

DEPARTAMENTO DE EVALUACIÓN, MEDICIÓN Y REGISTRO EDUCACIONAL



INFORME DEL CÁLCULO DE PUNTAJES PSU ADMISIÓN 2020

28 de Febrero 2020

Este trabajo fue parcialmente apoyado por la infraestructura de supercómputo del NLHPC (ECM-02)

DOCUMENTO TÉCNICO ADM2020

Contenido

1	Introducción	4
2	Descripción de la Población	6
2.1	Población Inscrita y Población que Rinde	6
2.2	Año de egreso de la enseñanza media	10
2.2.1	Lenguaje	10
2.2.2	Matemática	11
2.2.3	Ciencias	12
2.3	Género	13
2.3.1	Lenguaje	13
2.3.2	Matemática	14
2.3.3	Ciencias	15
2.4	Zona Geográfica	16
2.4.1	Lenguaje	17
2.4.2	Matemática	18
2.4.3	Ciencias	19
2.5	Dependencia del establecimiento educacional de egreso	20
2.5.1	Lenguaje	20
2.5.2	Matemática	21
2.5.3	Ciencias	22
2.6	Rama educacional del establecimiento educacional de egreso	23
2.6.1	Lenguaje	23
2.6.2	Matemática	24
2.6.3	Ciencias	25
2.7	Conclusiones	26
3	Descripción de las pruebas	27
3.1	Lenguaje	28
3.2	Matemática	29
3.3	Ciencias	30

4	Análisis con Teoría Clásica del Test	32
4.1	Comparación de Dificultades	34
4.1.1	Modelo de Regresión Lineal	34
4.2	Confiabilidad	37
4.2.1	Lenguaje	38
4.2.2	Matemática	38
4.2.3	Ciencias	39
5	Análisis con Teoría de Respuesta al Ítem	40
5.1	Modelo de Rasch	40
5.1.1	Método de Estimación	42
5.1.2	Supuestos del modelo Rasch	43
5.2	Resultado de la estimación de dificultades	44
5.2.1	Calibración sobre las dificultades de los ítems	44
5.2.2	Análisis de regresión sobre las dificultades de los ítems	44
5.2.3	Estimación final de Dificultades de los ítems	54
5.3	Obtención de Puntajes	55
5.3.1	Estimación de Habilidad	55
5.3.2	Transformación de Habilidad a Puntaje PSU	55
5.4	Ajuste del modelo	58
5.5	Confiabilidad EAP	60
6	Resultados	61
6.1	Distribución de Puntaje por Prueba	61
6.2	Puntaje de Historia	64
6.3	Puntajes Faltantes de Ciencias	65
6.4	Distribución de puntajes por aplicación	66
6.5	Densidades de distribución de puntajes por grupo	68
6.5.1	Dependencia	68
6.5.2	Año de egreso de la enseñanza media	70
6.5.3	Rama Educacional	72
6.5.4	Sexo	74

Capítulo 1

Introducción

El DEMRE ejecuta anualmente el proceso de selección de postulantes a las 41 universidades que componen el Sistema Único de Admisión, por encargo del Consejo de Rectores de las Universidades Chilenas (CRUCH). Esto incluye la construcción de las pruebas de selección, su aplicación, el análisis de los resultados y su publicación, además de la recepción de las postulaciones, su procesamiento y la asignación posterior de los postulantes a las vacantes ofrecidas.

El proceso de admisión 2020 se llevó a cabo en el contexto de un importante estallido social a nivel nacional, que obligó a postergar en dos oportunidades las fechas de rendición de pruebas y a realizar tres aplicaciones, enfrentando violentas acciones de sabotaje, que incluyeron la sustracción de folletos y hojas de respuestas, así como su destrucción. El resultado más dramático de estas acciones fue la suspensión irrecuperable de la PSU de Historia, Geografía y Ciencias Sociales, que fue filtrada.

La repetición de pruebas introdujo problemas para el cálculo de puntajes que no se pueden resolver mediante la Teoría Clásica de Test, que es la metodología utilizada en todos los procesos anteriores. Por una parte, las pruebas rendidas en distintas fechas no compartían ítems en un número suficiente para poder equipararlas y, por otra parte, las poblaciones que debieron rendir en distintas fechas son muy distintas. Las acciones de sabotaje que obligaron a suspender locales de aplicación y convocar a los postulantes asignados a esos locales a nuevas aplicaciones, se concentraron en comunas con evidentes características particulares: populares y especialmente comprometidas en el estallido social. De este modo, resulta evidente que el problema de seleccionar un método de cálculo de puntajes en este caso, no es solo un problema técnico de cumplimiento de sus supuestos de base, sino un problema de justicia y equidad. La población de examinados que debió rendir dos o tres veces las pruebas no puede sufrir un perjuicio adicional por una mala elección del método de cálculo de puntajes.

Para resolver este problema se utilizó el método de Rasch, es decir, un modelo de Teoría de Respuesta al Ítem que preserva el orden de precedencia de puntajes del método de la Teoría Clásica de Test, que es simple y que estima la habilidad de los examinados en la misma escala que la dificultad de los ítems. De este modo, utilizando toda la información histórica disponible acerca del comportamiento de los

ítems participantes en las tres aplicaciones de la PSU 2020, se pudo llegar a una mejor estimación de su dificultad que la que se pudiera haber obtenido de su sola aplicación en este proceso. Para estimar la dificultad de todos los ítems involucrados en estas aplicaciones, que utilizaría la posterior estimación de la habilidad mediante el método de Rasch, se realizó una calibración concurrente que utilizó la información de todas las aplicaciones piloto en que habían participado estos ítems y de las aplicaciones oficiales en las que habían participado otros ítems presentes en las aplicaciones piloto consideradas. De este modo, la dificultad de cada uno de los ítems participantes en todas las aplicaciones de una misma prueba, se estableció con toda la información disponible comportamiento.

