

Informe de la Evaluación Diagnóstica 2014
Dimensión Matemática
Unidad de Apoyo a la Enseñanza

Capítulo 1

Introducción

En 2014 se decide realizar una nueva prueba diagnóstica, dedicándole especial atención a la elaboración del instrumento de evaluación en la dimensión matemática.

En primer lugar se trabaja en el establecimiento de los estándares de contenido en la dimensión matemática. Para esto se conforma un grupo de discusión conformado por cinco docentes que se desempeñan o se han desempeñado en la docencia en Educación Media Superior. Por el CURE participan las docentes Ana Oribe y Alexandra Fregueiro.

Los estándares de contenido son definidos como la descripción de conocimientos o habilidades específicas que se espera que los examinados demuestren su dominio acorde a su edad, nivel o campo de estudio (Cizek, Bunch y Koons, 2004) o lo que los estudiantes deben saber y saber hacer en determinadas áreas (Tourón, 2009). Los estándares están asociados a las competencias que se pretende alcanzar.

Desarrollamos estándares de contenido en la dimensión matemática considerando un nivel suficiente en la prueba de Evaluación Diagnóstica que se aplicará a los estudiantes de ingreso en el Centro Universitario de la Región Este.

Se tomaron en cuenta y se revisaron los programas de Matemática de los últimos dos años de bachillerato tanto de Secundaria como UTU, tomándose como referencia de competencia mínima los contenidos de los programas de tercer año de bachillerato Social-humanístico de Secundaria y Turismo de UTU.

Los estándares de contenido quedaron establecidos de la siguiente forma:

1- Números y operaciones

- a. Operaciones, razón y proporción:
 - Resolver problemas de aritmética incluyendo porcentajes, razones y proporciones.
- b. Números complejos:
 - Realizar operaciones con números complejos.
 - Representar números complejos y sus operaciones en el plano complejo.
- c. Conteo:
 - Conocer los conceptos de: arreglo, permutación y combinación.
 - a. Simplificar expresiones racionales que contengan factoriales.
 - b. Calcular números combinatorios.
 - c. Resolver problemas de conteo utilizando números combinatorios.
 - d. Utilizar el diagrama de árbol en la resolución de problemas de conteo sencillos.
- e. Teoría elemental de los números naturales y enteros:
- f. Conocer las propiedades de los enteros y los naturales.
- g. Sistema de los números reales:
- h. Conocer las propiedades de los números reales.
- i. Operar con exponenciales y logaritmos.
- j. Clasificar números como racionales o irracionales.
- 2- Álgebra y funciones
 - a. Expresiones:
 - Sustituir y simplificar expresiones algebraicas.
 - Escribir expresiones en formas equivalentes.
 - b. Ecuaciones e inecuaciones:
 - Conocer operativamente el concepto de ecuaciones equivalentes.
 - Reconocer si un sistema lineal es determinado, indeterminado o incompatible.
 - Resolver problemas cuya solución conduce a un sistema de ecuaciones lineales y comprobar la validez de su solución en el contexto del problema que lo generó.
 - Resolver ecuaciones y desigualdades de una variable.
 - Resolver sistemas de ecuaciones e inecuaciones.
 - Resolver ecuaciones racionales y radicales.
 - Conocer el concepto de función algebraica.
 - Representar y resolver ecuaciones y desigualdades gráficamente.
 - c. Aritmética con polinomios y funciones racionales:
 - Realizar operaciones aritméticas en polinomios.
 - Entender la relación entre ceros y factores de un polinomio.

- Usar el teorema de identidad de polinomios para resolver problemas.
- d. Propiedades de las funciones:
 - Conocer la definición de función, dominio, codominio y recorrido.
 - Identificar la existencia del límite de una función en un punto de su dominio y calcularlo.
 - Obtener el límite de una función por aproximación de valores funcionales.
 - Calcular el límite de una función aplicando las propiedades de la suma, producto y/o división de las funciones.
 - Determinar el límite de una función dada por su gráfica.
 - Reconocer la continuidad de una función en un punto o en un intervalo a partir de su gráfica.
 - Reconocer la diferencia entre la existencia del límite de una función en un punto y su continuidad.
 - Determinar las asíntotas horizontales o verticales de las funciones cocientes de funciones polinómicas de primer grado.
 - Inferir la variación de una función a partir de la fórmula de la función y de su función derivada.
 - Conocer la relación entre derivabilidad y continuidad.
 - Determinar el crecimiento y decrecimiento de funciones.
 - Graficar funciones.
 - Comprender los conceptos de: incremento y cociente incremental de una función.
 - Reconocer la variación del cociente incremental de una función al variar el incremento de la variable.
 - Calcular el cociente incremental en un punto.
 - Deducir la derivada de las funciones polinómicas.
 - Integrar el concepto geométrico de recta tangente a una curva en uno de sus puntos.
 - Interpretar geoméricamente la derivada de una función en un punto.
 - Reconocer la derivada en un punto como indicador de la rapidez de una variación de la función en ese punto.
 - Comprender el concepto de función derivada.
 - Aplicar las fórmulas de derivación a la derivada de una función.
 - Construir la gráfica de una función a partir de condiciones dadas: límite en un punto, discontinuidades, variación, etc.
 - Deducir la variación de la función derivada del gráfico de una función utilizando el coeficiente angular de las rectas tangentes.

- e. Interpretar funciones:
 - Interpretar funciones que surjan de la aplicación en un contexto.
 - Analizar funciones usando diferentes representaciones.
- 3. Geometría
 - a. Plano Euclidiano:
 - Conocer las propiedades de los ángulos.
 - Conocer las posiciones de rectas en un plano: secantes y paralelas.
 - Propiedades de las rectas paralelas y perpendiculares.
 - b. Rectas, parábolas, círculos, simetría, transformaciones:
 - Conocer y calcular el área y perímetro de un polígono.
 - Conocer y calcular el área y circunferencia de un círculo.
 - Calcular el volumen de un cubo y cilindro.
 - Conocer y utilizar el teorema de Pitágoras y las propiedades especiales de los isósceles, equiláteros y triángulos rectángulos.
 - c. Congruencia:
 - Experimentar con transformaciones en el plano.
 - Entender la congruencia en términos de movimientos rígidos.
 - d. Semejanza, triángulos rectángulos y trigonometría:
 - Resolver problemas que involucren semejanza.
 - Definir razones trigonométricas y resolver problemas que involucren triángulos rectángulos.
 - Aplicar trigonometría a triángulos generales.
 - f. Geometría analítica en el plano:
 - Asignar coordenadas cartesianas a puntos en el plano.
 - Reconocer y determinar la Ecuación cartesiana de la recta y los Semi-planos que determina.
 - Determinar la distancia entre dos puntos del plano.
 - Reconocer y determinar la Ecuación de la circunferencia y el Círculo que compone.
 - Calcular la intersección de recta y circunferencia.
- 4. Análisis de datos, estadística y probabilidad
 - a. Media, mediana, moda, rango, rango intercuartil, desviación estándar, gráficos y diagramas:
 - b. definir los conceptos de población, muestra y muestra aleatoria.
 - c. Construir tablas de frecuencias relativas y absolutas, absolutas acumuladas y relativas acumuladas a partir de un con junto de datos.

- d. Presentar la información gráficamente a través de histogramas, polígonos de frecuencia, ojivas, etc.
- e. Interpretar tablas y gráficos.
- f. A partir de un conjunto de datos no agrupados, calcular: media, mediana, moda, cuartiles, varianza y desviación estándar.
- g. Definición de probabilidad clásica:
 - Resolver problemas mediante el uso de las propiedades de la probabilidad.
- h. Probabilidad condicional:
 - Entender la independencia y la probabilidad condicional y usarla para interpretar los datos.
- i. Comprender el concepto de variable aleatoria.
- j. Usar las reglas de la probabilidad para calcular la probabilidad de eventos compuestos en un modelo de probabilidad uniforme.
- k. Uso de la probabilidad para la toma de decisiones:
- l. Calcular los valores esperados y usarlos para resolver problemas.
- m. Usar probabilidad para evaluar los resultados de las decisiones.
 - Distribución binomial y distribución normal:
 - Resolver problemas de distribución binomial y distribución normal.
- a. Hallar la correlación e interpretar los resultados.
- b. Calcular la recta de regresión lineal entre dos variables e interpretar los resultados.

