

Análisis de la Prueba de Selección Universitaria de Matemática (PSUM)



CONSEJO DE RECTORES DE
LAS UNIVERSIDADES CHILENAS



Consejo de Rectores de las Universidades Chilenas, 2019
Alameda 1371, piso cuatro, Santiago de Chile
www.consejoderectores.cl

Equipo de trabajo:

Alejandra Venegas
María Elisa Zenteno
María Elena González

Coordinación y Edición General: Tatiana Diener Cabezas, Periodista

Diseño: Paulina Fuenzalida Vera

Análisis de la Prueba de Selección Universitaria de Matemática (PSUM)

Sistema Único de Admisión
Consejo de Rectores de las Universidades Chilenas
Agosto 2019

Índice

1. ANTECEDENTES	6
Objetivo del Estudio	8
2. ASPECTOS METODOLÓGICOS	9
Datos	10
Unidad de análisis	10
Análisis	11
Características de la prueba de Matemática	13
3. RESULTADOS	18
Grado de Dificultad	19
Confiabilidad	22
Validez Predictiva	23

4.	SÍNTESIS DE RESULTADOS	27
5.	CONCLUSIONES	30
6.	LIMITACIONES DEL ESTUDIO	33
7.	ANEXO	36
8.	GLOSARIO	46
9.	BIBLIOGRAFÍA	47

1 Antecedentes

Las Pruebas de Selección Universitaria (PSU) se aplican desde el año 2003 y, por mandato del Consejo de Rectores de las Universidades Chilenas (CRUCH), el organismo encargado de desarrollar las pruebas es el Departamento de Evaluación, Medición y Registro Educativo (DEMRE) de la Universidad de Chile.

En el año 2011, el CRUCH en conjunto con el Ministerio de Educación solicitaron una evaluación internacional de las PSU, la que se conoció durante 2013 a través del *Informe Pearson*. Desde la publicación de dicho informe hasta la fecha, el sistema de admisión universitario ha sido foco de debate.

Más recientemente, a fines de 2017, el CRUCH convocó a un Comité de Expertos¹ para evaluar el sistema de admisión y los instrumentos de selección. Durante 2018, el Comité centró su trabajo en analizar siete propuestas que planteó el DEMRE para introducir ajustes a las PSU y al sistema.

Una de las propuestas dice relación con diseñar y aplicar dos pruebas de Matemática, en vez de una que se utiliza en la actualidad. Al respecto, el informe final del Comité de Expertos recomienda separar la prueba en dos: Matemática Básica y otra de Matemática Avanzada. Esta recomendación se basó en un estudio realizado por DEMRE 2017, estudio que solo consideró a los estudiantes de la promoción.

A juicio del comité de expertos, la prueba básica consideraría los contenidos que se enseñan en el primer ciclo de enseñanza media (primero y segundo medio) y, también, los contenidos de 7° y 8° año. En tanto, la prueba avanzada incluiría los contenidos del segundo ciclo de enseñanza media (tercero y cuarto medio). Desde el Consejo de Rectores se estima que la separación de esta prueba requiere de un cuidadoso análisis y validación. Por lo mismo, el presente estudio tiene como finalidad contribuir con consideraciones técnicas en la discusión de los cambios a la PSU de Matemática (PSUM).

¹ El Comité de Expertas del Sistema Único de Admisión (SUA) discutió y evaluó la propuesta del DEMRE (2017). Los miembros activos del comité fueron Salomé Martínez, Alejandra Mizala, Verónica Santelices y Rebecca Zwick.

Objetivo del estudio

El presente estudio tiene como finalidad contribuir con el análisis metodológico y estadístico a la discusión de los cambios a la prueba de selección universitaria de matemáticas (PSUM).

2 Aspectos metodológicos

A continuación, se presentan cuatro secciones (Datos, Unidad de Análisis, Análisis y Características de la PSU de Matemática) que dan cuenta de la información y metodología usada para la realización de este informe.

Datos

Se consideraron datos correspondientes a estudiantes que rinden la PSU de Matemática (PSUM) el 2016 y participaron del Proceso de Admisión 2017. En esta cohorte se tiene información de caracterización de los ítems e información del estudiante. Además, para los análisis de predictibilidad se usaron datos correspondientes al rendimiento del primer año académico de los estudiantes en todas las universidades del Sistema Único de Admisión (SUA) del Consejo de Rectores (Anexo 1 y Anexo 2). Para ello, se utilizó el Promedio Ponderado Acumulado (PPA)², respecto a los créditos de cada curso del estudiante.

Unidad de análisis

Para evaluar los cambios propuestos a la PSU de Matemática se consideraron tres parámetros psicométricos (*grado de dificultad (GD)*, *confiabilidad* y *validez predictiva*), que se calcularon a partir de características de los estudiantes y/o agrupamiento de ítems, como se especifica a continuación:

♦ **Características del estudiante según su establecimiento de egreso:**

Dependencia: es una característica propia del establecimiento, que puede ser Particular Pagado, Particular Subvencionado o Municipal.

Rama: es una característica propia del establecimiento, que puede ser Científico Humanista (CH) o Técnico Profesional (TP).

² Las universidades consideran los cursos aprobados y reprobados para reportar el PPA.

♦ **Agrupamiento de ítems, según:**

Ciclos de Enseñanza: el primer ciclo corresponde a ítems del currículo de primero y segundo medio; el segundo ciclo corresponde a ítems del currículo de tercero y cuarto medio.

Eje temático: contenido o disciplina que evalúa ítem. En la PSUM corresponden a *Número, Álgebra, Geometría y Datos & Azar*.

Habilidad: puede ser Aplicación, Comprensión, o Análisis-Síntesis-Evaluación.

Competencia: puede ser Comprender o Resolver.

En tanto, para la *validez predictiva* la unidad de análisis fue a nivel de carrera. En este caso, se definió que la cantidad de estudiantes debía ser superior o igual a 30, de manera de tener variabilidad en la unidad de análisis al momento de hacer las correlaciones.

Análisis

Para realizar el análisis de los datos, se usaron los siguientes estadísticos: *grado de dificultad (GD): confiabilidad, y validez predictiva*. Se detallan sus principales características a continuación:

Grado de dificultad (GD): tal como su nombre lo indica, este parámetro corresponde al *grado de dificultad* del ítem. La definición que se usó para este informe fue la de Teoría Clásica, es decir, *grado de dificultad* como la proporción de respuestas correctas por ítem. En tanto, para obtener la dificultad del instrumento completo, se promediaron todos los GD de los ítems. Este indicador varía en un rango de 0 a 1, donde valores cercanos a 0, dan cuenta de una mayor dificultad del ítem.