El principal objetivo de este informe es presentar los estudios que fundamentan estas decisiones, las que fueron compartidas con los asesores externos: Verónica Santelices, David Torres e Inés Varas.

Para la completitud del informe se presentan también la descripción de la población y de las pruebas aplicadas, el comportamiento de los ítems de acuerdo a la Teoría Clásica de Test, y las decisiones de eliminación de ítems, que fueron compartidas con las asesoras externas Claudia Matus y Rosa Montaña.

DOCUMENTO TÉCNICO ADM2020

Capítulo 2

Descripción de la Población

En este capítulo se describe la población que estaba inscrita a rendir la prueba de selección universitaria (PSU) y los que se presentaron a rendir PSU en alguna de las tres aplicaciones del proceso de admisión 2020.

También se describe por año de egreso, género, zona geográfica, dependencia y rama del establecimiento educacional, tanto de los inscritos como de los que se presentaron a rendir cada una de las pruebas en este proceso de admisión 2020, en comparación a la admisión 2019.

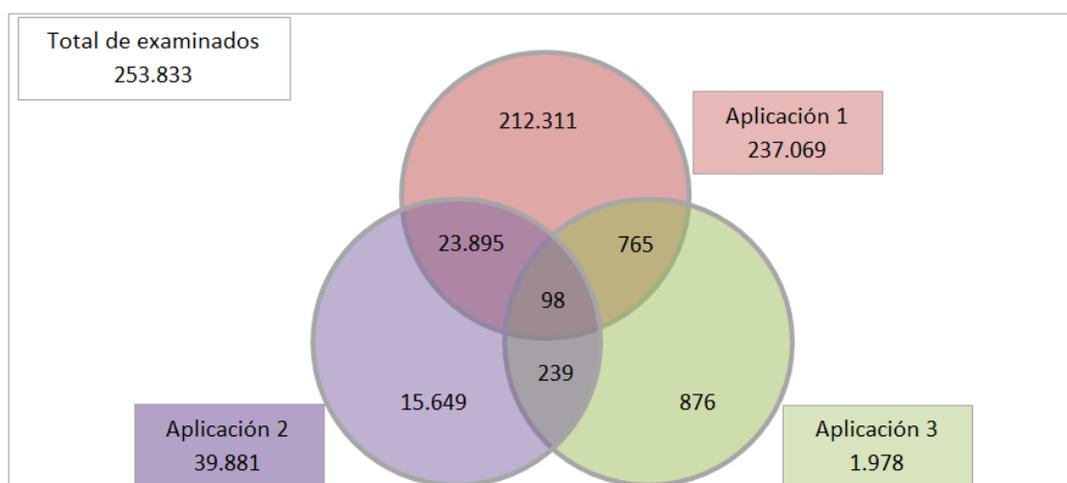
El objetivo de este capítulo exponer las diferencias entre las poblaciones que rinden en las distintas aplicaciones de acuerdo a los segmentos antes mencionados.

2.1 Población Inscrita y Población que Rinde

- Para el proceso de admisión 2020 se inscribieron 297.450 estudiantes para rendir la batería de pruebas de selección universitaria, 3.251 personas más que el proceso de admisión 2019, donde se inscribieron 294.199.

- En **Lenguaje**, 253.833 examinados al menos rindieron¹ una prueba (Ver Figura 2.1), porcentaje que corresponde al 85.3% del total de inscritos (297.450).
 - 237.069 rindieron la Aplicación 1.
 - 39.881 rindieron la Aplicación 2.
 - 1.978 rindieron la Aplicación 3.
- En la admisión 2019, el 90.7% del total de inscritos rindió la prueba de **Lenguaje** (266.939 de los 294.199).

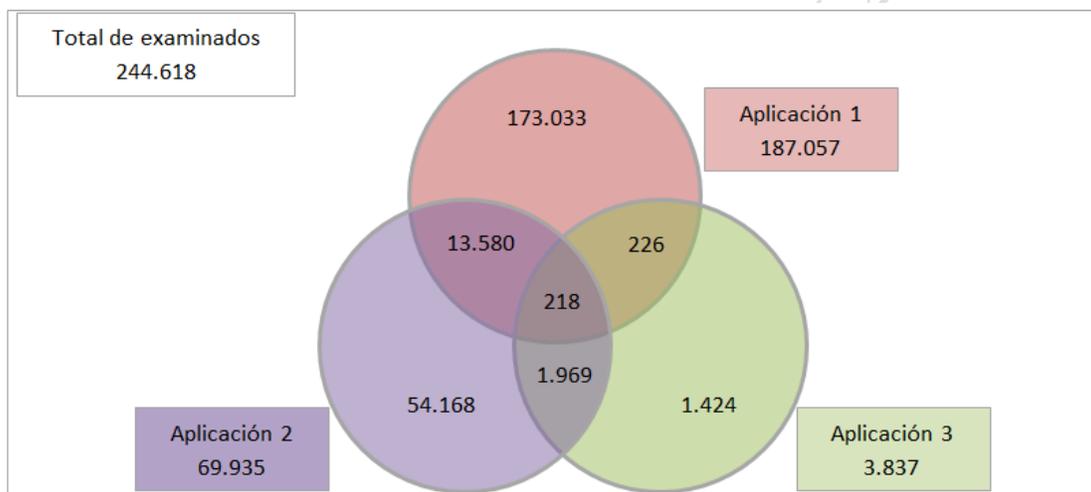
Figure 2.1: *Número de examinados por aplicación en Lenguaje*



¹Los números pueden variar de lo publicado en los días de rendición de las distintas aplicaciones, por tratarse de una estimación. El valor final de la asistencia se determina con el número de hojas de respuestas leídas.

- En **Matemática**, 244.618 examinados al menos rindieron² una prueba (Ver Figura 2.2), porcentaje que corresponde al 82.2% del total de inscritos (297.450).
 - 187.057 rindieron la Aplicación 1.
 - 69.935 rindieron la Aplicación 2.
 - 3.837 rindieron la Aplicación 3.
- En la admisión 2019, el 90.1% del total de inscritos rindió la prueba de **Matemática** (265.154 de los 294.199 inscritos).

