#J:#J)=0) .
    COMPUTE N0=N0+1 .
    COMPUTE V=X(#I:#I,K:K) .
    DO IF (N0=1) .
      COMPUTE R0={V} .
    ELSE .
      COMPUTE R0={R0;V} .
    END IF .
  END IF .

  DO IF (X(#I:#I,#J:#J)=1) .
    COMPUTE N1=N1+1 .
    COMPUTE V=X(#I:#I,K:K) .
    DO IF (N1=1) .
      COMPUTE R1={V} .
    ELSE .
      COMPUTE R1={R1;V} .
    END IF .
  END IF .
END LOOP .

DO IF (N1<>0) .
  COMPUTE MEDIA_1=CSUM(R1)/N1 .
ELSE .
  COMPUTE MEDIA_1=0 .
END IF .

DO IF (N0<>0) .
  COMPUTE MEDIA_0=CSUM(R0)/N0 .
ELSE .
  COMPUTE MEDIA_0=0 .
END IF .

COMPUTE N2=N0+N1 .
COMPUTE PU=(N1/N2)*100 . /*PROPORCION DE ACIERTOS AL ITEM*/
COMPUTE R=((MEDIA_1 - MEDIA_0)/SIGMA)*SQRT((N0*N1)/(N2*(N2-1))) . /*INDICE DE
DISCRIMINACION (CORRELACION BISERAL PUNTUAL)*/

COMPUTE STAT=R*SQRT((N2-2)/(1-R*R)) . /*ESTADISTICO DE PRUEBA DE LA CORRELACION
BISERAL PUNTUAL*/

COMPUTE Sig=1-tcdf(STAT,(N2-1)) .

COMPUTE RESULT_0={PU,R,STAT,Sig} .

DO IF (#J=1) .
  COMPUTE F={RESULT_0} .
ELSE .
  COMPUTE F={F;RESULT_0} .
END IF .
END LOOP .

/*CALCULO DE LA VARIANZA DE CADA ITEM Y LA CONFIABILIDAD DEL INSTRUMENTO DE
MEDICION*/
COMPUTE K2=K-1 .
COMPUTE ACIERT=F(:,1) .
COMPUTE ALPHA=(K2/(K2-1))*(1-((T((ACIERT/100))*(1-(ACIERT/100))))/(SIGMA*SIGMA)) .

PRINT F/format "f10.3"/title '      ANALISIS DE ITEMS BAJO TCT'
      /space 5 /RNames=NX1/CNames={"% ACIERTOS","DISCRIMINACION","ESTADISTICO"," P
VALUE"}.

PRINT {ACIERT,100-ACIERT} .
PRINT ALPHA .
END MATRIX .

```

B.2 Sintaxis Correlación Tetracórica

La sintaxis Correlación Tetracórica es desarrollada en el software estadístico SPSS, la cual sirve para correr el análisis factorial con esta matriz de correlación y probar la unidimensionalidad de la prueba.

```
*CORRELACION TETRACORICA .
```

```
*RECODIFICACION DE VALORES PERDIDOS .
```

```
RECODE  
  C1 TO C25 (SYSMIS=9) .  
EXECUTE .
```

```
SET MXLOOPS 150000 MITERATE 150000 .
```

```
MATRIX .  
GET X /FILE=* /NAMES=VARNAME/ VARIABLES=C1 TO C25.  
COMPUTE N=NROW(X) .  
COMPUTE K=NCOL(X) .  
COMPUTE NX1=varname(1:K).
```

```
COMPUTE RTETRA=IDENT(K) .
```

```
LOOP #J=1 TO K-1 .  
  LOOP #H=#J+1 TO K .  
    COMPUTE CONT1=0 .  
    COMPUTE CONT2=0 .  
    COMPUTE CONT3=0 .  
    COMPUTE CONT4=0 .  
    LOOP #I=1 TO N .  
      DO IF ((X(#I,#J)=1) AND (X(#I,#H)=0)) .  
        COMPUTE CONT1=CONT1+1 .  
        COMPUTE N1=1 .  
        DO IF (CONT1=1) .  
          COMPUTE VA={N1} .  
          ELSE .  
            COMPUTE VA={VA;N1} .  
        END IF .  
      END IF .  
      DO IF ((X(#I,#J)=1) AND (X(#I,#H)=1)) .  
        COMPUTE CONT2=CONT2+1 .  
        COMPUTE N2=1 .  
        DO IF (CONT2=1) .  
          COMPUTE VB={N2} .  
          ELSE .  
            COMPUTE VB={VB;N2} .  
        END IF .  
      END IF .  
      DO IF ((X(#I,#J)=0) AND (X(#I,#H)=0)) .  
        COMPUTE CONT3=CONT3+1 .  
        COMPUTE N3=1 .  
        DO IF (CONT3=1) .  
          COMPUTE VC={N3} .  
          ELSE .  
            COMPUTE VC={VC;N3} .  
        END IF .  
      END IF .
```