Confiabilidad: este parámetro permite conocer la consistencia interna que tiene el test.

La confiabilidad de la PSUM se evaluó mediante el coeficiente Alfa de Cronbach que mide la fiabilidad del instrumento.

Dado que en algunos agrupamientos (*eje temático, habilidad, competencia, ciclos de enseñanza*) se cuenta con poca cantidad de ítems, fue necesario realizar un ajuste mediante la fórmula de Spearman-Brown (ver ecuación 1.1) para proyectar la confiabilidad a un determinado número de ítems, debido a que se están comparando pruebas (o subpruebas en este caso), que poseen diferente número de ítems.

$$\frac{n\alpha}{1 + (n - 1)\alpha} \quad (1.1)$$

n: Cantidad de veces que se aumenta la longitud del test.

α: Coeficiente de fiabilidad del test.

Validez Predictiva: este indicador permite establecer una relación entre el test que se está validando y una medida de desempeño que se obtenga luego de la administración del test. En este caso se evalúa mediante el desempeño de los estudiantes en sus carreras respectivas al término del primer año universitario (Promedio Ponderado Acumulado, PPA). Este análisis provee de evidencia relevante para la validez de un instrumento, que como la PSU de Matemática, tiene como finalidad selección y admisión a las universidades del SUA.

Para *validez predictiva* se realizaron dos tipos de correlaciones. La primera es la que ha usado el Sistema Único de Admisión (SUA) en sus estudios de validez y, la segunda, es una versión simple, tal como se explica a continuación:

- ♦ **Correlación SUA:** para obtener un promedio general combinando de todas las carreras, se utilizará la transformación Z de Fischer (ecuación 1.2) que, usando los promedios ponderados, permite, mediante la transformación inversa tangente (ecuación 1.3), obtener la correlación promedio.

$$z = \frac{1}{2} \ln \left(\frac{1+p}{1-p} \right) \quad (1.2)$$

$$z = \tanh(\bar{Z}) \quad (1.3)$$

- ♦ **Correlación versión simple:** esta metodología de cálculo para la validez predictiva utiliza, en primer lugar, correlaciones simples a nivel carrera. Luego, para obtener una correlación general, se promedia ponderadamente según el número de alumnos.

$$\frac{\sum_i n_i r_i^*}{\sum_i n_i} \quad (1.4)$$

n_i : Cantidad de alumnos en la carrera i

r_i^* : Correlación corregida simple de la carrera i

Corrección por atenuación

Debido a la diferencia en confiabilidad presentada en los test separados por ciclo se utilizará la corrección por atenuación (ec.1.5) que permite obtener una estimación de la correlación entre la PSUM y el criterio (notas universitarias), que sea independiente de la confiabilidad del agrupamiento de ítems de la PSUM que se use en la comparación.

$$\frac{r_{xy}}{\sqrt{\alpha_{xx}\alpha_{yy}}} \quad (1.5)$$

r_{xy} : Correlación observada entre x e y

α_{xx} : Confiabilidad en el instrumento x

α_{yy} : Confiabilidad en el instrumento y

Características de la prueba de Matemática

La PSU de Matemática que se aplica para el proceso de admisión, tiene 4 formas diferentes (se identifican por los números 111, 112, 113 y 114). Cada una de las formas tiene 80 ítems, de los cuales 75 se consideran en el cálculo del puntaje final y cinco son de pilotaje.

Como se indicó en la sección de *Unidad de Análisis*, interesa observar los agrupamientos de ítems por ciclo de enseñanza. De este modo, en la Tabla 1 se muestra la distribución de ítems

para los dos ciclos de enseñanza (1° y 2° medio, y 3° y 4° medio) en cada una de las 4 formas de la prueba. Tal como se puede apreciar, el 73% de los ítems presentes en la PSUM para la Admisión 2017, corresponden a ítems del primer ciclo.

Tabla N° 1: Cantidad de Ítems en los dos ciclos de enseñanza

Forma PSUM	1° y 2° medio	3° y 4° medio	Total ítems
111	55	20	75
112	55	20	75
113	55	20	75
114	55	20	75

Para los análisis de la prueba de Matemática en esta sección se consideraron todos los estudiantes que participaron de la Admisión 2017 y que tenían información por rama y dependencia de egreso. En total, la muestra fue de 256.262 estudiantes (independiente de la forma de la PSUM que rindieron).

En los Gráficos 1, 2 y 3, se puede apreciar una distribución asimétrica en el comportamiento de las respuestas correctas de la PSUM. Esto ocurre tanto para la prueba actual como para las versiones separadas por ciclo de enseñanza.

Para complementar la información de los gráficos, la Tabla 2 muestra la *mediana* y el *promedio* de la proporción de respuestas correctas de la PSUM completa y de las versiones separadas por ciclo de enseñanza. Ambos estadísticos son mayores para los ítems del primer ciclo de enseñanza, en relación con los del segundo ciclo, lo que puede indicar que se trata de ítems de menor dificultad. Respecto al segundo ciclo de enseñanza, el *promedio* y *mediana* proporcional disminuye notoriamente, lo que muestra una dificultad mayor para los ítems de este grupo.

Gráfico N° 1: Proporción de estudiantes con respuestas correctas en la PSUM completa (75 ítems)

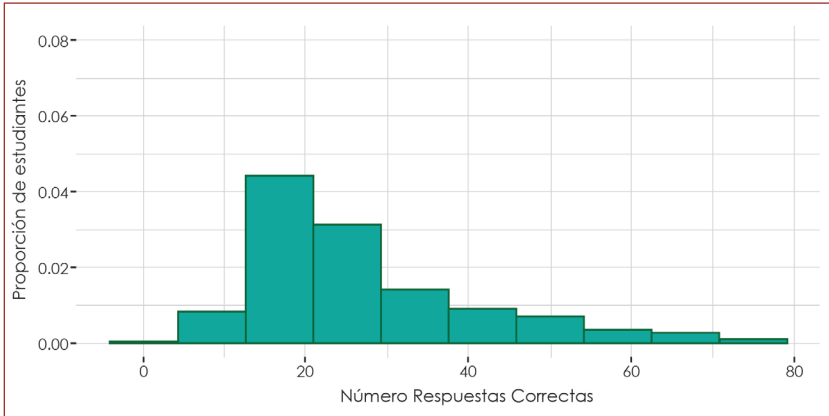


Gráfico N° 2: Proporción de estudiantes con respuestas correctas en la PSUM de primer ciclo de enseñanza (55 ítems)