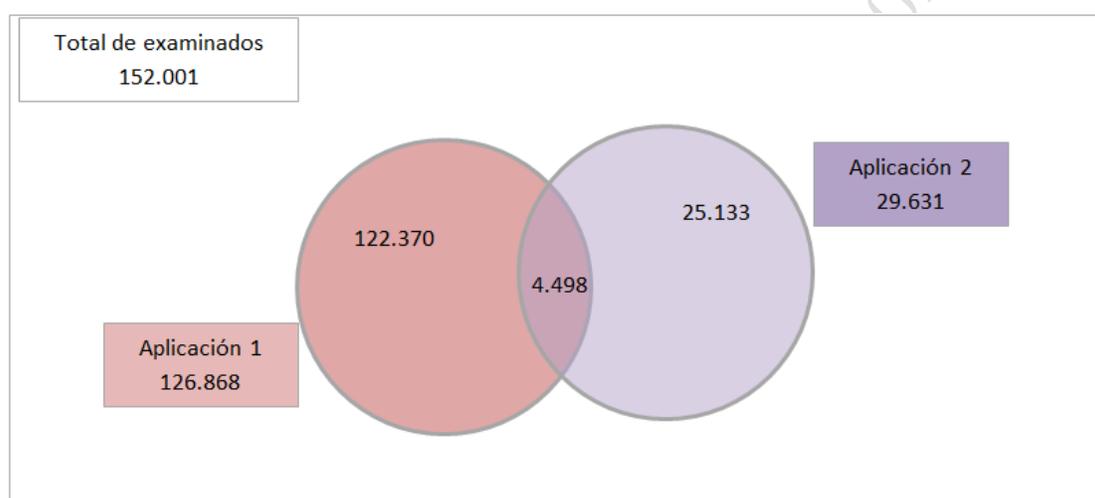
Figure 2.2: *Número de examinados por aplicación en Matemática*



²Los números pueden variar de lo publicado en los días de rendición de las distintas aplicaciones, por tratarse de una estimación. El valor final de la asistencia se determina con el número de hojas de respuestas leídas.

- En **Ciencias**, 152.001 examinados al menos rindieron³ una prueba (Ver Figura 2.3), porcentaje que corresponde al 81.6% del total de inscritos (186.359).
 - 126.868 rindieron la Aplicación 1.
 - 29.631 rindieron la Aplicación 2.
- En la admisión 2019, rindió el 87.9% del total de inscritos a la prueba de **Ciencias** (162.556 de los 184.983 inscritos).

Figure 2.3: *Número de examinados por aplicación en Ciencias*



³Los números pueden variar de lo publicado en los días de rendición de las distintas aplicaciones, por tratarse de una estimación. El valor final de la asistencia se determina con el número de hojas de respuestas leídas.

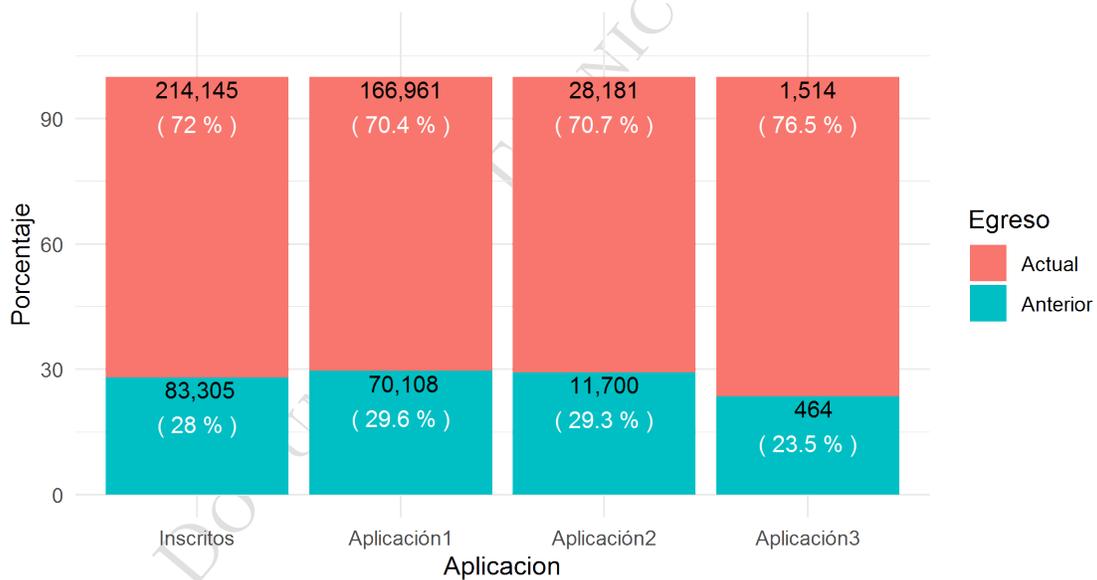
2.2 Año de egreso de la enseñanza media

El porcentaje de estudiantes inscritos que egresaron de la enseñanza media el año 2019 es 72%, porcentaje que coincide con el de la admisión anterior.

2.2.1 Lenguaje

- El porcentaje de recién egresado en la aplicación 1 de **Lenguaje** fue de 70.4%.
- El porcentaje de recién egresado en la aplicación 2 de **Lenguaje** fue de 70.7%.
- El porcentaje de recién egresado en la aplicación 3 de **Lenguaje** fue de 76.5%.
- El porcentaje de recién egresado que rindió la prueba de **Lenguaje** fue un 71.1%, mientras que en el proceso anterior correspondió al 71.7%.