1.1. Estándares de desempeño

Los estándares de desempeño se definen como la descripción del grado de desempeño de los examinados en diferentes categorías (Cizek, Bunch y Koons, 2004) y muchas veces son usados para informar sobre el desempeño de grupos de estudiantes y el progreso de los centros educativos o estados en vez de ser utilizados para tomar decisiones acerca de los estudiantes individuales (Linn, 2003).

1.1.1. Los niveles de desempeño

Para la dimensión matemática de nuestra prueba los niveles de desempeño fueron definidos por un panel de expertos conformado por cinco docentes, tres de ellos habían participado en el establecimiento de estándares de contenido,

se eligieron a los que presentaban mayor experiencia en la formulación de pruebas, otros dos docentes se unieron a este grupo y fueron seleccionados por tener antecedentes de participación en elaboración de pruebas a gran escala.

El panel de expertos acordó establecer tres niveles de desempeño, que se nombraron Insuficiente, Suficiente y Avanzado y establecieron las competencias que debe tener un estudiante para alcanzar cada uno de los niveles.

1.2. Selección de los ítems de la dimensión matemática

La creación y adaptación de ítems es una de las tareas más importantes para obtener un instrumento de medición fiable, por eso se decidió apelar a ítems ya testeados. Se eligió tomar como base la dimensión Matemática del SAT (Scholastic Assessment Test) porque esta es una prueba con más de 80 años de trayectoria dirigida a estudiantes de bachillerato con el fin de evaluar su preparación académica para el ingreso a la universidad. Es una prueba estandarizada desarrollada por The College Board.

La prueba en la dimensión Matemáticas quedó finalmente conformada por 42 ítems.

Capítulo 2

Análisis Psicométrico

2.1. Índice de dificultad

2.1.1. Índice de dificultad corregido

Fueron calculados los índices de dificultad (ID) corregidos para cada ítem. Se entiende por índice de dificultad de un ítem la proporción de sujetos que lo aciertan de aquellos que han intentado resolverlo. Es así que a medida que el índice aumenta (entre 0 y 1) significa que ha habido más aciertos por parte de los individuos que intentaron resolverlo (el ítem es más fácil).

En este caso se presentan los índices de dificultad corregidos, dado que, como sugiere Muñiz (1998), al ser ítems de múltiple opción, es necesario corregir los efectos del mero azar lo cual fue realizado con la siguiente fórmula:

$$ID = \frac{A - \frac{E}{K-1}}{N}$$

donde:

A: es el número de sujetos que acierta el ítem;

E: es el número de sujetos que fallan el ítem;

K: es el número de alternativas del ítem;

N: es el número de sujetos que intentan resolver el ítem.

Se recomienda que el índice de dificultad sea mayor a 0,5 y que la dificultad media sea 0,62, 0,67 y 0,75 para ítems de 4, 3 y 2 opciones respectivamente (Crocker y Algina en Abad et al., 2011). Son considerados como mejores ítems los que aportan más varianza al test y los que tienen valores medios para índice de dificultad (Abad et al., 2011).

Cabe aclarar que dado que fue utilizada la fórmula para el cálculo del índice de dificultad corregido, los ítems que no llegan a tener ni siquiera la cantidad esperada de aciertos al azar tienen un índice de dificultad corregido negativo. Estos ítems deben ser analizados para establecer por qué no se llega a la cantidad de aciertos al azar esperada, si es un problema en la formulación de los distractores o del propio ítem. También puede ser que un distractor coincida con un conocimiento apropiado por los sujetos en forma incorrecta. Hay que considerar que como el valor del índice de dificultad depende de la muestra, tener pocos ítems cuyo índice supere los 0,50 puede indicar la baja preparación de la muestra.

En la tabla que sigue pueden observarse los índices de dificultad corregidos para cada uno de los ítems.

	item	indif
[1,]	9	1.21103395
[2,]	10	1.03958333
[3,]	11	1.07996633
[4,]	12	1.23824214
[5,]	13	1.10233918
[6,]	14	1.02753304
[7,]	15	1.17733813
[8,]	16	1.24986673
[9,]	17	0.58112583
[10,]	18	0.25980392
[11,]	19	0.03313253
[12,]	20	0.32339450
[13,]	21	0.73469388
[14,]	22	0.82740586
[15,]	23	0.84274194
[16,]	24	0.64520958
[17,]	25	0.78456221
[18,]	26	0.60668790
[19,]	27	0.42886179
[20,]	28	0.74246231
[21,]	29	0.77132701
[22,]	30	0.42886179
[23,]	31	-0.32812500
[24,]	32	0.78456221
[25,]	34	0.25980392
[26,]	35	0.75490196
[27,]	36	0.80111111
[28,]	37	0.29716981
[29,]	38	-0.06168831
[30,]	39	0.66279070
[31,]	40	0.43548387
[32,]	41	0.35619469
[33,]	42	0.49060150
[34,]	43	0.67937853
[35,]	44	0.64880952
[36,]	45	0.65236686
[37,]	46	0.12777778

[38,]	47	0.00308642
[39,]	48	0.58986928
[40,]	49	0.00308642
[41,]	50	0.14010989
[42,]	51	0.28809524
[43,]	52	0.53873239
[44,]	53	0.37931034
[45,]	54	0.18684211
[46,]	55	-0.25746269
[47,]	56	0.40126050
[48,]	57	0.29716981
[49,]	58	0.37931034
[50,]	59	0.31481481

2.1.2. Índice de discriminación

El índice de discriminación se utiliza para saber si el ítem sirve para discriminar entre los que tienen altos y bajos valores en el constructo. Este indicador de discriminación informan si el ítem mide lo mismo que la prueba global, es decir, cuánto contribuye el ítem a medir lo mismo que la prueba (Abad et al, 2011).

Fueron calculados los índices de discriminación, en este caso las correlaciones biseriales puntuales (bivariadas) de los ítems por separado y de los grupos. Las correlaciones miden cómo están relacionadas las variables o los órdenes de los rangos.

En este caso fueron calculadas las correlaciones biseriales puntuales corregidas. Cuanto más se acerca el valor del índice a 1, mejor discrimina ese ítem entre quienes obtienen puntuaciones altas y bajas en cada dimensión del test. La bibliografía sugiere que los ítems que obtengan un índice menor a 0,20 sean descartados (Abad et al, 2011 y Schmeiser y Welch, 2006).

	item	indis
[1,]	9	0.442240161
[2,]	10	0.562026659
[3,]	11	0.633170078
[4,]	12	0.670667984
[5,]	13	0.311375877

[6,]	14	0.612068861
[7,]	15	0.373083879
[8,]	16	0.679070640
[9,]	17	0.029389328
[10,]	18	-0.047751687
[11,]	19	0.040142139
[12,]	20	-0.078926191
[13,]	21	0.027369276
[14,]	22	-0.015512429
[15,]	23	-0.003019783
[16,]	24	0.066068236
[17,]	25	-0.062920722
[18,]	26	0.014684153
[19,]	27	0.033342136
[20,]	28	0.111384063
[21,]	29	0.028675090
[22,]	30	0.009931953
[23,]	31	0.038244709
[24,]	32	0.012933539
[25,]	34	-0.022032983
[26,]	35	-0.021796433
[27,]	36	-0.114532296
[28,]	37	0.029965112
[29,]	38	0.058214513
[30,]	39	-0.043621980
[31,]	40	0.100016483
[32,]	41	0.084439547
[33,]	42	0.066035545
[34,]	43	-0.118171313
[35,]	44	0.070846145
[36,]	45	-0.008986063
[37,]	46	-0.018400733
[38,]	47	-0.009619404
[39,]	48	0.030544128
[40,]	49	0.036645957
[41,]	50	0.077595630
[42,]	51	-0.063012749
[43,]	52	-0.025802942

[44,]	53	-0.042168759
[45,]	54	0.081302112
[46,]	55	-0.036864158
[47,]	56	0.049374697
[48,]	57	-0.086649378
[49,]	58	0.018574516
[50,]	59	-0.013039088

2.2. Fiabilidad

Se entiende por fiabilidad el grado de estabilidad, precisión o consistencia interna que presenta el test como instrumento de medición de un rasgo determinado.