```

DO IF ((X(#I,#J)=0) AND (X(#I,#H)=1)) .
  COMPUTE CONT4=CONT4+1 .
  COMPUTE N4=1 .
  DO IF (CONT4=1) .
    COMPUTE VD={N4} .
    ELSE .
    COMPUTE VD={VD;N4} .
  END IF .
END IF .
END LOOP .
COMPUTE A=CSUM(VA) .
COMPUTE B=CSUM(VB) .
COMPUTE C=CSUM(VC) .
COMPUTE D=CSUM(VD) .
DO IF ((A*D)<(B*C)) .
  COMPUTE RTET=COS((3.1416/180)*(180/(1+SQRT((B*C)/(A*D)))))) .
ELSE .
  COMPUTE RTET=COS((3.1416/180)*(180/(1+SQRT((A*D)/(B*C)))))) .
END IF .
COMPUTE RTETRA(#J,#H)=RTET .
COMPUTE RTETRA(#H,#J)=RTET .
END LOOP .
END LOOP .
PRINT RTETRA / FORMAT "F7.4" / TITLE '  CORRELACION TETRACORICA' /CNAMES=NX1 .
PRINT DET(RTETRA) .
SAVE RTETRA /OUTFILE 'C:\Archivos de programa\SPSS\TETRACORICA.sav' .
END MATRIX .

GET
FILE='C:\Archivos de programa\SPSS\TETRACORICA.sav'.

```

B.3 Script Método Jackknife

El script método Jackknife se ha elaborado en el software estadístico R

```
#Cargar libreria bootstrap
```

```
require(bootstrap)
```

```
#Errores por regiones naturales mediante jackknife
```

```
#####
```

```

hab<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABNAC.csv",sep=",")
habrn1<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABCOSTA.csv",sep=",")
habrn2<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABSIERRA.csv",sep=",")
habrn3<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABSELVA.csv",sep=",")
habfem<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABFEM.csv",sep=",")
habmasc<-read.table("E:/CURSO DE TITULACION/BASES DE DATOS/HABMASC.csv",sep=",")
theta=hab[,1]
theta_r1=habrn1[,1]
theta_r2=habrn2[,1]
theta_r3=habrn3[,1]
theta_r4=habfem[,1]
theta_r5=habmasc[,1]
r<-jackknife(theta,mean)$jack.se
r1<-jackknife(theta_r1,mean)$jack.se
r2<-jackknife(theta_r2,mean)$jack.se
r3<-jackknife(theta_r3,mean)$jack.se
r4<-jackknife(theta_r4,mean)$jack.se
r5<-jackknife(theta_r5,mean)$jack.se
ern<-c(r,r1,r2,r3,r4,r5)
error_rn=matrix(ern,6,1)
error_rn

```

B.3 Código Modelo TRI dicotómico de dos parámetros o modelo de Birnbaum

El código para el modelo TRI dicotómico de dos parámetros ha sido elaborado en el software estadístico WinBugs

```

model
{
  for (i in 1 : n)
  {
    for (j in 1 : l)
    {
      m[i,j] <- a[j]*(u[i] - b[j])
      p[i,j] <- exp(m[i,j])/(1 + exp(m[i,j]))
      y[i,j] ~ dbern(p[i,j])
    }
  }

  #priors F for item parameters
  for (j in 1:l)
  {
    a[j] ~ dlnorm(0,1);
    b[j] ~ dnorm(0,1);
  }

  #prior for latent variable
  for (i in 1:n)
  {
    u[i] ~ dnorm(0,1)
  }

  # mean and standard deviation for latent variable
  mu<-mean(u[])
  du<-sd(u[])
}

Data list(n=1223, l=25,
          y= structure(
            .Data = c(1,1,0,0,1,0,1,0,1,1,1,0,1,1,1,1,1,1,0,0,1,0,0,1,0,
                      1,1,0,0,0,0,0,1,0,1,1,0,1,0,1,1,1,1,0,0,0,0,0,1,0,
                      .....
                      .....
                      .....
                      1,0,0,1,0,0,0,1,0,0,0,0,0,0,1,1,0,0,1,0,0,0,1,0),
            .Dim = c(1223,25))
          )