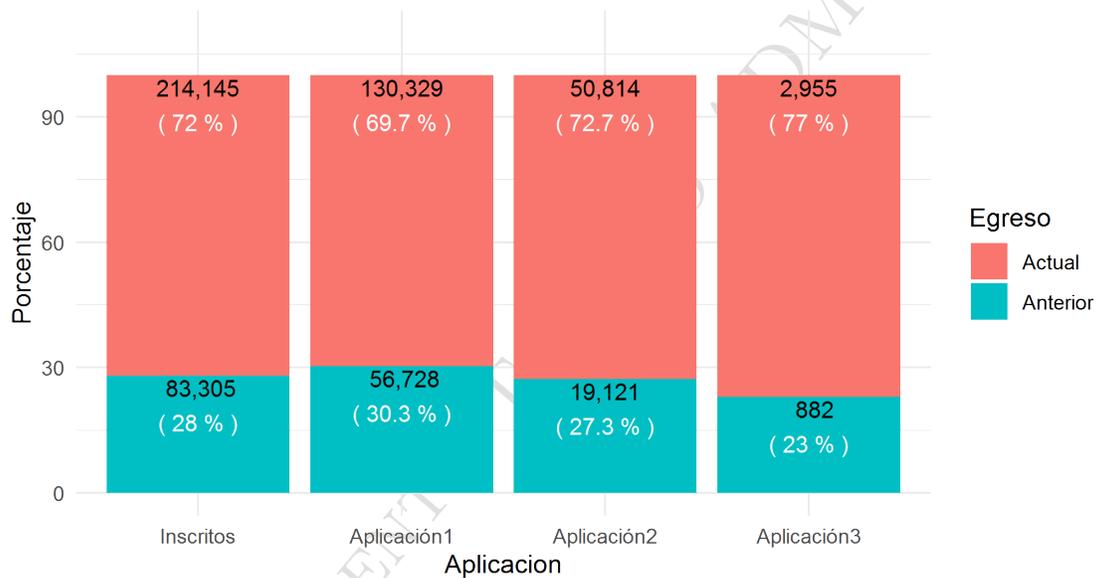
Figure 2.4: *Porcentaje de la promoción del año en cada aplicación de Lenguaje*



2.2.2 Matemática

- El porcentaje de recién egresado en la aplicación 1 de **Matemática** fue de 69.7%.
- El porcentaje de recién egresado en la aplicación 2 de **Matemática** fue de 72.7%.
- El porcentaje de recién egresado en la aplicación 3 de **Matemática** fue de 77%.
- El porcentaje de recién egresado que rindió la prueba de **Matemática** fue un 70.8%, mientras que en el proceso anterior correspondió al 71.6%.

Figure 2.5: *Porcentaje de la promoción del año en cada aplicación de Matemática*

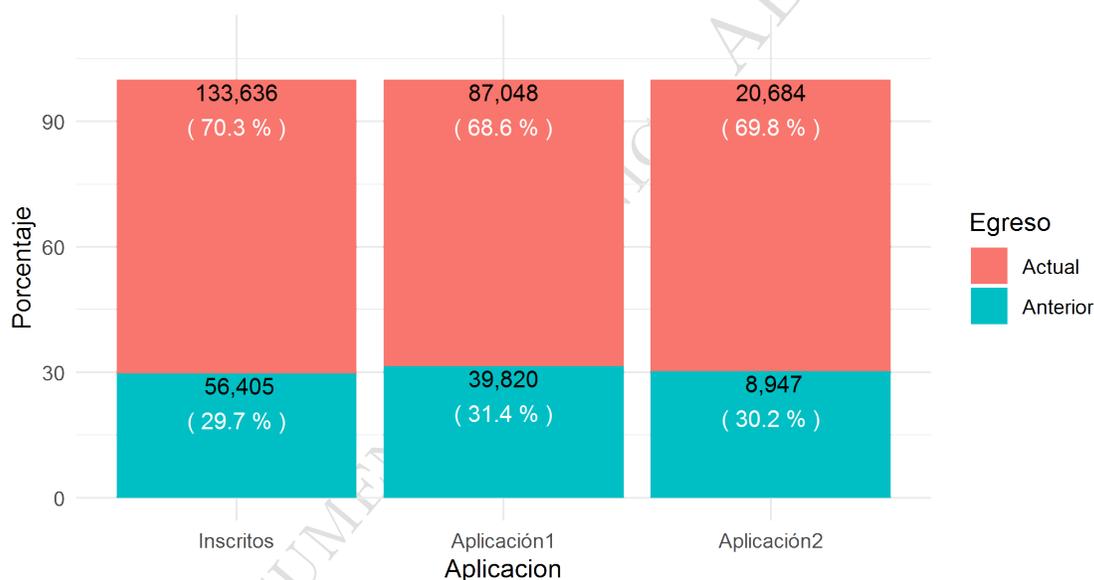


2.2.3 Ciencias

El porcentaje de estudiantes inscritos que egresaron de la enseñanza media el año 2019 es 70.3%, con respecto a la admisión anterior, 130.112 eran promoción del año, porcentaje que coincide con el del proceso anterior.

- El porcentaje de recién egresado en la aplicación 1 de **Ciencias** fue de 68.6%.
- El porcentaje de recién egresado en la aplicación 2 de **Ciencias** fue de 69.8%.
- El porcentaje de recién egresado que rindió la prueba de **Ciencias** fue un 69.3%, mientras que el año pasado correspondió al 70%.