2.2.1. Estimación de la fiabilidad mediante el coeficiente Alpha

El Alfa de Cronbach es un coeficiente de consistencia interna, basado en el promedio de las correlaciones entre los ítems. Entre las ventajas de esta medida se encuentra la posibilidad de evaluar cuánto mejoraría (o empeoraría) la fiabilidad de la prueba si se excluyera un determinado ítem. Cuanto más cercano a 1 es el coeficiente, mejor la correlación entre los ítems analizados.

Se calculó el coeficiente alfa para todos los ítems exceptuando el ítem en cuestión de forma de poder tomar decisiones con posterioridad para descartar o mantener los ítems en la prueba.

Se obtuvo un Alfa de 0.02 para la prueba de matemática. Este valor se puede considerar muy bueno.

2.3. Validez

2.3.1. Validez de constructo

Análisis Factorial Exploratorio

Para determinar si las dimensiones establecidas a priori son unidimensionales se realizó un análisis factorial para cada dimensión (aquí se presenta para la dimensión matemática). Se utilizó el análisis de componentes principales, que difiere del análisis factorial ya que el primero es un modelo descriptivo de datos y el segundo es un modelo estructural. Generalmente se utiliza el análisis de componentes principales como una forma similar del análisis factorial (Revelle, 2013, p.38)

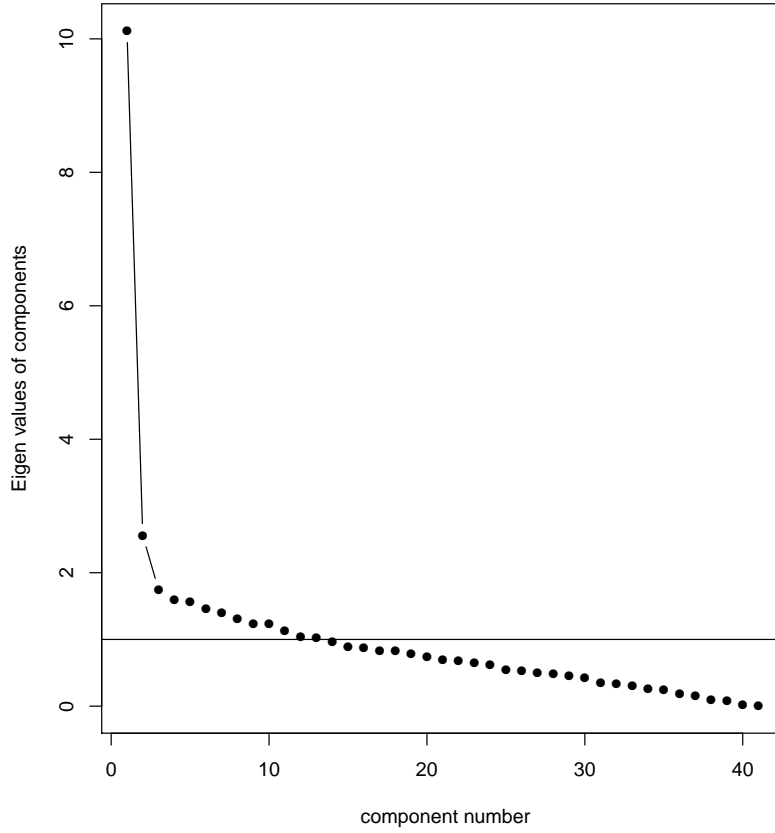
Dimensión Matemática

Análisis de Componentes Principales con rotación varimax

Loadings:

	PC1
[1,]	
[2,]	
[3,]	0.357
[4,]	0.533
[5,]	0.593
[6,]	0.611
[7,]	
[8,]	0.682
[9,]	0.492
[10,]	0.578
[11,]	0.524
[12,]	0.638
[13,]	0.630
[14,]	0.416
[15,]	
[16,]	0.472
[17,]	0.530
[18,]	0.509
[19,]	0.499

Scree plot



[20,] 0.485
[21,] 0.561
[22,] 0.715
[23,] 0.514
[24,] 0.598
[25,] 0.525
[26,] 0.563
[27,] 0.544
[28,] 0.505
[29,] 0.486
[30,] 0.474

[31,]
[32,] 0.460
[33,] 0.427
[34,] 0.448
[35,] 0.709
[36,]
[37,] 0.377
[38,] 0.658
[39,]
[40,]
[41,] 0.554

	PC1
SS loadings	10.121
Proportion Var	0.247

Capítulo 3

Calibración de la prueba y establecimiento del punto de corte

Para calibrar nuestra prueba nos basamos en la Teoría de Respuesta al Ítem (TRI) y se utilizó el modelo logístico de 2 parámetros, donde se modeliza la probabilidad de responder correctamente al ítem i por una persona que tenga habilidad θ_j mediante:

$$P_i(\theta_j) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}} \quad (3.1)$$

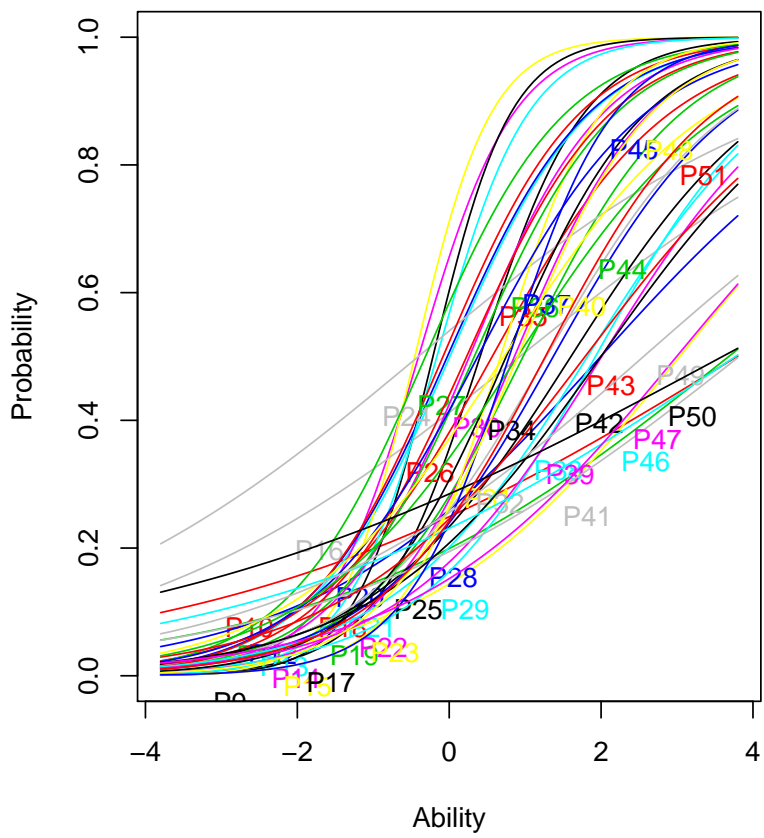
donde el parámetro a_i se denomina discriminación del ítem i y el parámetro b_i dificultad del ítem. Para estimar los parámetros de los ítems y las habilidades de los estudiantes utilizamos el paquete ltm de R.

Con el objetivo de lograr la mayor independencia entre la dificultad empírica y los puntos de corte a determinar se utilizó el método propuesto por García, Abad, Olea y Aguado (2013), ya que los ítems son diseñados, o en nuestro caso, seleccionados en base a los estándares de desempeño establecidos para clasificar a los estudiantes.