Inits list(b = c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))

```

Instrumento de Medición

Comprensión Lectora

5°
Grado

Prueba
1

DATOS DE LA INSTITUCION EDUCATIVA :

Nombre de la Institución Educativa

Apellido paterno

Apellido materno

Departamento

Provincia

Nombres

Grado

Distrito

CÓDIGO DE I.E.

Sección

CÓDIGO DE ESTUDIANTE

INDICACIONES PARA RESPONDER LA PRUEBA

1. En esta prueba encontrarás diversos textos con sus respectivas preguntas. Lee con calma y atención cada texto y cada pregunta.
2. Para responder las preguntas, vuelve a leer el texto, si es necesario.
3. Si te demoras mucho en responder alguna pregunta, pasa a la siguiente. Cuando termines, podrás regresar a las preguntas que no has respondido.
4. Para responder la pregunta, marca la respuesta que crees que es correcta.
5. Si te has equivocado en marcar tu respuesta de una pregunta, puedes marcar otra; pero tienes que anular la respuesta anterior.
6. Recuerda que las preguntas tienen sólo una respuesta verdadera.
7. Durante la lectura, puedes subrayar, marcar o dejar anotaciones en los textos, si deseas.

Si no entiendes algo, levanta la mano y consulta a la persona que está tomando la prueba.

ESPERA LAS INDICACIONES PARA COMENZAR

Martes, 18 de septiembre de 2007

Flores para Lima

El 21 de septiembre se abre en Lima la feria Perúflora 2007, en la que se exhibirán especies diferentes y se analizarán las perspectivas de desarrollo que tiene el sector.

Uno de los momentos más esperados en Nueva York todos los años es cuando su tienda por departamentos más importante, Macy's, hace un gran show de flores. La tienda, localizada en el corazón de la ciudad, se viste de flores exóticas. Durante toda una semana la población abarrotó el lugar para verlas y, luego, comprarlas.

Esto refleja una de las tendencias más en boga del momento: la creciente demanda de flores (tanto tradicionales como exóticas) en ciudades como Nueva York, Londres, París, Milán y Tokio. Cabe precisar que gran parte de dicha oferta es importada.

En Latinoamérica, los países favoritos de donde se piden las flores son Colombia y Ecuador. En Colombia, por ejemplo, se han especializado en las ventas a Europa. Allí ya han desarrollado cadenas productivas, incluyendo aviones propios.

Aunque el Perú tiene el potencial para entrar a competir fuerte en ese mercado, aún no lo ha hecho. El país tiene los suficientes microclimas para producir todo tipo de flores. La selva es un diamante en bruto por explotar.

Las principales empresas exportadoras peruanas comercializaron 1,304 toneladas de flores en el exterior en 2006. Entre estas, destacan "Vivero los Inkas" y "Florisert". Entre los países de destino de la producción local figuran EE.UU, Italia y Holanda, con el 52%, 20% y 13% del total, respectivamente.

FERIA. Por eso, con la llegada a Lima de Perúflora 2007, que se realizará en el Parque Central de Miraflores del 21 al 23 de setiembre, organizada por la Asociación Peruana de Arquitectura del Paisaje y la Municipalidad de Miraflores, los interesados en el negocio podrán ver todas las oportunidades que tiene el país en el desarrollo de este sector.

Con el fin de activarlo y presentar todas las perspectivas en el ámbito local e internacional, se darán conferencias sobre los retos para exportar, cuáles son las flores con más demanda, qué tipos se pueden cultivar en el Perú, entre otros temas.

Es, sin duda, una magnífica oportunidad para entrar a un negocio con grandes perspectivas de crecimiento y que, además, es cuidadoso con el medio ambiente.



La feria permitirá conocer las

Ahora responde las preguntas sobre la noticia que has leído.

1. ¿En dónde queda la tienda por departamentos Macy's?
 - A. En Nueva York
 - B. En Lima
 - C. En Colombia
 - D. En París

2. ¿Qué porcentaje del total de toneladas de flores exportadas por el Perú son vendidas a Holanda?
 - A. El 13%
 - B. El 20%
 - C. El 21%
 - D. El 52%

3. Lee con atención el siguiente párrafo del texto:

“En Latinoamérica, los países favoritos de donde se piden las flores son Colombia y Ecuador. En Colombia, por ejemplo, se han especializado en las ventas a Europa. Allí ya han desarrollado cadenas productivas, incluyendo aviones propios.”

¿Para qué crees que el autor del texto menciona que los negocios de flores en Colombia cuentan con aviones propios?