Figure 2.6: *Porcentaje de la promoción del año en cada aplicación de Ciencias*



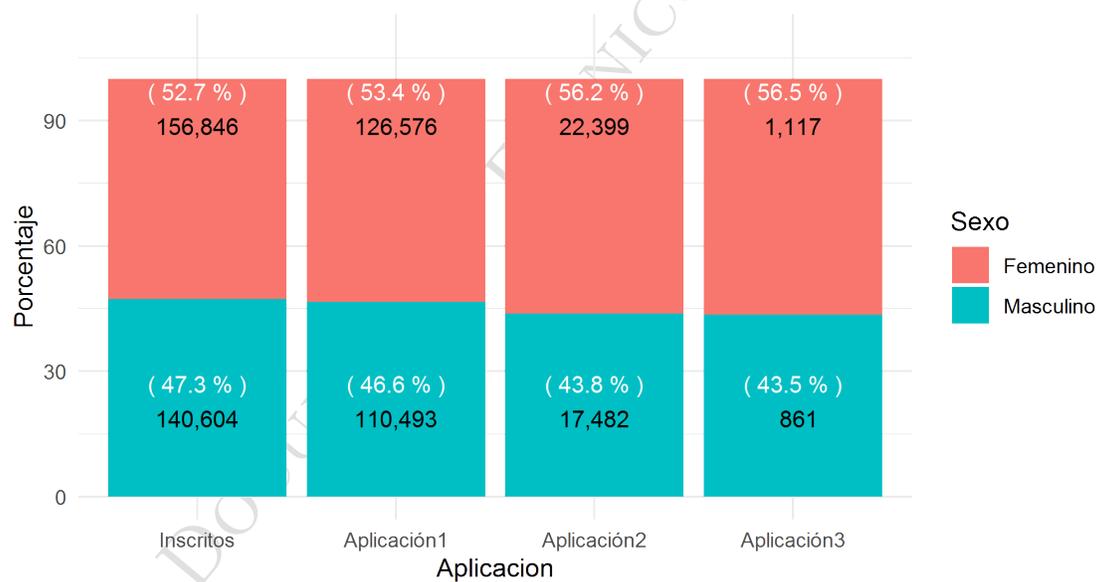
2.3 Género

El porcentaje de estudiantes inscritos de género femenino en este proceso corresponde a 52.7%. En el proceso anterior este porcentaje fue de 52.6%.

2.3.1 Lenguaje

- El porcentaje de examinadas de género femenino en la aplicación 1 de **Lenguaje** fue de 53.4%.
- El porcentaje de examinadas de género femenino en la aplicación 2 de **Lenguaje** fue de 56.2%.
- El porcentaje de examinadas de género femenino en la aplicación 3 de **Lenguaje** fue de 56.5%.
- El porcentaje de examinadas de género femenino que rindió la prueba de **Lenguaje** fue de un 53.5%, mientras que el año pasado correspondió al 53%.

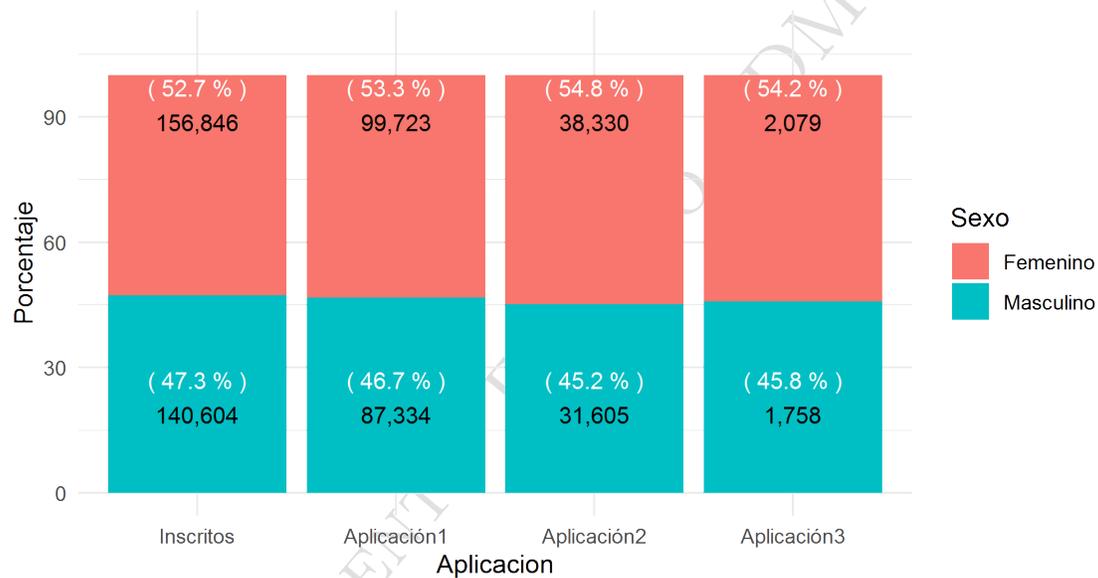
Figure 2.7: *Distribución según el género de los postulantes en cada aplicación de Lenguaje*



2.3.2 Matemática

- El porcentaje de examinadas de género femenino en la aplicación 1 de **Matemática** fue de 53.3%.
- El porcentaje de examinadas de género femenino en la aplicación 2 de **Matemática** fue de 54.8%.
- El porcentaje de examinadas de género femenino en la aplicación 3 de **Matemática** fue de 54.2%.
- El porcentaje de examinadas de género femenino que rindió la prueba de **Matemática** fue un 53.6%, mientras que el año pasado correspondió al 53%.

Figure 2.8: *Distribución según el género de los postulantes en cada aplicación de Matemática*