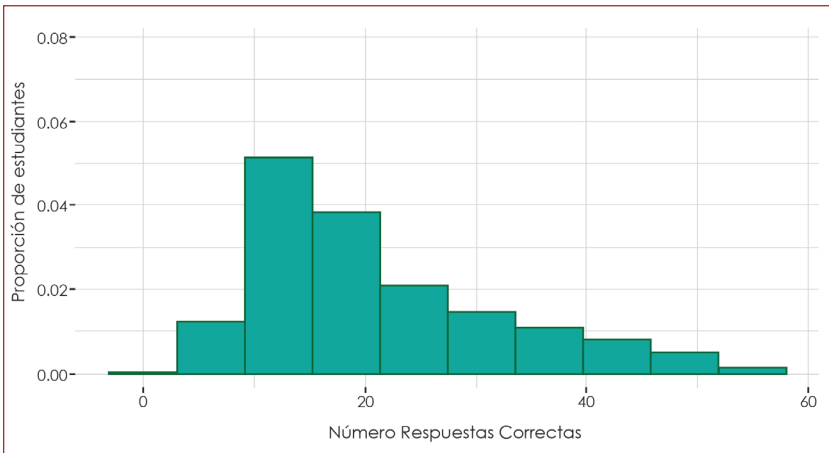


Gráfico N° 3: Proporción de estudiantes con respuestas correctas en la PSUM de segundo ciclo de enseñanza (20 ítems)

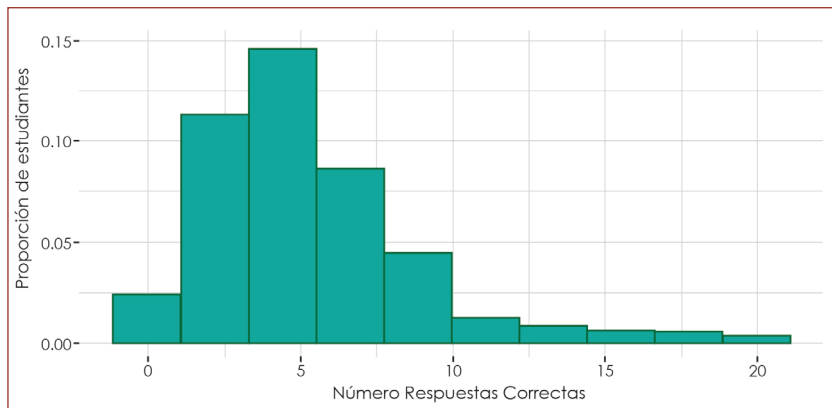


Tabla N° 2: Promedio y Mediana de respuestas correctas de la PSUM completa y las versiones separadas por ciclo de enseñanza

Tipo de Prueba	Promedio	Mediana	Promedio proporcional a la cantidad de ítems	Mediana proporcional a la cantidad de ítems
PSUM (75 ítems)	27	22	0,36	0,29
Primer ciclo de enseñanza (55 ítems)	21	18	0,38	0,32
Segundo ciclo de enseñanza (20 ítems)	5	5	0,25	0,25

Los Gráficos 4 y 5 muestran la proporción de estudiantes con respuestas correctas, según rama (CH y TP) y dependencia (PP, PS y Municipal).

En el caso de los estudiantes TP, se observa un promedio y una mediana de las respuestas correctas con valores más bajos, en comparación a los estudiantes de CH. En la Tabla 3 se aprecia esta diferencia, que es significativa (se realizó un test de comparación de medias, donde se obtuvo un p-valor inferior a 0,05).

Gráfico N° 4: Proporción de estudiantes con respuestas correctas de la PSUM completa según rama del establecimiento de egreso (75 ítems)

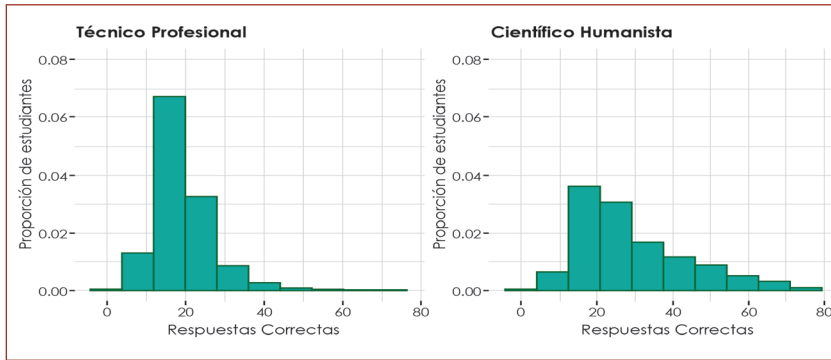


Gráfico N° 5: Proporción de estudiantes con respuestas correctas de la PSUM completa según dependencia del establecimiento de egreso (75 ítems)

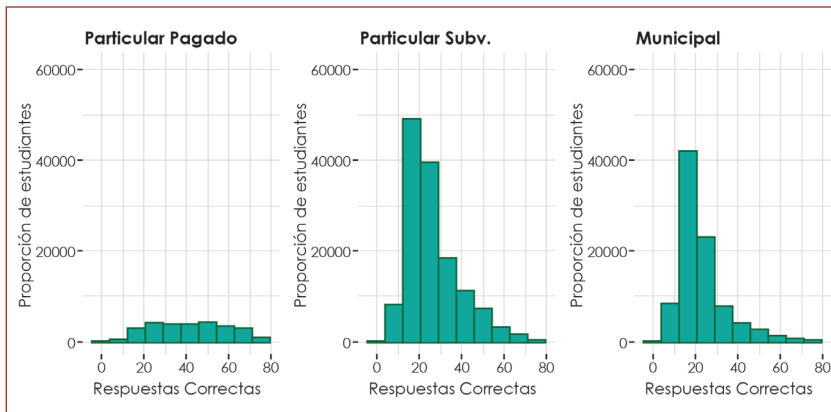


Tabla N° 3: Promedio y Mediana de respuestas correctas de la PSUM completa según rama y dependencia del establecimiento

	Rama		Dependencia		
	TP N=71.062	CH N=185.200	Particular Pagado N=26.903	Particular Subvencionado N=139.109	Municipal N=90.250
Promedio	20	29	42	26	23
Mediana	18	25	42	23	19

3 Resultados

En esta sección se presentarán los cálculos del *grado de dificultad* y *confiabilidad* por rama y dependencia de establecimiento de egreso. Además, se estima la validez predictiva mediante correlaciones.

Los cálculos de *dificultad* y *confiabilidad*, que se muestran en las Tablas 4, 5, 6 y 7, corresponden a solo una de las formas de la PSUM, la forma 111³. Las tablas con los resultados de las tres formas restantes de la PSUM (112, 113 y 114) se encuentran en el Anexo N°3, N°4, N°5, N°6, N°7, N°8, N°9, N°10 y N°11. Cabe decir que los resultados entre formas son similares.