- A. Para dar a entender que los aviones son imprescindibles en los negocios de importaciones.
 - B. Para resaltar la idea de que los negocios de flores traen mucho dinero a sus comerciantes.
 - C. Para mostrar que los países a los que se exportan las flores están muy lejos del Perú.
 - D. Para demostrar que Colombia es un país muy desarrollado en comparación con el Perú.

4. ¿Cuál fue el propósito principal de realizar la feria Perúflora 2007?
 - A. Impulsar a las familias peruanas a cultivar flores y cuidar el medio ambiente.

- B. Exponer al público diferentes especies de flores que crecen en el Perú.
 - C. Activar el mercado de exportación de flores de nuestro país.
 - D. Vender una gran cantidad de flores en la ciudad de Lima.
5. ¿Qué quiere decir el autor del texto cuando menciona que “la selva es un diamante en bruto por explotar”?
- A. Que en la selva hay una gran explotación de flores para la venta en el exterior.
 - B. Que es necesario educar a la población de la selva para que se desarrolle.
 - C. Que la selva es un lugar con mucho potencial para el cultivo de flores.
 - D. Que en la selva hay mucha riqueza proveniente de los diamantes.

PASA AL SIGUIENTE TEXTO

La censura en Internet

El hecho de que Internet permita posibilidades de comunicación irrestrictas es valiosísimo, y por eso, ¿es pertinente la aparición de censuras o legislaciones precisas que la normen o puede existir el libre albedrío en un medio de comunicación que, por su naturaleza, apunta a la anarquía?

Internet se ha convertido en la tribuna por excelencia de la libertad de expresión. Debido a que no existe ninguna entidad que vigile y norme los contenidos circulantes, cualquier persona u organismo tiene la posibilidad de expresar sus ideas o creencias a través de la red de redes, incluso aquellas que son proscritas por la sociedad y que, a lo largo de los años, han sufrido la censura y el control de los gobiernos y de las instituciones reconocidas, por ser cuestionables desde el punto de vista moral y político.

El asunto es espinoso y difícil, más aún si tomamos en cuenta que los contenidos ‘peligrosos’ de la red son de los más diversos tipos. Grupos terroristas, sanguinarios e intolerantes como el MRTA o la ETA se hacen presentes paulatinamente en Internet, usándola como caja de resonancia de sus planteamientos y como receptora de ‘donaciones’ en efectivo para la causa. Los principios de los grupos neonazis han encontrado también una tribuna abierta en el ciberespacio y fomentan la discriminación y la xenofobia a partir de la premisa de la superioridad de la raza aria.

Los fanáticos religiosos tampoco han perdido tiempo en la carrera de información mundial a través de la red. Aparentemente, muchas sectas consiguen gran parte de sus fanáticos y adeptos por medio de Internet. Tal es el caso de la secta La Puerta del Cielo, cuyos miembros se suicidaron masivamente al creer que, tras el cometa Halle Bopp, se ocultaba una nave espacial que los transportaría a un reino superior.

Las medidas propuestas para solucionar el dilema de la libertad irrestricta de información han sido bastante diversas. Por un lado, algunas instituciones han propuesto fórmulas legislativas de prevención basadas en licencias o registros de abonados. Muchos gobiernos, por su parte, promueven, sin empacho, la necesidad de censura y mano dura contra cualquier intento insurgente. Por último, algunos sociólogos e investigadores, teniendo en cuenta que los vetos y prohibiciones pueden ser contraproducentes, se inclinan más bien por responder con las mismas armas: un web site con contenidos inmorales o intolerantes puede ser atacado por medio de otros donde se brinde la cara de la información que ellos no presentan.

El eterno dilema humano entre la libertad y la censura seguirá su curso y permitirá saber si el ser humano es capaz de autonormarse o si requiere de limitaciones impuestas por algún organismo de control, pues el problema no es realmente Internet, sino más bien el mundo que la naturaleza humana ha creado. Internet no es más que una vía de acceso por la cual acceder más fácilmente a él. Los diversos problemas de la sociedad ya estaban allí desde antes.

Ahora responde algunas preguntas sobre el texto que acabas de leer.

6. Según el texto, ¿qué es lo que proponen ciertos sociólogos e investigadores para solucionar los problemas causados por la libertad irrestricta de la información en Internet?