Los intervalos de habilidad que se encuentran en cada nivel, se presentan en la siguiente tabla:

Categoría	θ
Insuficiente	<0.565
Suficiente	≥ 0.565 y < 2.23
Avanzado	≥ 2.23

Características del Ítem (2PL)



3.1. Resultados

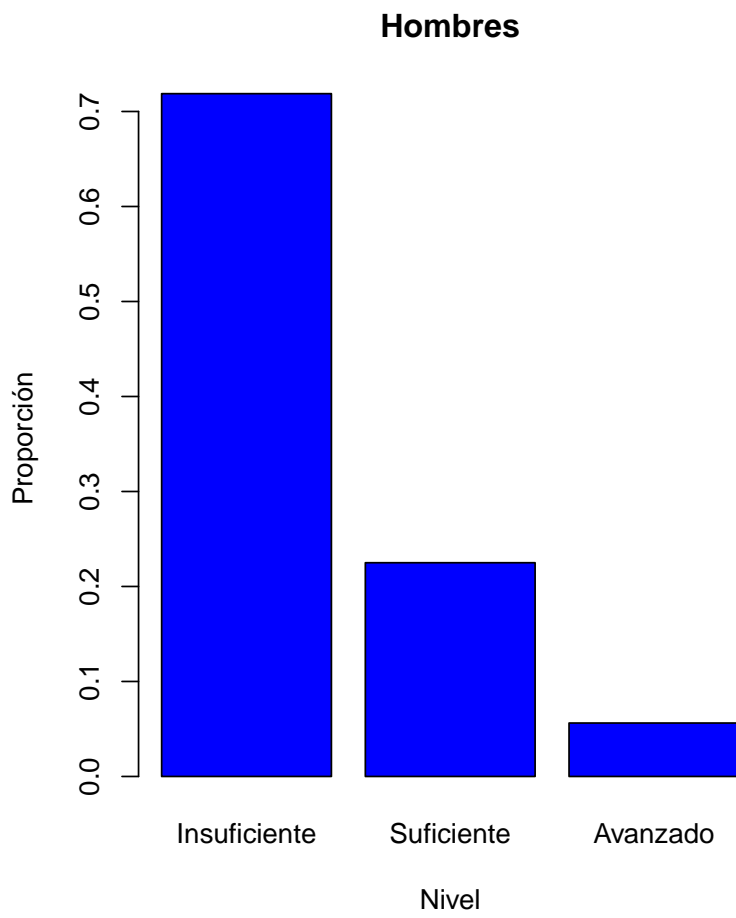
Al clasificar a los estudiantes con estos puntos de corte y se obtuvo que el 77,72 % se ubicó en el nivel insuficiente, el 19,56 % en el suficiente y el 2,72 % en el nivel avanzado.

3.2. Clasificación según sexo

Hombres

Realizaron la prueba 160 estudiantes de sexo masculino. Los porcentajes de estudiantes clasificados en cada nivel son:

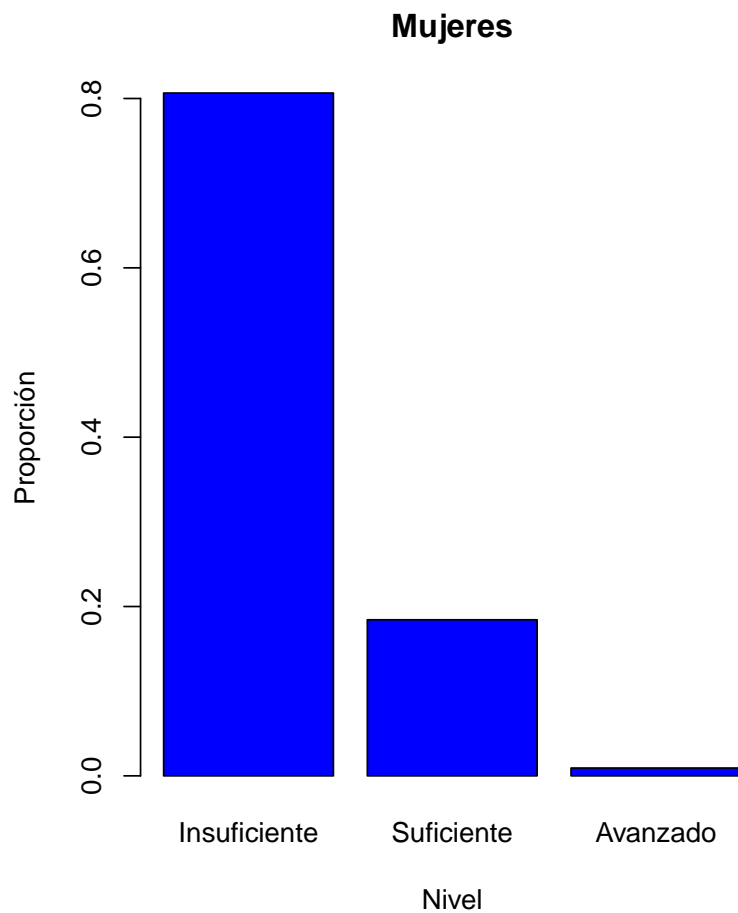
Insuficiente 71.88 Suficiente 22.5 Avanzado 5.62



Mujeres

Realizaron la prueba 217 estudiantes de sexo femenino. Los porcentajes de estudiantes clasificados en cada nivel son:

Insuficiente 80.65 Suficiente 18.43 Avanzado 0.92

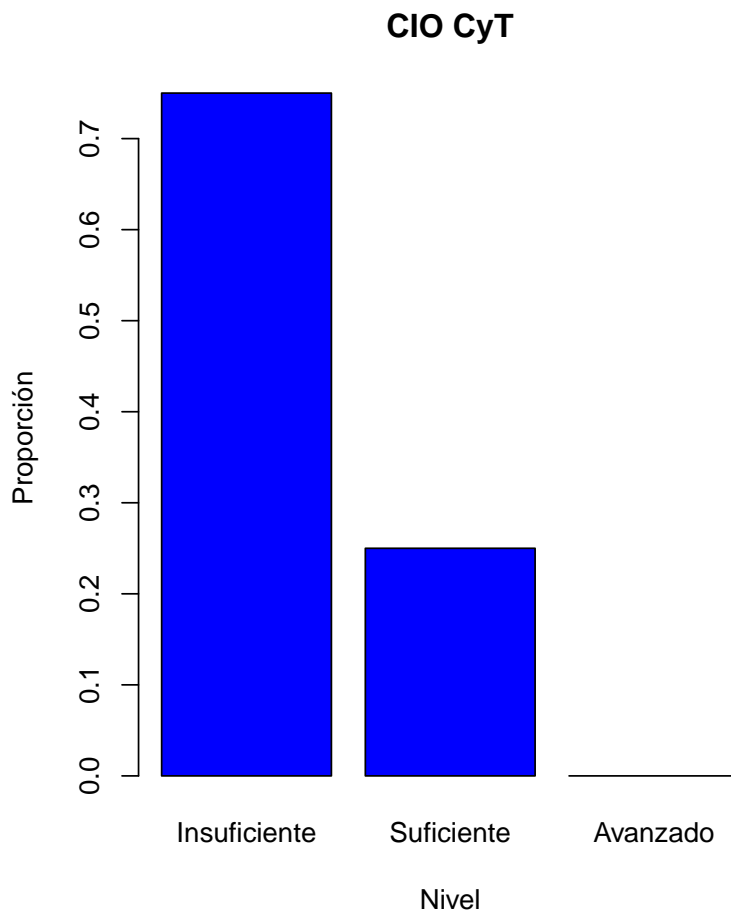


3.3. Clasificación según carrera

Ciclo Inicial Optativo Ciencia y Tecnología

Realizaron la prueba 36 estudiantes del CIO CyT. Los porcentajes de estudiantes clasificados en cada nivel son:

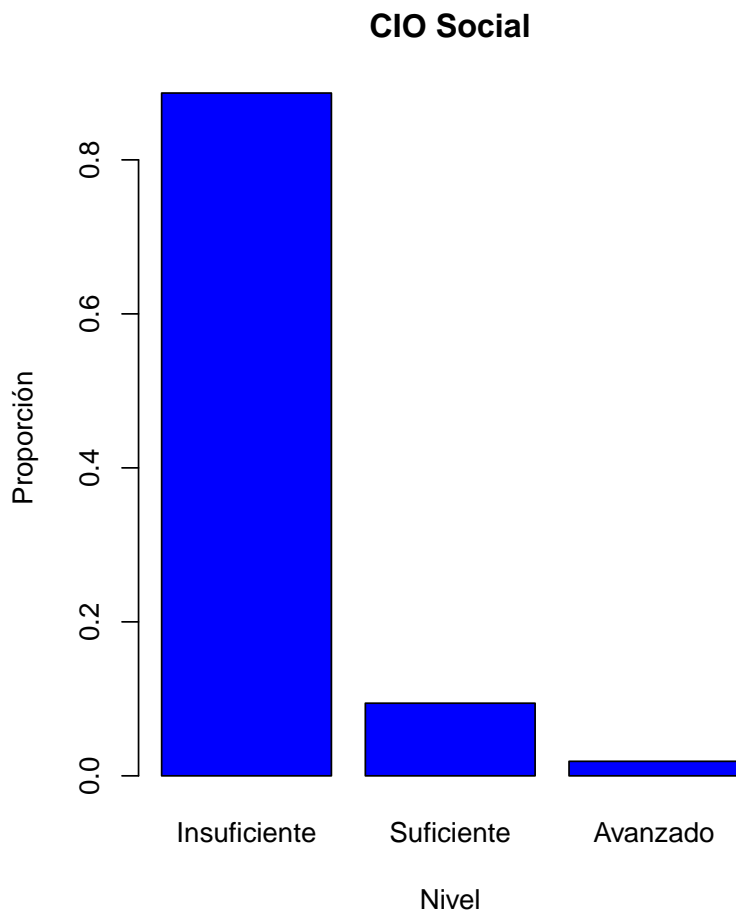
Insuficiente 75 Suficiente 25 Avanzado 0



Ciclo Inicial Optativo Social

Realizaron la prueba 53 estudiantes del CIO Social. Los porcentajes de estudiantes clasificados en cada nivel son:

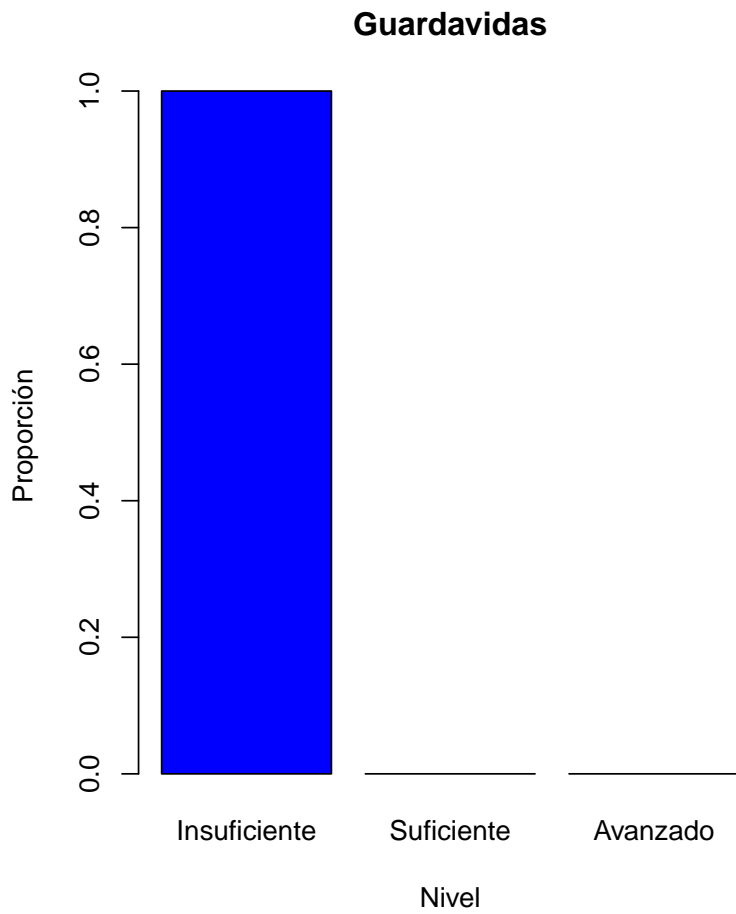
Insuficiente 88.68 Suficiente 9.43 Avanzado 1.89



Curso de Guardavidas

Realizaron la prueba 1 estudiantes del Curso de Guardavidas. Los porcentajes de estudiantes clasificados en cada nivel son:

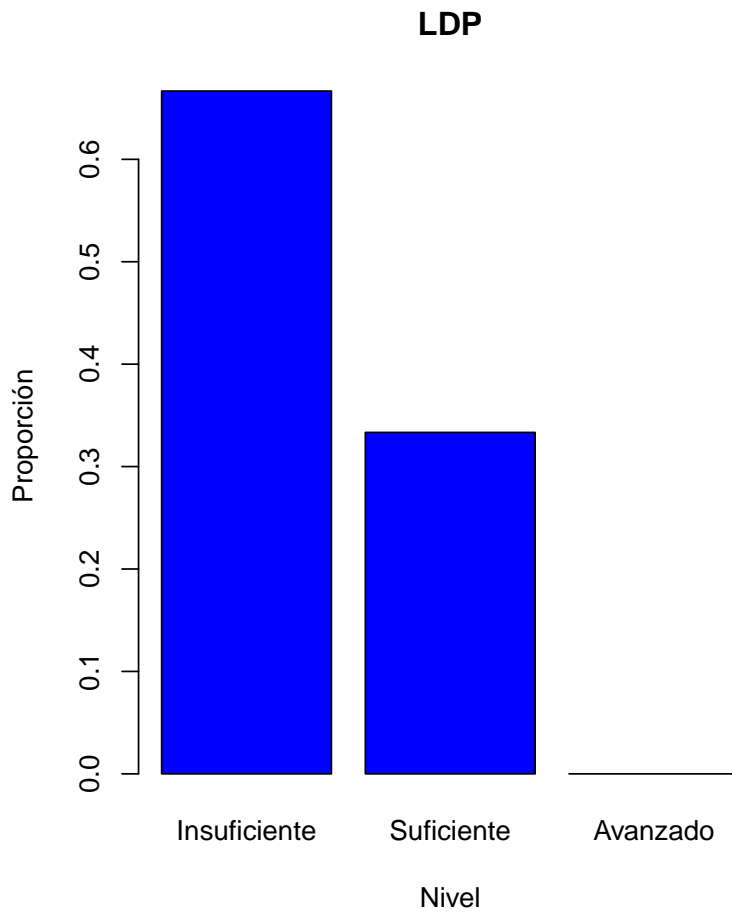
Insuficiente 100 Suficiente 0 Avanzado 0



Licenciatura en Diseño de Paisaje

Realizaron la prueba 21 estudiantes de la LDP. Los porcentajes de estudiantes clasificados en cada nivel son:

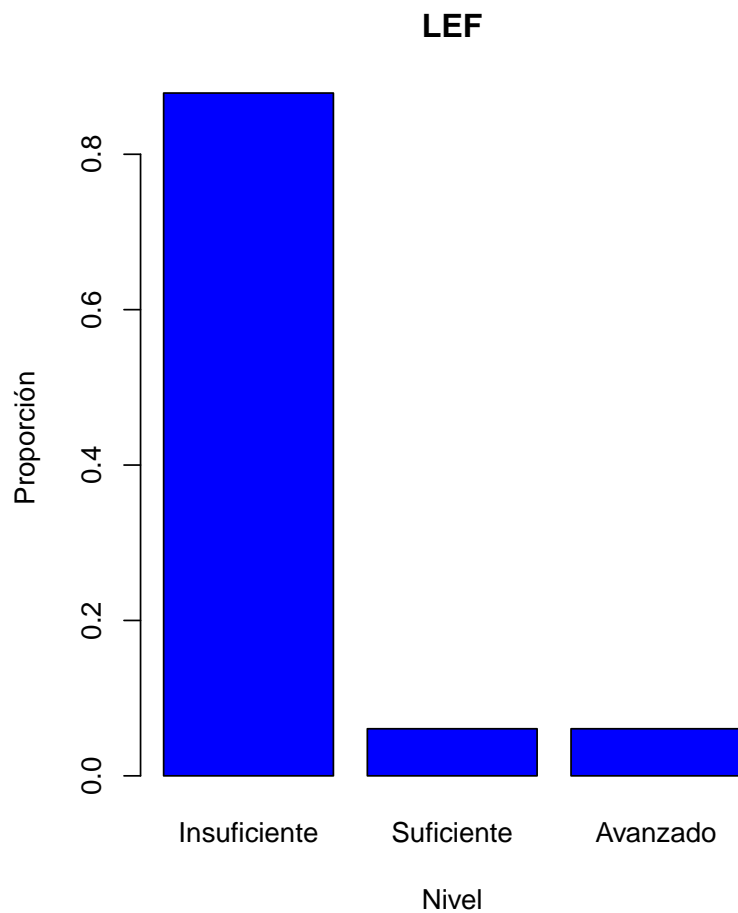
Insuficiente 66.67 Suficiente 33.33 Avanzado 0



Licenciatura en Educación Física

Realizaron la prueba 33 estudiantes de la LEF. Los porcentajes de estudiantes clasificados en cada nivel son:

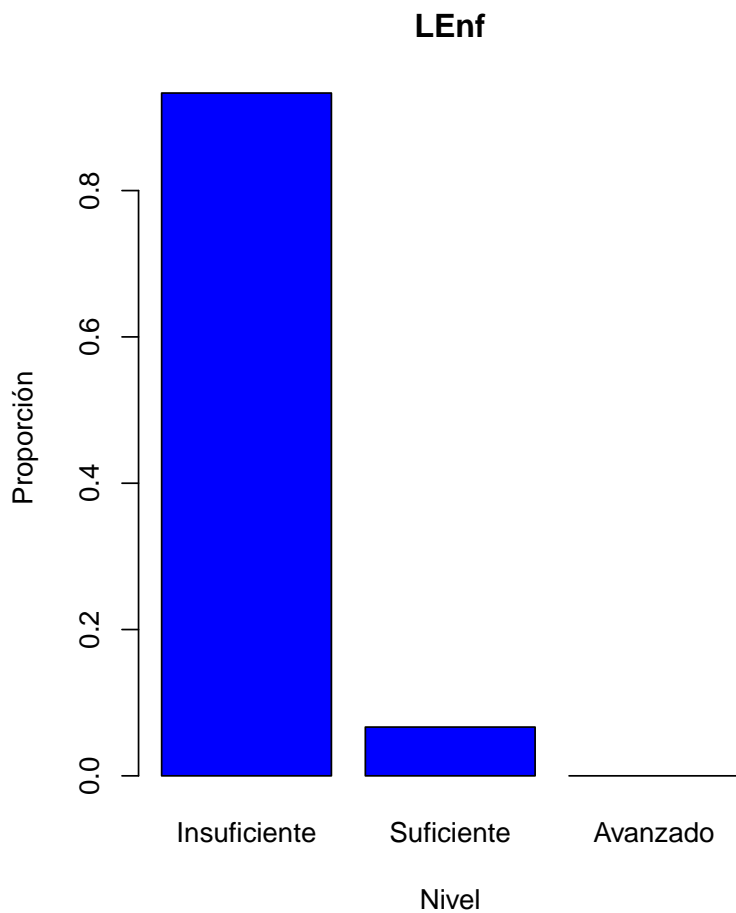
Insuficiente 87.88 Suficiente 6.06 Avanzado 6.06



Licenciatura en Enfermería

Realizaron la prueba 15 estudiantes de la LEnf. Los porcentajes de estudiantes clasificados en cada nivel son:

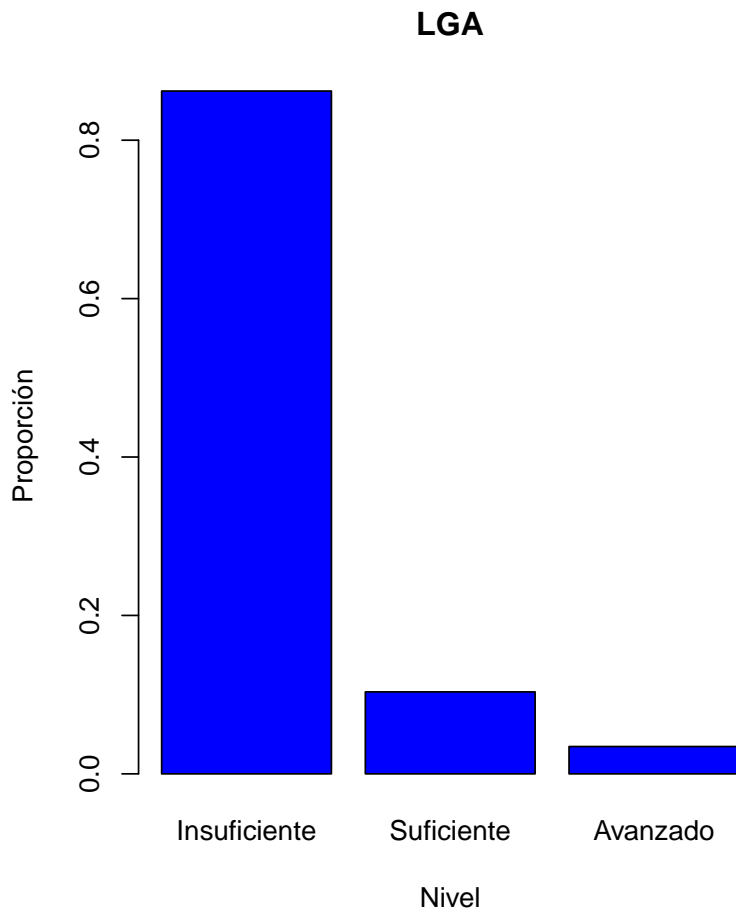
Insuficiente 93.33 Suficiente 6.67 Avanzado 0



Licenciatura en Gestión Ambiental

Realizaron la prueba 29 estudiantes de LGA. Los porcentajes de estudiantes clasificados en cada nivel son:

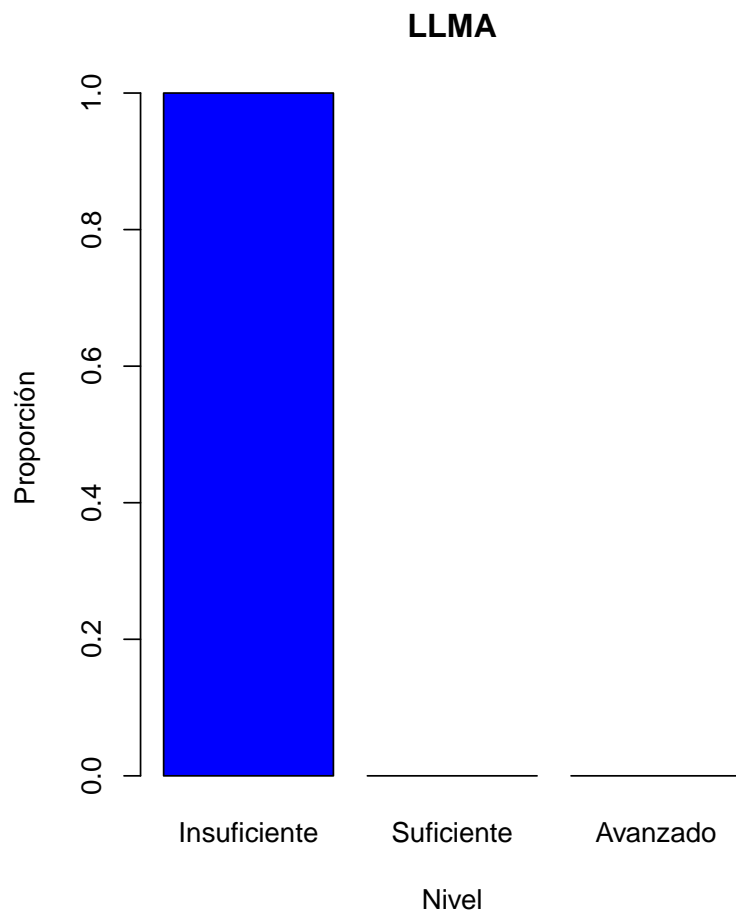
Insuficiente 86.21 Suficiente 10.34 Avanzado 3.45