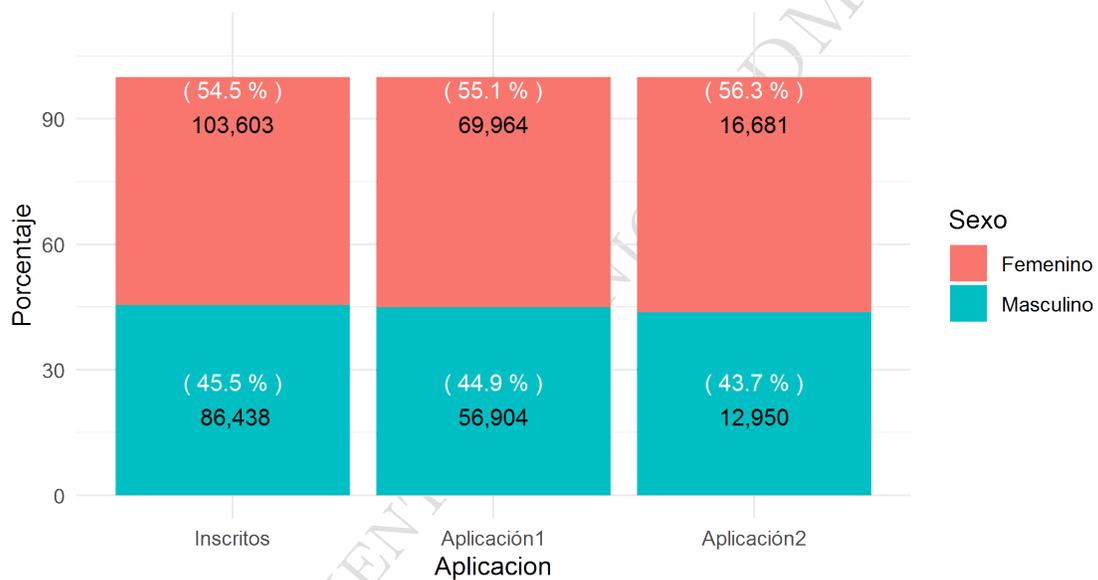


2.3.3 Ciencias

El porcentaje de estudiantes inscritos de sexo femenino en este proceso corresponde a 54.5%.

- El porcentaje de examinadas de género femenino en la aplicación 1 de **Ciencias** fue de 55.1%.
- El porcentaje examinadas de género femenino en la aplicación 2 de **Ciencias** fue de 56.3%.
- El porcentaje examinadas de género femenino que rindió la prueba de **Ciencias** fue un 55.2%, mientras que el año pasado correspondió al 54.5%.

Figure 2.9: *Distribución según el género de los postulantes en cada aplicación de Ciencias*



2.4 Zona Geográfica

Las 16 regiones del país fueron divididas en tres zonas geográficas que se describen a continuación:

- Zona Norte: Corresponde a las regiones del norte de Chile,
 - Región de Arica y Parinacota
 - Región de Tarapacá
 - Región de Antofagasta
 - Región de Atacama
 - Región de Coquimbo.
- Zona Centro: Corresponde a las regiones del centro de Chile,
 - Región de Valparaíso
 - Región de Metropolitana de Santiago
 - Región del Libertador General Bernardo O'Higgins
 - Región del Maule
 - Región del Ñuble.
- Zona Sur: Corresponde a las regiones del sur de Chile,
 - Región del Biobío
 - Región de La Araucanía
 - Región de Los Ríos
 - Región de Los Lagos
 - Región de Aysén del General Carlos Ibáñez del Campo
 - Región de Magallanes y la Antártica Chilena.

2.4.1 Lenguaje

- Zona Norte: El porcentaje de inscritos en esta zona corresponde a 12.8%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 11.8%, 25.1% y 0.9% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Lenguaje** fue de un 12.6%, mientras que en la admisión 2019 correspondió al 12.5%.

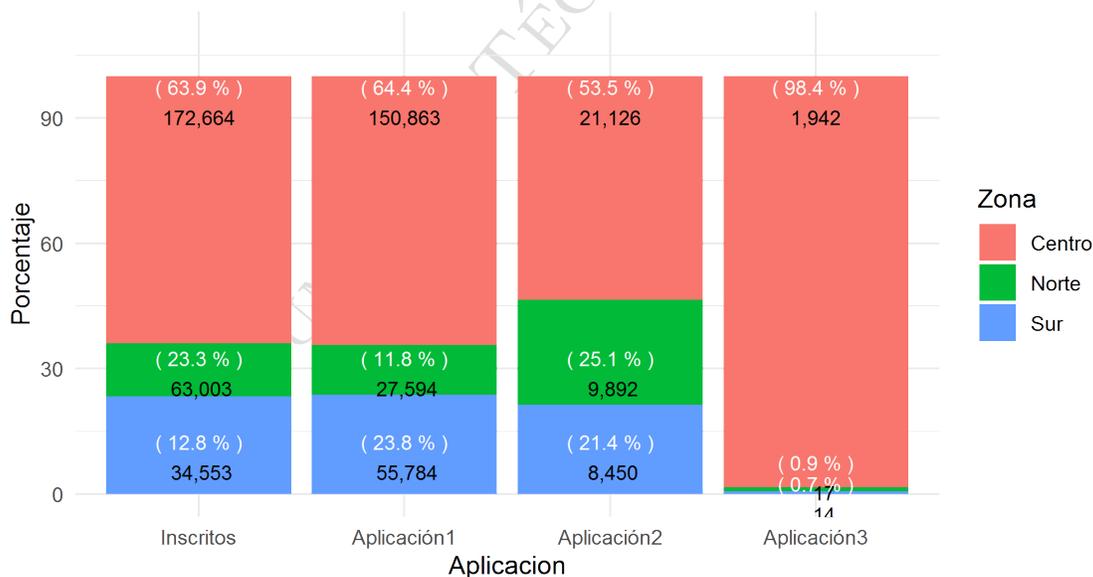
- Zona Centro: El porcentaje de inscritos en esta zona corresponde a 63.9%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 64.4%, 53.5% y 98.4% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Lenguaje** fue de un 64%, mientras que en la admisión 2019 correspondió al 64.2%.

- Zona Sur: El porcentaje de inscritos en esta zona corresponde a 23.3%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 23.8%, 21.4% y 0.7% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Lenguaje** fue un 23.5%, mientras que en la admisión 2019 correspondió al 23.3%.

Figure 2.10: *Distribución según la zona geográfica de los postulantes en cada aplicación de Lenguaje*



2.4.2 Matemática

- Zona Norte: El porcentaje de inscritos en esta zona corresponde a 12.8%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 9.2%, 20.6% y 0.7% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Matemática** fue un 12.4%, mientras que en la admisión 2019 correspondió al 12.5%.