En tanto, la estimación de las correlaciones para *validez predictiva* se hace a nivel general para todos los estudiantes, sin considerar las formas de la PSUM.

Grado de dificultad

En la Tabla 4 se presentan los grados de dificultad promedio para la PSUM, tanto a nivel general como según los agrupamientos de ítems definidos (*ciclo de enseñanza, eje, habilidad y competencia*), según rama y dependencia del establecimiento educacional de egreso.

Globalmente, se constata que los ítems de la PSU de Matemática son de alta dificultad (su porcentaje de respuesta correcta se sitúa claramente por debajo del 50%). En este contexto general, se observa que la dificultad es mayor para los egresados de establecimientos Técnico Profesional (TP), en comparación con los de establecimientos Científico-Humanistas (CH).

La PSUM también es más difícil para quienes egresan de establecimientos municipales, en comparación con que estudiantes de establecimientos particulares subvencionados y, especialmente, en comparación con los de establecimientos particulares pagados (estos últimos son el único grupo donde el porcentaje de respuestas correctas supera el 50%).

³ Dado que para cada una de las formas de la PSUM se presentan dos tablas extensas, se optó por mostrar los resultados de solo una de las formas en esta sección.

Es importante advertir que en los valores de *grados de dificultad* pueden incidir factores que no son posibles de observar mediante este indicador (GD). Por ejemplo, al comparar ítems de distintos *ejes temáticos*, no necesariamente se puede concluir que un eje es más difícil que otro, pues en uno de los ítems pueden influir aspectos como las habilidades cognitivas requeridas para entender y responder, o la oportunidad que ha existido para aprender los contenidos, entre otras explicaciones posibles. Por ello, se debe tener precaución al realizar comparaciones e interpretar la información (GD).

Tabla N° 4: Grado de dificultad para la PSUM forma 111

Ítems		Todos los estudiantes	Rama		Dependencia		
			TP ⁴	CH	Particular Pagado	Particular Subvencionado	Municipal
		64.125	17.701	46.424	6.717	34.820	22.588
PSUM completa	75 ítems	0,35	0,25	0,38	0,55	0,34	0,29
Ciclo	Primer Ciclo (55 ítems)	0,37	0,27	0,41	0,59	0,37	0,32
	Segundo Ciclo (20 ítems)	0,28	0,21	0,3	0,45	0,27	0,24
Eje	Números (16 ítems)	0,38	0,27	0,42	0,61	0,38	0,32
	Álgebra (19 ítems)	0,37	0,26	0,41	0,60	0,36	0,3
	Geometría (20 ítems)	0,35	0,27	0,38	0,54	0,34	0,3
	Datos y Azar (20 ítems)	0,31	0,24	0,33	0,49	0,30	0,26
Habilidad	Aplicación (38 ítems)	0,39	0,29	0,43	0,60	0,38	0,33
	Comprensión (16 ítems)	0,34	0,24	0,38	0,57	0,34	0,28
	Análisis, Síntesis y Evaluación (21 ítems)	0,27	0,21	0,3	0,45	0,34	0,24
Competencia	Comprender (58 ítems)	0,34	0,25	0,37	0,55	0,33	0,28
	Resolver (17 ítems)	0,38	0,3	0,41	0,57	0,38	0,33

4 Cabe mencionar que los grados de dificultad para la rama TP de recién egresados, son similares a los obtenidos para toda la muestra TP. Los estudiantes TP recién egresados son el 72% (12.805) del total de estudiantes TP.

En la Tabla 5 se presenta la distribución del *grado de dificultad* de los ítems según rangos de dificultad para la forma 111 de la prueba (ver en Anexo 6, 7 y 8 los rangos de dificultades para las otras formas de la PSUM). Tal como se puede apreciar, más de dos tercios de los ítems poseen una dificultad inferior a 0,42. Dado que ningún ítem supera el 62% de respuestas correctas, uno de los aspectos que debiera revisarse es el proceso de ensamblaje de la prueba. Esto, porque no se observan ítems con baja dificultad.

Tabla N° 5: Grado de dificultad por rango para la PSUM forma 111

Rango de Dificultad para los ítems	Cantidad de ítems	Porcentaje acumulado
0,12-0,22	15	20%
0,22-0,32	11	35%
0,32-0,42	27	71%
0,42-0,52	15	91%
0,52-0,62	7	100%
Total	75	-

En la Tabla 6, donde se compara el rango de dificultad de los ítems según el ciclo de enseñanza, se observa que en el primer y segundo ciclo prevalecen ítems de alta dificultad, aunque en el primer ciclo se pueden encontrar algunos ítems de menor dificultad.

Tabla N° 6: Grado de dificultad por rango para la PSUM forma 111, según ciclo de enseñanza

Rango de Dificultad	Primer Ciclo de enseñanza	Segundo Ciclo de enseñanza	Porcentaje acumulado 1er ciclo	Porcentaje acumulado 2° ciclo
0,12-0,22	9	6	16%	30%
0,22-0,32	5	6	25%	60%
0,32-0,42	21	6	66%	90%
0,42-0,52	13	2	87%	100%
0,52-0,62	7	0	100%	-
Total	55	20	-	-

Confiabilidad

El cálculo de la confiabilidad se basó en el coeficiente alfa de Cronbach, a partir de los ítems presentes en la forma 111. En los Anexos 9, 10 y 11 se pueden encontrar las tablas con la misma información calculada para las formas restantes.

En términos generales, el conjunto de 75 ítems muestra una alta confiabilidad (0,93), tal como ha sido reportado en los informes técnicos de las PSU elaborados por DEMRE. En cuanto a los establecimientos educacionales por dependencia, en los tres casos (PP, PS y Municipal), la confiabilidad se muestra próxima o superior a 0,9.

Sin embargo, al comparar los resultados según rama escolar, se constata que en los estudiantes que egresan de la enseñanza Técnico Profesional, la confiabilidad es claramente menor (0,73), si se compara con la enseñanza Científico Humanista, cuya confiabilidad es de 0,93. Esta diferencia podría deberse a la menor variabilidad de desempeños que existen entre los estudiantes TP, ya que en la PSUM sus resultados se concentran en puntajes más bajos, en comparación a los CH. Es importante mencionar que en la teoría de la medición se afirma que la confiabilidad disminuye a medida que se reduce la variabilidad del grupo en que se aplica una prueba.

Con respecto a la comparación de los resultados entre ciclos de enseñanza, se observa una diferencia entre el primer y segundo ciclo, donde la confiabilidad es más alta en el primer caso. No obstante, dado que la PSUM posee un número significativamente más elevado de ítems asociados al primer ciclo -en comparación al segundo ciclo-, esta diferencia podría reflejar el hecho que la confiabilidad aumenta cuando se incrementa el número de ítems.