Licenciatura en Lenguajes y Medios Audiovisuales

Realizaron la prueba 2 estudiantes de la LLMA. Los porcentajes de estudiantes clasificados en cada nivel son:

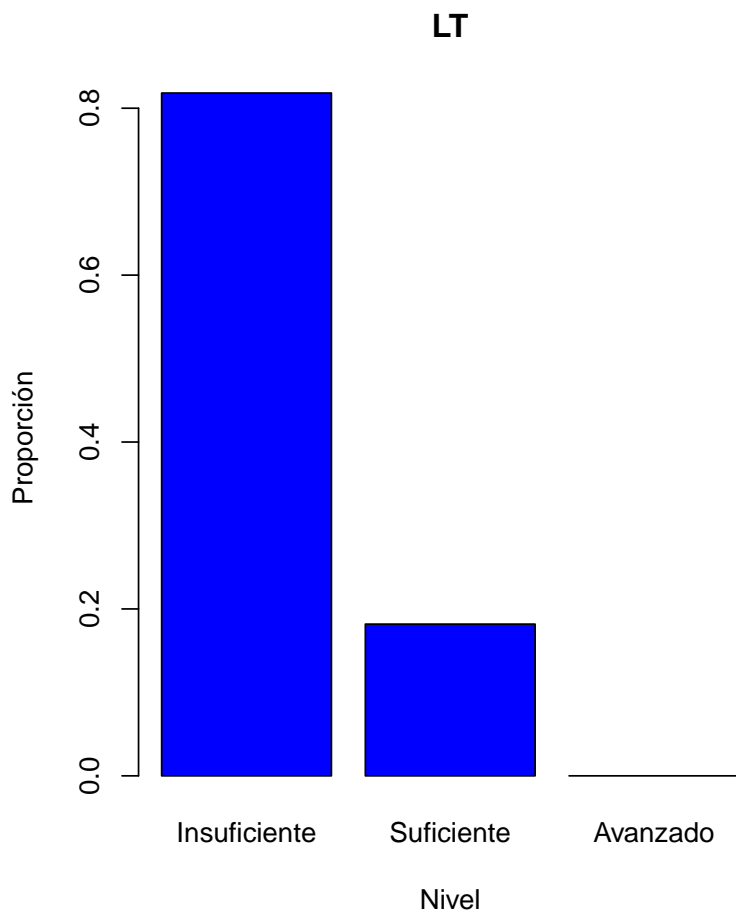
Insuficiente 100 Suficiente 0 Avanzado 0



Licenciatura en Turismo

Realizaron la prueba 44 estudiantes de LT. Los porcentajes de estudiantes clasificados en cada nivel son:

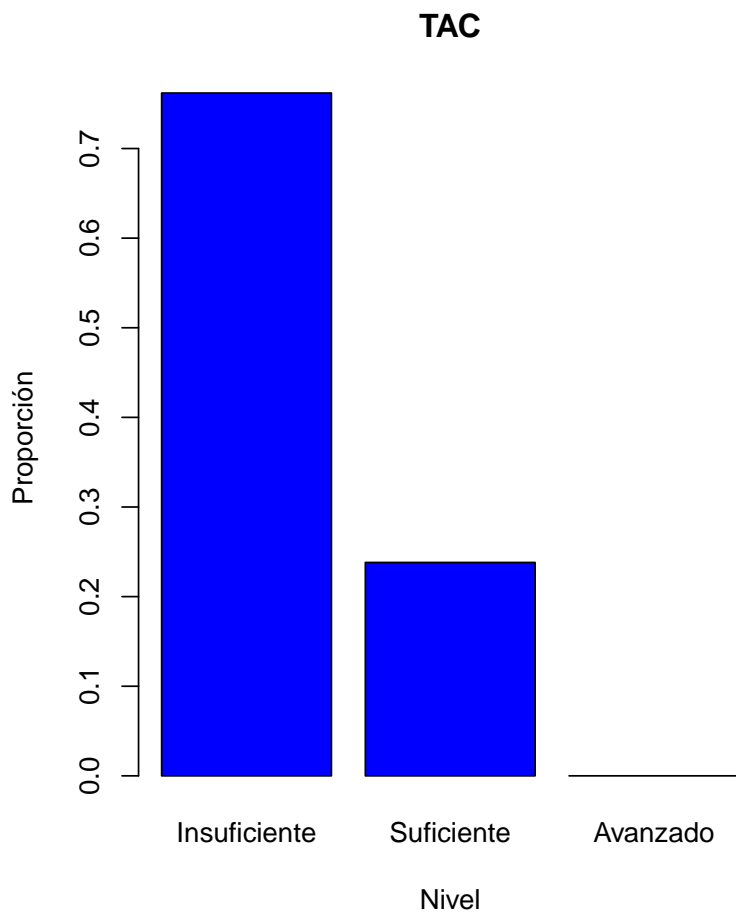
Insuficiente 81.82 Suficiente 18.18 Avanzado 0



Tecnólogo en Administración y Contabilidad

Realizaron la prueba 105 estudiantes del TAC. Los porcentajes de estudiantes clasificados en cada nivel son:

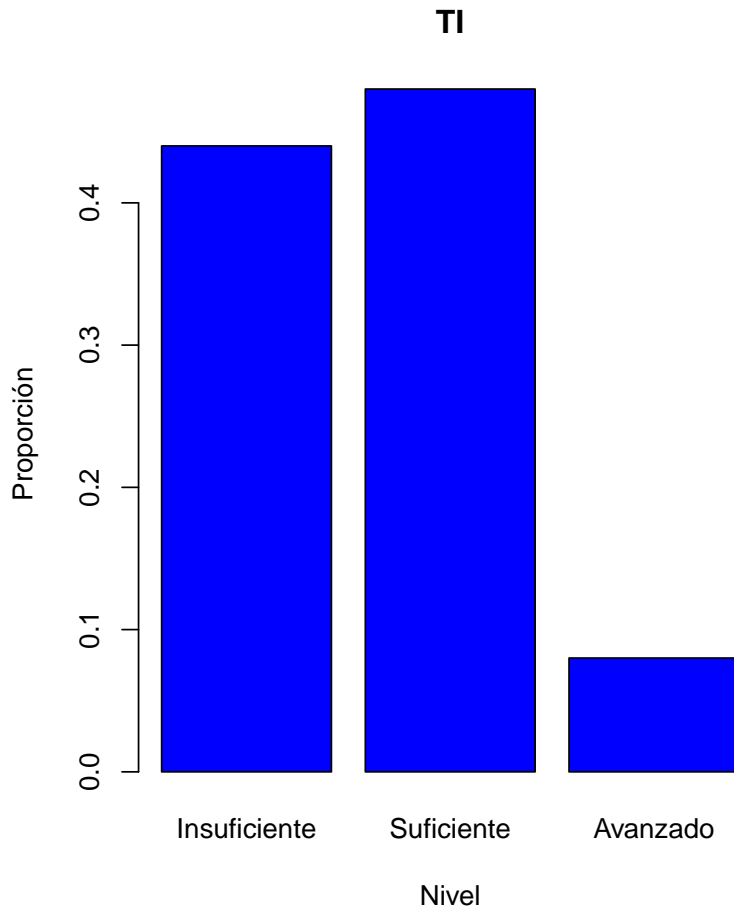
Insuficiente 76.19 Suficiente 23.81 Avanzado 0



Tecnólogo en Informática

Realizaron la prueba 25 estudiantes del TI. Los porcentajes de estudiantes clasificados en cada nivel son:

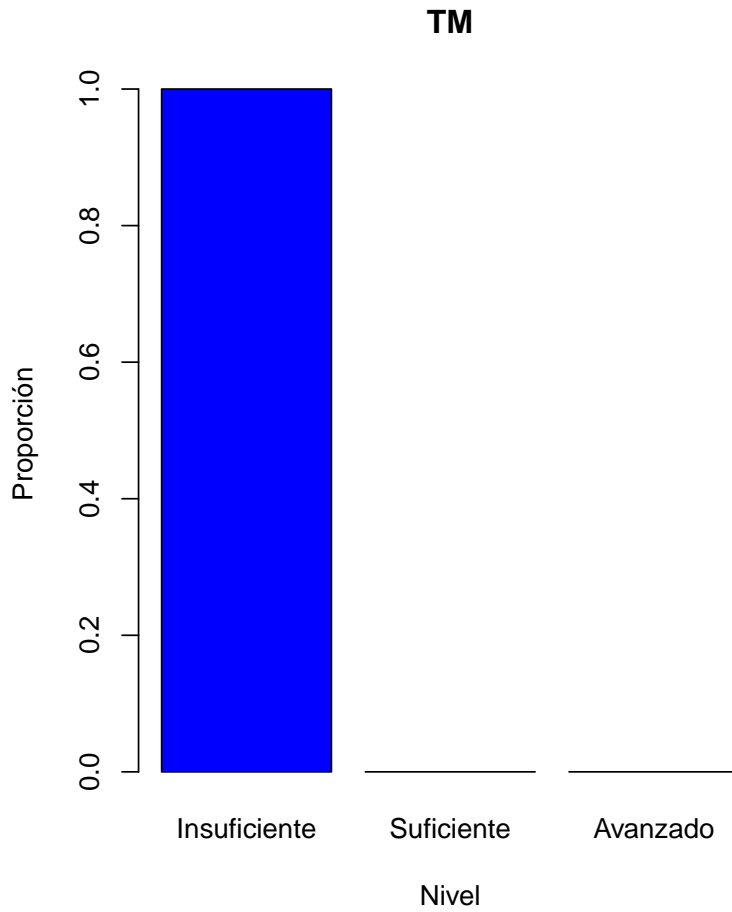
Insuficiente 44 Suficiente 48 Avanzado 8



Tecnólogo en Minería

Realizaron la prueba 1 estudiantes del TM. Los porcentajes de estudiantes clasificados en cada nivel son:

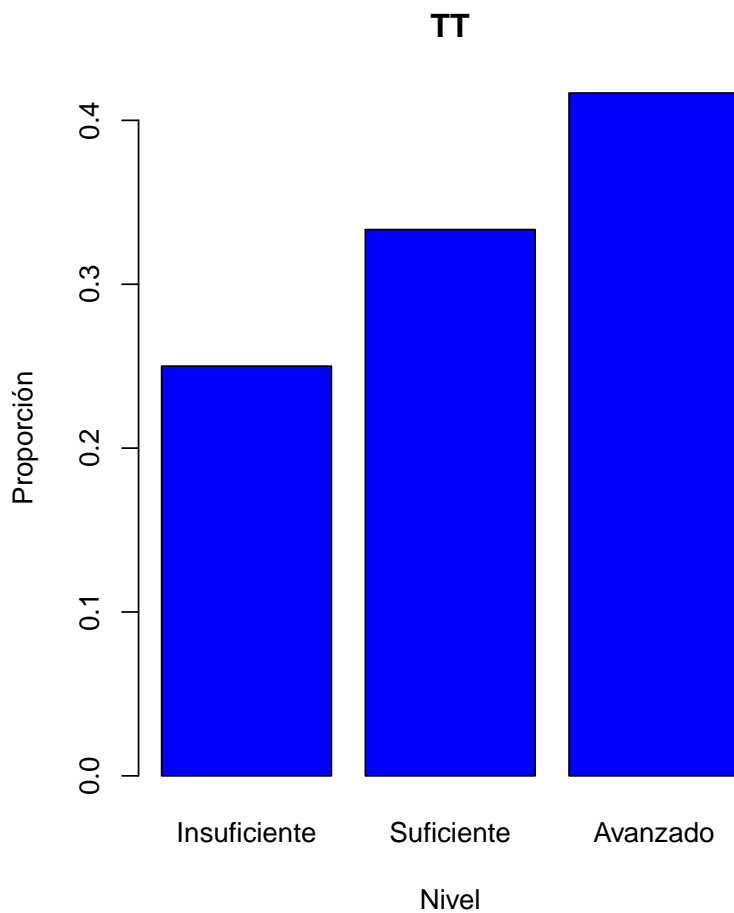
Insuficiente 100 Suficiente 0 Avanzado 0



Tecnólogo en Telecomunicaciones

Realizaron la prueba 12 estudiantes del TT. Los porcentajes de estudiantes clasificados en cada nivel son:

Insuficiente 25 Suficiente 33.33 Avanzado 41.67



3.4. Referencias

Abad, F., Olea, J., Ponsoda, V., García, C. (2011). Medición en Ciencias Sociales y de la Salud. Madrid: Síntesis.

Cizek, G.J. (2001). *Setting performance standards: concepts, methods and perspectives*. Mahwah: NJ. Lawrence Erlbaum Associates.

Cizek, G. J., Bunch, M. B. (2007). *Standard Setting. A guide to establishing and evaluating performance standards on tests*. Thousand Oak, CA: Sage Publications.

CollegeBoard (2012). *Getting Ready for the SAT Subjects Tests*. Disponible en: www.collegeboard.org/subjecttests. Recuperado: 28/01/14.

CollegeBoard (2014). *SAT Practice Questions Math*. Disponible en: <http://sat.collegeboard.org/> Recuperado: 28/01/14.

García-Arancil, A.; Palomares-Montero, D. (2012). Indicadores para la evaluación de las instituciones universitarias: validación a través del método Delphi. *Revista Española de Documentación Científica*, 35, 1, pp.119-144.

García, P.E.; Abad, F.J.; Olea, J. y Aguado, D. (2013). A new IRT-based standard setting method: Application to elath-Listening. *Psicothema*, 25 (2), pp.238-244.

Geist, M.R. (2006). Using the Delphi method to engage stakeholders: a comparison of two studies. *Evaluation and Program Planning*, 33, 147-154.

Hambleton, R.K. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. En G.J. Cizek (Ed.) *Setting performance standards: Concepts, methods, and perspectives*. Mahwah. New York: Lawrence Erlbaum.

Linn, R.L. (1994). The likely impact of performance standards as a function of uses: From rhetoric to sanctions. Paper presented at the National Center for Education Statistics and National Assessment Governing Board Joint conference on Standard Setting for Large Scale Assessments, Washington, DC.

Linn, R. (2003). Performance Standards: Utility for Different Uses of Assessments. *Education Policy Analysis Archives*, 11 (31). Disponible en: <http://epaa.asu.edu/epaa/v11n31/>. Fecha de consulta: 26/11/2013.

Muñiz, J. (2000). *Teoría Clásica de los Tests*. Madrid: Pirámide.

North, B. y Jones, N. (2009). Maintaining standards across languages, contexts and administrations by exploitin teacher judgment and IRT scaling. Disponible en: <http://www.coe.int/t/dg4/linguistic/manuell_en.asp>.

Pérez Juste, R. (2006). *La Evaluación de Programas Educativos*. La Muralla. Madrid.

Prieto, G. y Delgado, A. (1996). Construcción de Ítems. En J. Muñiz (Coord.) (1996). *Psicometría*. Madrid: Universitas.

Schmeiser y Welch (2006). En Abad, F., Olea, J., Ponsoda, V., García, C. (2011). *Medición en Ciencias Sociales y de la Salud*. Madrid: Síntesis.

Universidad de la República (2011). Ordenanza de Estudios de Grado y otros Programas de Formación Terciaria. Resolución No. 4 del CDC del 30/08/2011. Disponible en http://www.cse.edu.uy/sites/www.cse.edu.uy/files/documentos/ORDENANZA_04_DEL_CDC_DEL_30_08_2011.pdf