- A. Formular leyes de prevención basadas en licencias o registros de abonados.
- B. Crear sitios web que brinden información en contra de otros sitios web inmorales.
- C. Poner mano dura y censurar cualquier intento insurgente de publicar contenidos inmorales.
- D. Usar Internet como caja de resonancia de sus planteamientos y para recaudar donaciones.

7. ¿Cuál es la idea principal del texto?

- A. No existe ninguna entidad gubernamental que vigile y norme los contenidos circulantes en Internet.
- B. La Internet replantea al mundo el eterno dilema humano entre la libertad irrestricta de expresión y la censura.
- C. Es necesario tomar medidas drásticas que impidan la difusión de contenidos peligrosos o inmorales en Internet.
- D. La Internet ha permitido que diversos grupos de terroristas y fanáticos religiosos difundan sus ideas sin restricciones.

8. Lee atentamente la siguiente oración extraída del texto:

“...cualquier persona u organismo tiene la posibilidad de expresar sus ideas o creencias a través de la red de redes, incluso aquellas que son proscritas por la sociedad y que, a lo largo de los años, han sufrido la censura y el control de los gobiernos y de las instituciones reconocidas, por ser cuestionables desde el punto de vista moral y político.”

La palabra subrayada del texto se refiere a:

- A. las ideas o creencias
- B. las personas o redes
- C. las censuras
- D. las instituciones

Lee con atención el siguiente párrafo extraído del texto y responde las preguntas:

“El eterno dilema humano entre la libertad y la censura seguirá su curso y permitirá saber si el ser humano es capaz de autonormarse o si requiere de limitaciones impuestas por algún organismo de control, pues el problema no es realmente Internet, sino más bien el mundo que la naturaleza humana ha creado. Internet no es más que una vía de acceso por la cual acceder más fácilmente a él. Los diversos problemas de la sociedad ya estaban allí desde antes.”

9. ¿A qué se refiere la palabra “él”?

- A. Internet
- B. El mundo
- C. El ser humano
- D. El organismo de control

10. ¿A qué se refiere la palabra “antes”?

- A. A antes de que la humanidad pueda autonormarse.
- B. A antes de que hubiera un organismo de control.
- C. A antes de que apareciera la Internet.
- D. A antes de la censura.

PASA AL SIGUIENTE TEXTO

Una noche de verano

El hecho de que Henry Armstrong estuviera enterrado no era un motivo suficientemente convincente como para demostrarle que estaba muerto: siempre había sido un hombre difícil de persuadir. El testimonio de sus sentidos le obligaba a admitir que estaba realmente enterrado. Su posición —tendido boca arriba con las manos cruzadas sobre su estómago y atadas con una débil soga que rompió fácilmente sin que se alterase la situación—, el estricto confinamiento de toda su persona, la negra oscuridad y el profundo silencio constituían una evidencia imposible de contradecir y Armstrong lo aceptó sin perderse en cavilaciones.

Pero, muerto... no. Solamente estaba enfermo, muy enfermo, aunque, con la apatía del inválido, no se preocupó demasiado por la extraña suerte que le había correspondido. No era un filósofo, sino simplemente una persona vulgar, dotada en aquel momento de una patológica indiferencia; el órgano que le había dado ocasión de inquietarse estaba ahora aletargado. De modo que sin ninguna aprensión por lo que se refiriera a su futuro inmediato, se quedó dormido y todo fue paz para Henry Armstrong.

Pero algo todavía se movía en la superficie. Era aquella una oscura noche de verano, rasgada por frecuentes relámpagos que iluminaban unas nubes, las cuales avanzaban por el este preñadas de tormenta. Aquellos breves y relampagueantes fulgores proyectaban una fantasmal claridad sobre los monumentos y lápidas del camposanto. No era una noche propicia para que una persona normal anduviera vagabundeando alrededor de un cementerio, de modo que los tres hombres que estaban allí, cavando en la tumba de Henry Armstrong, se sentían razonablemente seguros.

Dos de ellos eran jóvenes estudiantes de una Facultad de Medicina que se hallaba a unas millas de distancia; el tercero era un gigantesco negro llamado Jess. Desde hacía muchos años Jess estaba empleado en el cementerio en calidad de sepulturero, y su chanza favorita era la de que «conocía todas las ánimas del lugar». Por la naturaleza de lo que ahora estaba haciendo, podía inferirse que el lugar no estaba tan poblado como su libro de registro podía hacer suponer.