Para evaluar cómo se comportaría la PSUM con ítems del segundo ciclo, si el número de ítems fuera equivalente al del primer ciclo, se definió aplicar la fórmula de Spearman-Brown para simular la confiabilidad. De esta manera, dado que son 20 ítems, se simuló que se triplicaran, es decir, que fueran 60, lo que se asemeja a

los 55 de primer ciclo. El resultado de este análisis muestra que la confiabilidad de los ítems del segundo ciclo proyectada a 60 ítems es muy semejante a la del primer ciclo.

Tabla N° 7: Confiabilidad promedio para forma 111

Ítems		Todos	Rama		Dependencia		
			TP	CH	Part. Pagado	Part. Subv.	Municipal
		N= 64.125	N=17.701	N=46.424	N=6.717	N=34.820	N=22.588
Actual	Todos los ítems (75 ítems)	0,93	0,74	0,93	0,96	0,91	0,89
Ciclo	Primer Ciclo (55 ítems)	0,91	0,73	0,92	0,94	0,9	0,87
	Segundo Ciclo (20 ítems)	0,72	0,21	0,76	0,86	0,64	0,57
	Segundo Ciclo (Simulando 60 ítems con la fórmula de Spearman-Brown)	0,89	0,44	0,9	0,95	0,84	0,8

Validez Predictiva

Como se indicó en la sección Análisis, los cálculos se realizarán con dos metodologías distintas, *Correlación SUA* y *Correlación versión simple*. Esta última se encuentra en el Anexo 12 y 13 (tablas).

Para las correlaciones de validez predictiva se consideraron carreras con al menos 30 alumnos y con información del rendimiento del primer año académico. En total, la muestra fue de 61.418 estudiantes pertenecientes a 35 universidades.

► Correlación SUA

En la Tabla 8 se presentan diferentes correlaciones, calculadas entre el Promedio Ponderado Acumulado (PPA) de primer año y distintos agrupamientos de ítems de la PSUM. En términos generales, la mayor correlación se logra con la PSUM.

En el caso de la comparación entre los dos ciclos de enseñanza, el número de ítems es desigual, lo que, como se explicó previamente, se asocia a la menor confiabilidad de los ítems del segundo ciclo. Por este motivo se decidió aplicar la corrección por atenuación en los dos ciclos. Los resultados, que se muestran en la tercera columna de la Tabla 8, indican que, en el caso del puntaje de los ítems del segundo ciclo, que inicialmente aparecen con menor capacidad predictiva que los ítems del primer ciclo, presentan predictibilidad equivalente a los del primer ciclo cuando se aplica la corrección por atenuación.

Tabla N° 8: Correlaciones del PPA con agrupamiento de ítems en PSUM

Ítems	PPA	PPA (aplicando corrección por atenuación)
PSUM (75 ítems)	0,28	-
PSUM Primer Ciclo (55 ítems)	0,26	0,27
PSUM Segundo ciclo (20 ítems)	0,22	0,26
PSUM EJE-Algebra (19 ítems)	0,24	-
PSUM EJE-Geometría (20 ítems)	0,21	-
PSUM EJE-Números (16 ítems)	0,22	-
PSUM EJE-Datos y azar (20 ítems)	0,18	-

La correlación promedio entre la PSUM y el desempeño universitario por sub áreas OCDE (Organización para la Cooperación y Desarrollo Económico), se presenta en la Tabla 9. Allí se puede observar que las carreras de Ingeniería, Tecnología y Ciencias son las que muestran índices más altos de validez predictiva. Cabe destacar que, cuando se aplica la corrección por atenuación (Tabla 10) a los ítems del primer ciclo y a los del segundo ciclo, la correlación promedio con el desempeño universitario es muy semejante entre los dos ciclos, con la sola excepción de un par de áreas donde, en términos generales, la PSUM es menos predictiva, como en Derecho y Servicio Social.

Tabla N° 9: Correlación del PPA con PSUM

Área	PPA con PSUM (75 ítems)
Ingeniería y profesiones afines	0,43
Ciencias Físicas	0,43
Informática	0,39
Protección del Medio Ambiente	0,38
Industria y Producción	0,37
Veterinaria	0,34
Agricultura, Silvicultura y Pesca	0,34
Servicios de Seguridad	0,33
Servicios Personales	0,31
Ciencias de la Vida	0,31
Matemáticas y Estadísticas	0,29
Arquitectura y Construcción	0,27
Servicios de Transporte	0,26
Salud	0,25
Enseñanza Comercial y Administración	0,24
Humanidades	0,23
Formación de Personal Docente	0,21
Ciencias Sociales y del Comportamiento	0,18
Derecho	0,16
Periodismo e Información	0,14
Servicios Sociales	0,13
Artes	0,13

Tabla N° 10: Correlación del PPA con PSUM

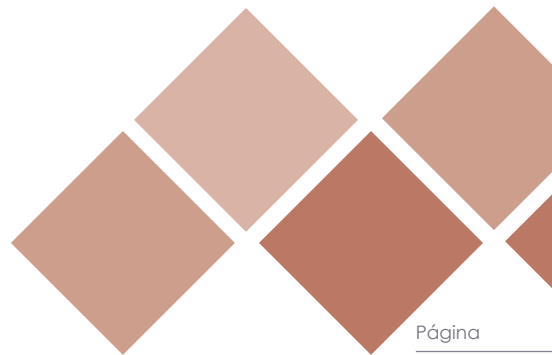
Área	PPA con primer ciclo de enseñanza (aplicando Corrección por atenuación)	PPA con segundo ciclo de enseñanza (aplicando Corrección por atenuación)
Ingeniería y profesiones afines	0,43	0,44
Ciencias Físicas	0,43	0,42
Informática	0,38	0,40
Protección del Medio Ambiente	0,38	0,18
Industria y Producción	0,37	0,37
Veterinaria	0,37	0,26
Agricultura, Silvicultura y Pesca	0,34	0,29
Servicios de Seguridad	0,34	0,20
Servicios Personales	0,32	0,28
Ciencias de la Vida	0,30	0,31
Matemáticas y Estadísticas	0,29	0,28
Arquitectura y Construcción	0,27	0,25
Servicios de Transporte	0,27	0,16
Salud	0,25	0,34
Enseñanza Comercial y Administración	0,23	0,22
Humanidades	0,21	0,26
Formación de Personal Docente	0,21	0,15
Ciencias Sociales y del Comportamiento	0,18	0,16
Derecho	0,16	0,01
Periodismo e Información	0,15	0,19
Servicios Sociales	0,15	0,08
Artes	0,12	0,25