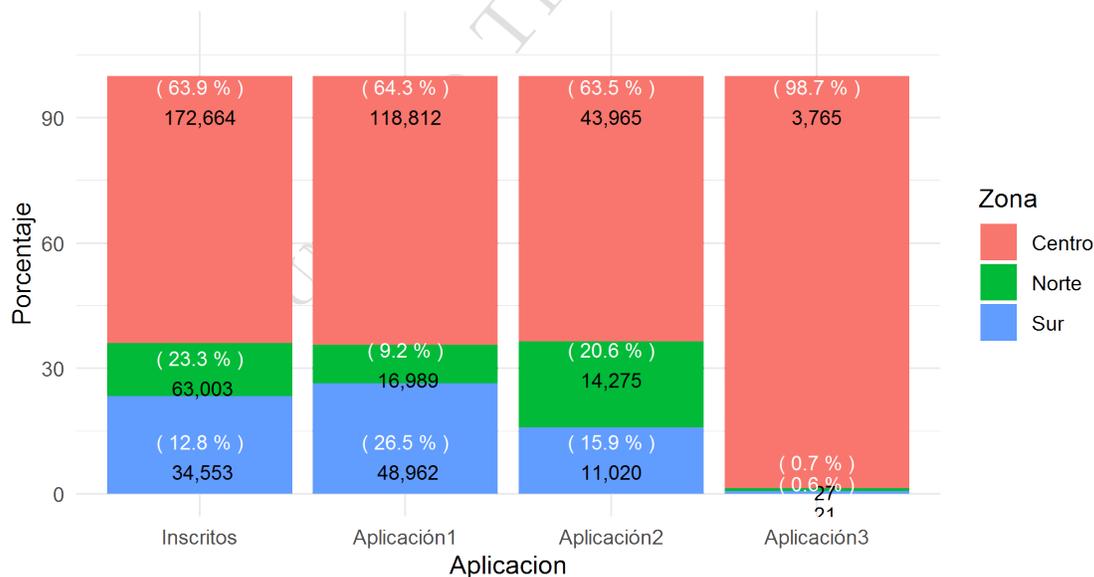
- Zona Centro: El porcentaje de inscritos en esta zona corresponde a 63.9%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 64.3%, 63.5% y 98.7% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Matemática** fue de un 63.7%, mientras que en la admisión 2019 correspondió al 64.2%.

- Zona Sur: El porcentaje de inscritos en esta zona corresponde a 23.3%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 26.5%, 15.9% y 0.6% respectivamente.

En general, el porcentaje de examinados de esta zona que rindió la prueba de **Matemática** fue de un 23.8%, mientras que en la admisión 2019 correspondió al 23.3%.

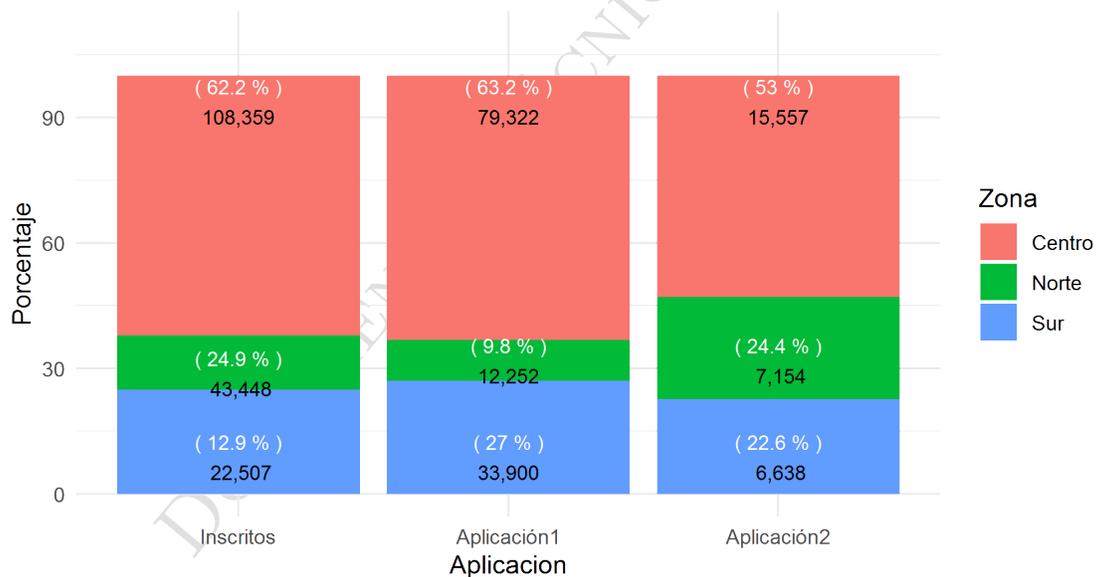
Figure 2.11: Distribución según la zona geográfica de los postulantes en cada aplicación de Matemática



2.4.3 Ciencias

- Zona Norte: El porcentaje de inscritos en esta zona corresponde a 12.9%. Mientras, los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 9.8% y 24.4% respectivamente. En general, el porcentaje de examinados de esta zona que rindió la prueba de **Ciencias** fue un 12.6%, mientras que en la admisión 2019 correspondió al 12.7%.
- Zona Centro: El porcentaje de inscritos en esta zona corresponde a 62.2%. Mientras, los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 63.2% y 53% respectivamente. En general, el porcentaje de examinados de esta zona que rindió la prueba de **Ciencias** fue de un 61.9%, mientras que en la admisión 2019 correspondió al 62.4%.
- Zona Sur: El porcentaje de inscritos en esta zona corresponde a 24.9%. Mientras, los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 27% y 22.6% respectivamente. En general, el porcentaje de examinados de esta zona que rindió la prueba de **Ciencias** fue un 25.5%, mientras que en la admisión 2019 correspondió al 24.8%.