Al otro lado del muro, apartados de la carretera, podían verse un caballo y un carruaje ligero, esperando.

El trabajo de excavación no resultaba difícil; la tierra con la cual había sido rellena la tumba unas horas antes ofrecía poca resistencia, y no tardó en quedarse amontonada a uno de los lados de la fosa. El levantar la tapadera del ataúd requirió más esfuerzo, pero Jess era práctico en la tarea y terminó por colocar cuidadosamente la tapadera sobre el montón de tierra, dejando al descubierto el cadáver, ataviado con terno negro y camisa blanca.

En aquel preciso instante, un relámpago zigzagueó en el aire, desgarrando la oscuridad, y casi inmediatamente estalló un fragoroso trueno. Arrancado de su sueño, Henry Armstrong incorporó tranquilamente la mitad superior de su cuerpo hasta quedar sentado.

Profiriendo gritos inarticulados, los hombres huyeron, poseídos por el terror, cada uno de ellos en una dirección distinta. Dos de los fugitivos no hubieran regresado por nada del mundo. Pero Jess estaba hecho de otra pasta.

Con las primeras luces del amanecer, los dos estudiantes, pálidos de ansiedad y con el terror de su aventura latiendo aun tumultuosamente en su sangre, llegaron a la Facultad.

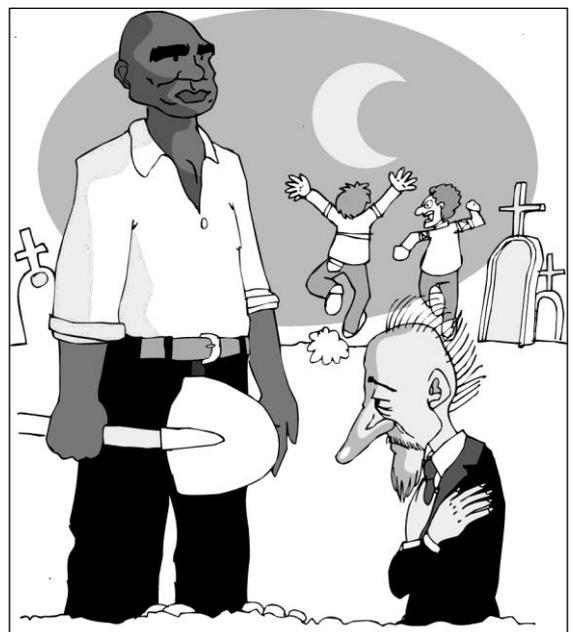
— ¿Lo has visto? —exclamó uno de ellos—.

— ¡Dios! Sí... ¿Qué vamos a hacer?

Se encaminaron a la parte de atrás del edificio, donde vieron un carruaje ligero con un caballo uncido y atado a una verja, cerca de la sala de disección. Maquinalmente, los dos jóvenes entraron en la sala. Sentado en un banco, a oscuras, vieron al negro Jess. El negro se puso de pie, sonriendo, todo ojos y dientes.

— Estoy esperando mi paga —dijo—.

Desnudo sobre una larga mesa, yacía el cadáver de Henry Armstrong. Tenía la cabeza manchada de sangre y arcilla por haber recibido un golpe de azada.



Ahora responde las siguientes preguntas sobre el cuento que acabas de leer. Para contestarlas, regresa al cuento y utiliza la información que necesites.

11. ¿Qué hizo Henry Armstrong cuando se dio cuenta de que estaba enterrado bajo tierra?
- A. Incorporó la mitad superior de su cuerpo.
 - B. Se quedó dormido tranquilamente.
 - C. Intentó desesperadamente de desenterrarse.
 - D. Fue alcanzado por un fragoroso trueno.
12. ¿Por qué Jess y los dos estudiantes no tuvieron mayor dificultad al momento de cavar la tierra de la tumba de Henry Armstrong?
- A. Porque Jess era un sepulturero con mucha experiencia.
 - B. Porque los tres hombres se ayudaron en la tarea.
 - C. Porque habían enterrado a Henry Armstrong hacía poco tiempo.
 - D. Porque Jess era un hombre muy fuerte que no tenía miedo a nada.
13. ¿Por qué en el cuento se afirma que “Por la naturaleza de lo que ahora estaba haciendo, ¿podía inferirse que el lugar no estaba tan poblado como su libro de registro podía hacer suponer”?
- A. Porque Jess se dedicaba a comerciar con cadáveres.
 - B. Porque Jess no sabía llevar bien las cuentas.
 - C. Porque Jess había asesinado a mucha gente sin registrarla.
 - D. Porque Jess en realidad no conocía a todas las ánimas del lugar.
14. Lee atentamente el siguiente fragmento del cuento:
- “El hecho de que Henry Armstrong estuviera enterrado no era un motivo suficientemente convincente como para demonstrarle que estaba muerto: siempre había sido un hombre difícil de persuadir”.
- ¿A quién se refiere la parte subrayada de la palabra?
- A. A Jess
 - B. Al cadáver
 - C. Al narrador
 - D. A Henry Armstrong
15. Henry Armstrong murió:
- A. enterrado vivo.
 - B. de un fuerte golpe en la cabeza.
 - C. de una grave enfermedad.
 - D. al recibir el impacto de un relámpago.

PASA AL SIGUIENTE TEXTO

El siguiente texto se publicó en un periódico. Léelo atentamente y luego responde las preguntas.

Día Mundial de la Alimentación

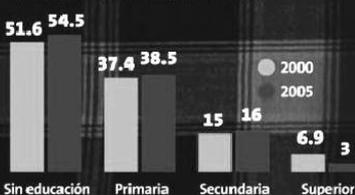