4 Síntesis de resultados

Síntesis de resultados

- a) Con respecto a la dificultad de la PSU de Matemática, se constata que es una prueba globalmente difícil, puesto que, tanto su puntaje general como todos sus ejes, muestran un porcentaje de respuesta correcta inferior al 50% (de hecho, más del 60% de los ítems tienen una dificultad inferior a 0,42). En este contexto, y tal como se podría esperar a partir de la información acerca de las brechas existente según rama educativa y dependencia del establecimiento educacional, se observa que la prueba es aún más difícil para egresados de establecimientos Técnico Profesionales y Municipales.
- b) Complementariamente, se observa que la dificultad es algo mayor en los ítems asociados al segundo ciclo de enseñanza media. Sin embargo, dado que este ciclo presenta un número menor de ítems, no es posible extraer una conclusión más clara sobre este aspecto.
- c) En términos de confiabilidad, tanto la prueba actual en su conjunto, como si se consideran las dos versiones -según ciclos de enseñanza-, muestran altos grados de confiabilidad. Sin embargo, debido a que la versión PSUM basada en el segundo ciclo contiene menos preguntas que la prueba del primer ciclo, fue necesario aplicar la fórmula de Spearman-Brown para hacer comparables los resultados. Con este ajuste, se obtiene que la confiabilidad de ambas partes, si tuvieran un número equivalente de ítems, sería parecida.

- d)** Respecto a la validez predictiva (correlación de la prueba con el desempeño en la universidad al término del primer año), la PSUM presenta un valor global que es razonable. Cuando se comparan las dos versiones de la prueba, según ciclo de enseñanza, se obtienen estimaciones de validez predictiva muy parecidas, tanto a nivel global como por carrera, si se aplica el ajuste por atenuación (para corregir la desigual confiabilidad, asociada al diferente número de ítems que poseen). En términos generales, se observa que la prueba de Matemática posee una buena capacidad predictiva del rendimiento de primer año.



5 Conclusión

Este estudio presenta información relevante con respecto a la PSU de Matemática, con especial foco en una posible diferenciación de una prueba basada en el primer ciclo de la enseñanza media (1° y 2° medio), versus una basada en el segundo ciclo (3° y 4° medio).

En términos generales, se evidencia que la prueba actual posee *confiabilidad* y *capacidad predictiva* aceptable. Además, se puede observar que algo similar ocurre con los ítems, separados por primer y segundo ciclo de enseñanza.

Entre los aspectos abordados en este informe (*grado de dificultad, confiabilidad y validez predictiva*), se puede sostener que el más relevante es el de la dificultad de la prueba, que se observa desajustado para una proporción importante de los(as) examinados(as). Los ítems presentan un alto nivel de *dificultad* para la mayor parte de quienes rinden la PSUM.

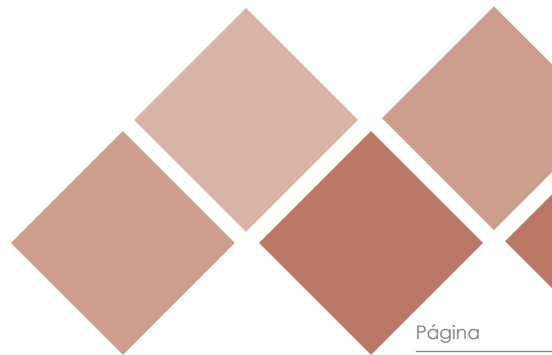
Para poder discriminar correctamente a los(as) estudiantes con puntajes inferiores, es importante identificar ítems con menor dificultad que permitan distinguir a estudiantes con mayor conocimiento en Matemática para este grupo (puntajes inferiores).

Dado que se trata de una prueba selectiva, debe ser capaz de diferenciar entre examinados(as) con un manejo más avanzado de la matemática, pero ello no se debe excluir la posibilidad de acercar la dificultad de la prueba a la distribución de los(as) examinados(as).

De esta manera, se sugiere considerar la opción de que la prueba mantenga un porcentaje de preguntas de alta dificultad (que hagan posible distinguir a postulantes a carreras con mayor base matemática), al tiempo que se agregan ítems de menor dificultad (no necesariamente fáciles), para poder medir con mayor precisión en la región de puntajes medios de la prueba.

Un avance en esa dirección no implica necesariamente que las brechas de puntaje asociadas a dependencia educativa o modalidad de la enseñanza media se reduzcan, pero al menos se contará con evidencia más pertinente para hacer comparaciones en una región de puntajes (medios y medios-bajos) donde en la actualidad la PSU de Matemática posee menor capacidad informativa.

En definitiva, es necesario revisar el proceso de diseño y ensamblaje de la PSUM y, por otro lado, avanzar en la realización de estudios que muestren con claridad lo que evalúa el contenido de cada ítem.



6 Limitaciones del estudio

Limitaciones del estudio

- a) En primer lugar, cabe destacar que para este estudio se efectuó un análisis a partir de la información disponible de una prueba que no fue concebida para comparar los dos ciclos de enseñanza (primer ciclo corresponde a ítems del currículo de primero y segundo medio; el segundo ciclo corresponde a ítems del currículo de tercero y cuarto medio).

De hecho, como se ha advertido, incluye un número muy desigual de preguntas del primer y segundo ciclo, lo que ha obligado a emplear procedimientos que permitieran hacer comparables los resultados, según procedimientos establecidos en la teoría de la medición. Si se avanzara en la dirección de crear dos pruebas, se requeriría contar con estudios empíricos acerca del funcionamiento de pruebas construidas para ser aplicadas como instrumentos diferenciados.

- b) En segundo lugar, considerando que la propuesta de contar con dos pruebas y una de ellas con contenidos hasta segundo ciclo de enseñanza media de Matemática -que incluya ítems de 7° y 8° año de enseñanza básica⁵-, se requiere contar con información del comportamiento de estos ítems, lo que supone un pilotaje de los mismos.

⁵ Esta propuesta fue planteada en el informe final del Comité de Expertas del Sistema Único de Admisión del Consejo de Rectores. Ver: <https://www.consejo-derectores.cl/informes>.

- c) Además en este estudio no se cuenta con información de los ítems a nivel de pilotaje, es decir, el comportamiento de éstos en términos de dificultad, discriminación y otros parámetros. Tampoco se tiene información respecto al proceso de construcción de cada ítem ni al ensamblaje de la prueba.

7 Anexos

Anexo 1: Universidades del CRUCH

Universidad
Universidad de Chile
Pontificia Universidad Católica de Chile
Universidad de Concepción
Pontificia Universidad Católica de Valparaíso
Universidad Técnica Federico Santa María
Universidad de Santiago de Chile
Universidad Austral de Chile
Universidad Católica del Norte
Universidad de Valparaíso
Universidad de Antofagasta
Universidad de La Serena
Universidad del Bío-Bío
Universidad de La Frontera
Universidad de Magallanes
Universidad de Talca
Universidad de Atacama
Universidad de Tarapacá
Universidad Arturo Prat
Universidad Metropolitana de Ciencias de la Educación
Universidad de Playa Ancha de Ciencias de la Educación
Universidad Tecnológica Metropolitana
Universidad de Los Lagos
Universidad Católica del Maule
Universidad Católica de la Santísima Concepción
Universidad Católica de Temuco
Universidad de Aysén ⁶
Universidad de O'Higgins
Universidad Diego Portales
Universidad Alberto Hurtado

⁶ Cabe puntualizar que para los cálculos de *validez predictiva* se consideraron carreras con al menos 30 estudiantes con información de la PSUM y rendimiento académico (notas de primer año de carrera). Debido a que la Universidad de Aysén no contaba con dicha información, no fue posible incluirla en el presente estudio.

Anexo 2: Universidades privadas adscritas al Sistema Único de Admisión (SUA)⁷

Universidad
Universidad Mayor
Universidad Finis Terrae
Universidad Andrés Bello
Universidad Adolfo Ibáñez
Universidad de Los Andes
Universidad del Desarrollo
Universidad Católica Silva Henríquez
Universidad Autónoma de Chile
Universidad San Sebastián
Universidad Central
Universidad Academia Humanismo Cristiano
Universidad de O'Higgins

⁷ Es importante señalar que el estudio se desarrolló a partir de información del Proceso de Admisión 2017 (PSU 2016) y el rendimiento académico del primer año universitario (2017). Dado que hay universidades que se incorporaron posterior a estas fechas, no pudieron ser incluidas en el análisis. Es el caso de las universidades Autónoma de Chile, San Sebastián, Central, Academia Humanismo Cristiano y de O'Higgins.

Anexo 3: Grado de dificultad para PSUM forma 112

Ítems		Todos los estudiantes	Rama		Dependencia		
			TP	CH	Part. Pagado	Part. Subv.	Municipal
		64.135	17.812	46.323	6.763	34.819	22.553
PSUM completa	75 ítems	0,36	0,27	0,40	0,58	0,36	0,31
Ciclo	Primer Ciclo (55 ítems)	0,40	0,29	0,44	0,62	0,39	0,34
	Segundo Ciclo (20 ítems)	0,26	0,20	0,29	0,45	0,25	0,23
Eje	Números (16 ítems)	0,40	0,30	0,44	0,63	0,40	0,34
	Álgebra (19 ítems)	0,38	0,27	0,42	0,61	0,38	0,32
	Geometría (20 ítems)	0,36	0,28	0,39	0,56	0,35	0,31
	Datos y Azar (20 ítems)	0,32	0,24	0,35	0,52	0,31	0,27
Habilidad	Aplicación (38 ítems)	0,38	0,28	0,42	0,61	0,37	0,32
	Comprensión (16 ítems)	0,40	0,29	0,45	0,64	0,40	0,34
	Análisis, Síntesis y Evaluación (21 ítems)	0,30	0,23	0,32	0,47	0,29	0,27
Competencia	Comprender (59 ítems)	0,35	0,26	0,39	0,56	0,34	0,30
	Resolver (16 ítems)	0,40	0,31	0,44	0,62	0,40	0,34

Anexo 4: Grado de dificultad para PSUM forma 113

Ítems		Todos los estudiantes	Rama		Dependencia		
			TP	CH	Part. Pagado	Part. Subv.	Municipal
		64.000	17.745	46.255	6.759	34.617	22.624
PSUM completa	75 ítems	0,35	0,26	0,38	0,55	0,34	0,3
Ciclo	Primer Ciclo (55 ítems)	0,37	0,27	0,41	0,59	0,37	0,32
	Segundo Ciclo (20 ítems)	0,28	0,21	0,3	0,45	0,27	0,24
Eje	Números (16 ítems)	0,39	0,27	0,43	0,61	0,38	0,33
	Álgebra (19 ítems)	0,37	0,26	0,41	0,59	0,36	0,31
	Geometría (20 ítems)	0,35	0,27	0,38	0,54	0,34	0,3
	Datos y Azar (20 ítems)	0,31	0,24	0,33	0,49	0,3	0,26
Habilidad	Aplicación (38 ítems)	0,39	0,29	0,43	0,6	0,38	0,33
	Comprensión (16 ítems)	0,34	0,25	0,38	0,56	0,34	0,29
	Análisis, Síntesis y Evaluación (21 ítems)	0,28	0,21	0,31	0,45	0,27	0,24
Competencia	Comprender (59 ítems)	0,34	0,25	0,37	0,54	0,33	0,29
	Resolver (16 ítems)	0,38	0,3	0,41	0,57	0,38	0,33

Anexo 5: Grado de dificultad para PSUM forma 114

Ítems		Todos los estudiantes	Rama		Dependencia		
			TP	CH	Part. Pagado	Part. Subv.	Municipal
		64.002	17.804	46.198	6.664	34.853	22.485
PSUM completa	75 ítems	0,36	0,27	0,39	0,56	0,35	0,3
Ciclo	Primer Ciclo (55 ítems)	0,39	0,29	0,43	0,61	0,39	0,33
	Segundo Ciclo (20 ítems)	0,26	0,2	0,29	0,44	0,25	0,22
Eje	Números (16 ítems)	0,4	0,3	0,44	0,62	0,4	0,34
	Álgebra (19 ítems)	0,38	0,26	0,42	0,6	0,37	0,32
	Geometría (20 ítems)	0,35	0,27	0,38	0,54	0,35	0,3
	Datos y Azar (20 ítems)	0,31	0,24	0,34	0,51	0,31	0,27
Habilidad	Aplicación (38 ítems)	0,38	0,28	0,41	0,6	0,37	0,32
	Comprensión (16 ítems)	0,4	0,29	0,44	0,63	0,4	0,33
	Análisis, Síntesis y Evaluación (21 ítems)	0,29	0,22	0,32	0,46	0,29	0,25
Competencia	Comprender (59 ítems)	0,35	0,25	0,38	0,55	0,34	0,29
	Resolver (16 ítems)	0,4	0,31	0,44	0,61	0,4	0,34

Anexo 6: Grado de dificultad por rango para la PSUM forma 112

Rango de Dificultad de los ítems	Cantidad de ítems	Porcentaje acumulado
0,12 - 0,22	12	16%
0,22 - 0,32	19	41%
0,32 - 0,42	16	63%
0,42 - 0,52	20	89%
0,52 - 0,62	8	100%
Total	75	-