El Día Mundial de la Alimentación se celebra mañana, 16 de octubre, con el lema "Invertir en la agricultura para lograr la seguridad alimentaria". La frase sirve para reflexionar sobre el papel de la agricultura para que se pueda poner fin al hambre y a la problemática de la desnutrición. Actualmente la desnutrición infantil afecta al 24.1% de niños entre 0 y 5 años y es una de las prioridades del Gobierno, que se ha fijado como meta al 2011 reducirla en cinco puntos porcentuales. Para ayudar a esa tarea, el Instituto Nacional de Salud (INS) recordó que consumir alimentos andinos incrementa el valor nutritivo y garantiza la seguridad alimentaria, así como el desarrollo físico y mental de las personas, especialmente de los niños y gestantes.

Requerimiento calórico

- Según una encuesta realizada por el INS, el consumo promedio de energía en los peruanos alcanza solo el 76.7% de las recomendaciones mundiales, es decir, 1,576.7 kilocalorías.
- Las calorías provenientes de la grasa se encuentran dentro de los valores recomendados. Sin embargo, la proporción de calorías que proviene del consumo de carbohidratos (74%) se encuentra cerca del límite máximo recomendado (75%).
- El estudio indica que la ingesta de hierro es insuficiente para el caso de mujeres porque cubre menos del 50% de sus requerimientos. El consumo de colesterol es adecuado, pero la fibra dietaria insuficiente.

Desnutrición crónica

En niños menores de 5 años según nivel de educación de la madre



Prevalencia de anemia

En niños menores de cinco años En mujeres en edad fértil



ALIMENTOS DE SELVA

- Los productos con gran cantidad de vitamina C son la yuca, camote, mani, guayaba, piña, cacao, papaya, vainilla; camu camu, piña, carambola y guayaba.
- La carne de paiche seco, la carachama, zúngaro del monte y los huevos de charapa motelo tienen gran cantidad de proteínas.
- Entre los alimentos energéticos están: arroz, mote pelado, maíz amarillo y frijoles.
- El aguaje, mangos, pijuayo, cocona y zapote cuentan con buenas cantidades de carotenos.



Productos saludables
El nutricionista del INS Iván Gómez Sánchez explicó que entre los alimentos nutritivos resaltan:

SOBREPESO
■ El mismo estudio del INS reveló que el 35.3% de la población peruana tiene sobrepeso y son las mujeres las más afectadas (39.1% de casos).
■ El sobrepeso es menor en las personas menores de 29 años (23%). El grupo más afectado con es el que se ubica entre 40 y 49 años (41.3%).

PREVALENCIA DE DELGADEZ
■ La prevalencia de delgadez en los peruanos es baja y alcanza solo al 2.7% de la población.
■ El mayor nivel de delgadez se presenta en el grupo entre 20 y 29 años.

OBESIDAD
■ Afecta al 16.5% de la población peruana.
■ Las mujeres tienen mayor prevalencia de obesidad (20.3%) en comparación con los varones (12.6%).
■ Según los estratos, la mayor prevalencia se reporta en la costa (20.2%) y en Lima Metropolitana (18.7%) y la menor prevalencia en la sierra rural (9.2%).

16. Según el texto, ¿qué porcentaje de personas con sobrepeso son mujeres?

- A. 23%
- B. 35,3%
- C. 39,1%

D. 41,3%

17. Según el texto, ¿cuál de los siguientes alimentos tiene altos niveles de carotenos?

- A. La carachama
- B. La yuca
- C. El maíz amarillo
- D. La cocona

18. Según el texto, ¿en cuál de los siguientes años la prevalencia de anemia en mujeres en edad fértil fue mayor?

- A. 2006
- B. 2005
- C. 2000
- D. 1996

19. Según el texto, ¿qué quiere decir la oración siguiente?

“El estudio indica que la ingesta de hierro es insuficiente para el caso de las mujeres porque cubre menos del 50% de sus requerimientos.”

- A. Que las mujeres solo consumen la mitad del hierro que necesitan.
- B. Que las mujeres solo ingieren la mitad del hierro que los hombres.
- C. Que solo el 50% de lo que consumen las mujeres contiene hierro.
- D. Que las mujeres tienen mayores requerimientos de hierro que los hombres.

20. ¿Cuál es el propósito de incluir la información acerca de la desnutrición crónica en relación al nivel de educación de la madre?

- A. Dar a entender que las mujeres mal alimentadas no pueden llegar a tener una buena educación.
- B. Llamar la atención sobre el hecho de que las madres del Perú tienen bajos niveles educativos.
- C. Resaltar la idea de que las mujeres con mejor educación tienen menos hijos que las mujeres sin educación.
- D. Destacar la idea de que el nivel de educación de las madres es un factor relacionado con la desnutrición de sus hijos.

21. ¿Cuál es el propósito principal del texto?

- A. Informar sobre los bajos niveles de nutrición de la población del país, sobre todo de las gestantes y niños.
- B. Aconsejar el consumo de productos alimenticios propios del Perú que tienen grandes valores nutritivos