Anexo 7: Grado de dificultad por rango para la PSUM forma 113

Rango de Dificultad de los ítems	Cantidad de ítems	Porcentaje acumulado
0,12 - 0,22	15	20%
0,22 - 0,32	12	36%
0,32 - 0,42	26	71%
0,42 - 0,52	16	92%
0,52 - 0,62	6	100%
Total	75	-

Anexo 8: Grado de dificultad por rango para la PSUM forma 114

Rango de Dificultad de los ítems	Cantidad de ítems	Porcentaje acumulado
0,12 - 0,22	13	17%
0,22 - 0,32	17	40%
0,32 - 0,42	18	64%
0,42 - 0,52	19	89%
0,52 - 0,62	8	100%
Total	75	-

Anexo 9: Confiabilidad promedio para PSUM forma 112

Ítems		Todos	Rama		Dependencia		
			TP	CH	Particular Pagado	Particular Subvencionado	Municipal
Cantidad		64.135	17.812	46.323	6.763	34.819	22.553
Actual	Todos los ítems (75 ítems)	0,93	0,75	0,94	0,96	0,91	0,89
Ciclo	Primer Ciclo: 1° y 2° medio (55 ítems)	0,92	0,75	0,93	0,94	0,9	0,88
	Segundo Ciclo: 3° y 4° medio (20 ítems)	0,73	0,22	0,76	0,87	0,65	0,58

Anexo 10: Confiabilidad promedio para PSUM forma 113

Ítems		Todos	Rama		Dependencia		
			TP	CH	Particular Pagado	Particular Subv.	Municipal
Cantidad		64.000	17.745	46.255	6.759	34.617	22.624
Actual	Todos los ítems (75 ítems)	0,93	0,73	0,93	0,96	0,91	0,88
Ciclo	Primer Ciclo: 1° y 2° medio (55 ítems)	0,91	0,73	0,92	0,94	0,89	0,87
	Segundo Ciclo: 3° y 4° medio (20 ítems)	0,71	0,23	0,75	0,86	0,63	0,56

Anexo 11: Confiabilidad promedio para PSUM forma 114

Ítems		Todos	Rama		Dependencia		
			TP	CH	Particular Pagado	Particular Subv.	Municipal
Cantidad		64.002	17.804	46.198	6.664	34.853	22.485
Actual	Todos los ítems (75 ítems)	0,93	0,75	0,94	0,96	0,91	0,89
Ciclo	Primer Ciclo: 1° y 2° medio (55 ítems)	0,92	0,75	0,92	0,94	0,9	0,88
	Segundo Ciclo: 3° y 4° medio (20 ítems)	0,72	0,22	0,76	0,86	0,65	0,58

Anexo 12: Correlación versión simple del PPA con agrupamiento de ítems en PSUM

Ítems	PPA
PSUM (75 ítems)	0,27
PSUM Primer Ciclo (55 ítems)	0,23
PSUM Segundo ciclo (20 ítems)	0,20
PSUM EJE-Algebra (19 ítems)	0,21
PSUM EJE-Geometría (20 ítems)	0,17
PSUM EJE-Números (16 ítems)	0,25
PSUM EJE-Datos y azar (20 ítems)	0,21

Anexo 13: *Correlación del PPA con agrupamiento de ítems según sub-área*

Área	PPA con PSUM (75 ítems)	PPA con primer ciclo de enseñanza (aplicando Corrección por atenuación)	PPA con segundo ciclo de enseñanza (aplicando Corrección por atenuación)
Ingeniería y profesiones afines	0,42	0,41	0,42
Ciencias Físicas	0,42	0,42	0,41
Informática	0,38	0,37	0,39
Protección del Medio Ambiente	0,37	0,37	0,33
Industria y Producción	0,36	0,36	0,35
Veterinaria	0,34	0,36	0,25
Agricultura, Silvicultura y Pesca	0,33	0,32	0,28
Servicios de Seguridad	0,32	0,32	0,26
Servicios Personales	0,3	0,32	0,20
Ciencias de la Vida	0,3	0,29	0,29
Matemáticas y Estadísticas	0,26	0,26	0,26
Arquitectura y Construcción	0,25	0,25	0,24
Servicios de Transporte	0,25	0,26	0,16
Salud	0,25	0,24	0,24
Enseñanza Comercial y Administración	0,23	0,23	0,21
Humanidades	0,22	0,21	0,25
Formación de Personal Docente	0,2	0,21	0,14
Ciencias Sociales y del Comportamiento	0,17	0,18	0,16
Derecho	0,16	0,15	0,18
Periodismo e Información	0,13	0,12	0,16
Servicios Sociales	0,13	0,15	0,01
Artes	0,13	0,14	0,07



Glosario y Bibliografía

8. Glosario

CH: corresponde a establecimientos educacionales de la rama Científico Humanista.

TP: corresponde a establecimientos educacionales de la rama Técnico Profesional.

PP: corresponde a establecimientos educacionales de dependencia Particular Pagado.

PS: corresponde a establecimientos educacionales Particular Subvencionado.

Municipal: corresponde a establecimientos educacionales de dependencia Municipal.

CRUCH: Consejo de Rectores de las Universidades Chilenas.

DEMRE: Departamento de Evaluación, Medición y Registro Educacional.

PPA: Promedio Ponderado Acumulado.

PSUM: Prueba de Selección Universitaria de Matemática.

9. Bibliografía

- ♦ DEMRE. (2017). Evidencia preliminar para la reducción de contenidos en la PSU de matemáticas. Santiago: Chile.
- ♦ Geiser, S., & Santelices, M.V. (2007). Validity of high school grades in predicting student success beyond the freshman year: School record vs. standardized tests as indicators of four-year college outcomes. Research & Occasional Paper Series CSHE.6.07. Center for Studies on Higher Education, University of California, Berkeley. Recuperado de http://cshe.berkeley.edu/publications/docs/ROPS.GEISER._SAT_6.12.07.pdf.
- ♦ Manzi, J., Bosch, A., Bravo, D., del Pino, G., Donoso, G. & Pizarro, R. (2010). Validez diferencial y sesgo en la predictividad de las pruebas de admisión a las universidades chilenas (PSU). Revista Iberoamericana de Evaluación Educativa, 3(2), 30-48. Recuperado de <http://www.rinace.net/riee/numeros/vol3-num2/art2.pdf>.
- ♦ Pearson Education (2013). Final Report Evaluation of the Chile PSU.
- ♦ Sistema Único de Admisión (2018). Validez predictiva de los factores de selección universitaria: Procesos de Admisión 2013 a 2015.



CONSEJO DE RECTORES DE
LAS UNIVERSIDADES CHILENAS