- C. Resaltar la importancia de la agricultura como solución a los problemas económicos del país.
- D. Mostrar a partir de datos estadísticos que la desnutrición ha disminuido desde el año 2000.

PASA AL SIGUIENTE TEXTO

El siguiente texto apareció publicado en una revista. Léelo con atención y luego responde las preguntas.

Entendemos por "racismo" toda práctica social en la que se discrimine de alguna manera a ciertos grupos sociales, por el solo hecho de poseer determinados rasgos físicos que se consideran no deseables e inferiores.

En el Perú, es evidente que las imágenes presentadas como deseables por los medios masivos de comunicación (en la televisión, prensa, revistas, radio, etc.) no corresponden con la apariencia de la mayoría de peruanos. Aquí, el racismo se dirige no en contra de una minoría, sino en contra de una mayoría, por lo que tiene efectos más desintegradores. Esto tiene como consecuencia primordial que, al ser el modelo socialmente deseado distinto al de la mayoría de peruanos, el racismo pueda dirigirse en contra de uno mismo. Y eso es lo que lo hace tan perverso y corrosivo. El racismo hace difícil aceptarse como se es, lo cual convierte a nuestro racismo en desgarrador.

Se ha intentado negar la existencia del racismo en nuestro país apelando al discurso del mestizaje. Este discurso es aparentemente democrático, pero en el fondo es útil a las élites para ocultar la desigualdad social. La mayoría de peruanos acepta que "todos tenemos de todo", pero, al mismo tiempo, sabe que cada uno tiene algo más de un lado que de otro; en otras palabras, que las mezclas son de diverso grado. Se reconoce el fuerte mestizaje, pero también se mantiene una identificación, a grandes rasgos comprobable en la realidad, entre tipos raciales, y grados de poder y estatus. Ya que el poder económico, político y social ha estado tradicionalmente en manos de blancos, cualquier persona de otra raza que logre avanzar en la complicada escalera social es vista como más blanca de lo que era anteriormente ("la plata blanquea" es un dicho popular).

22. Según lo que entiendes del texto, ¿qué quiere decir que “la plata blanquea”?

- A. Que los blancos tienen más dinero que las personas de otras razas.
- B. Que cuando alguien tiene dinero se le aclara el color de la piel.
- C. Que la gente es menos racista con las personas que tienen dinero.
- D. Que las personas pobres no tienen oportunidad de cambiar su color.

23. Según el autor del texto, ¿por qué el racismo en el Perú tiene efectos más desintegrados que el racismo de otros lugares?

- A. Porque en el Perú todos tenemos un poco de todo debido al fuerte mestizaje, sin embargo, existen desigualdades.
- B. Porque las imágenes presentadas por los medios masivos no reflejan la apariencia de las mayorías.
- C. Porque a pesar del mestizaje se mantiene una identificación entre tipos raciales y grados de poder y estatus.
- D. Porque no se dirige en contra de las minorías, sino en contra de las mayorías de los peruanos.

24. Lee con atención el siguiente párrafo extraído del texto:

“Ya que el poder económico, político y social ha estado tradicionalmente en manos de blancos, cualquier persona de otra raza que logre avanzar en la complicada escalera social es vista como más blanca de lo que era anteriormente”.

Del párrafo anterior se deduce que:

- A. En el Perú, solo las personas blancas pueden lograr avanzar en la complicada escalera social.
- B. El racismo en el Perú no solo está relacionado con los rasgos físicos, sino también con el dinero y el poder.
- C. Las personas ricas y poderosas política y socialmente en el Perú son todas de raza blanca.
- D. La escalera social se determina principalmente por el grado de mestizaje que muestra cada uno de los individuos que la conforman.

25. ¿En cuál de los siguientes casos se manifiesta el racismo de los medios de comunicación?

- A. En los comerciales de productos de belleza solo se presentan modelos con cuerpos perfectos y bellos rostros.
- B. La mayoría de narradores de noticias de programas informativos considerados “serios” son de raza blanca.
- C. Las actrices de televisión más famosas no quieren interpretar en las telenovelas los papeles de empleadas del hogar.

Una presentadora de televisión que tiene orígenes humildes y es

FIN DE LA PRUEBA