Figure 2.12: *Distribución según la zona geográfica de los postulantes en cada aplicación de Ciencias*



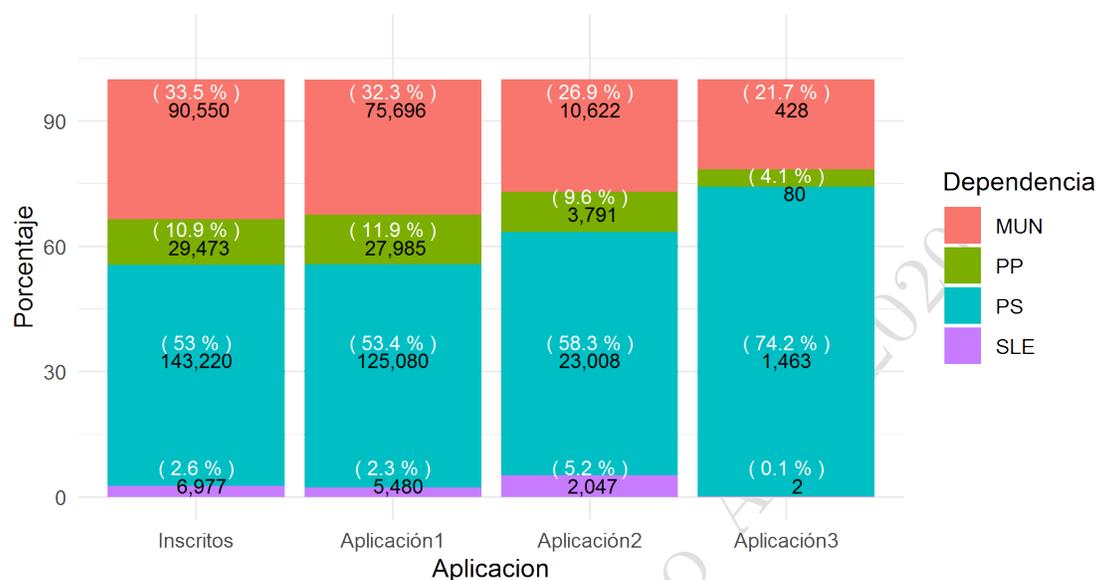
2.5 Dependencia del establecimiento educacional de egreso⁴

2.5.1 Lenguaje

- Particular Pagado: El porcentaje de inscritos es 10.9%. En las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 11.9%, 9.6% y 4.1% respectivamente. En general, el porcentaje de examinados que rindió **Lenguaje** fue de un 11.7%, mientras que en la admisión 2019 fue de un 11.4%.
- Particular Subvencionado: El porcentaje de inscritos es 53%. En las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 53.4%, 58.3% y 74.2% respectivamente. En general, el porcentaje de examinados que rindió **Lenguaje** fue de un 53.6%, mientras que en la admisión 2019 fue un 53.1%.
- Municipal: El porcentaje de inscritos es de 33.5%. En las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 32.3%, 26.9% y 21.7% respectivamente. En general, el porcentaje de examinados que rindió **Lenguaje** fue de un 32.2%, mientras que en admisión de 2019 fue un 34.4%.
- Servicio Local de Educación: El porcentaje de inscritos es de 2.6%. En las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 2.3%, 5.2% y 0.1% respectivamente. En general, el porcentaje de examinados que rindió **Lenguaje** fue de un 2.5%, mientras que en la admisión 2019 fue de un 1.1%.

⁴En este análisis se consideró a todos los estudiantes que tuviesen información del establecimiento educacional del último año de enseñanza media, por lo cual se excluyeron todos los casos de extranjeros, convalidación de estudios entre otros.

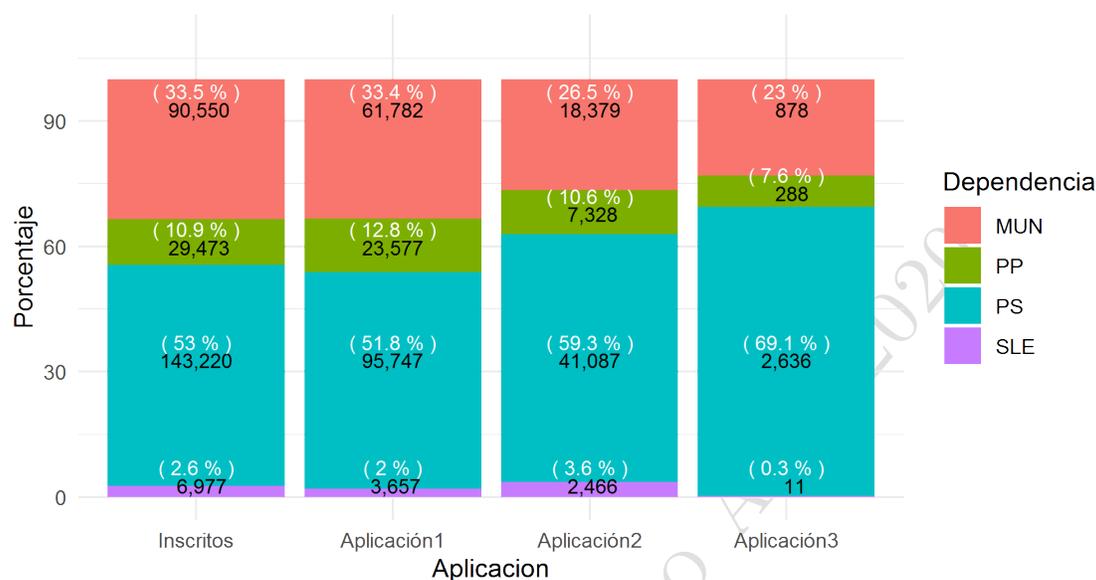
Figure 2.13: *Distribución por la dependencia del establecimiento de egreso de los postulantes en cada aplicación de Lenguaje*



2.5.2 Matemática

- Particular Pagado: El porcentaje de inscritos en esta dependencia corresponde a 10.9%. Los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 12.8%, 10.6% y 7.6% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Matemática** fue un 12%, mientras que en la admisión 2019 correspondió al 11.4%.
- Particular Subvencionado: El porcentaje de inscritos en esta dependencia corresponde a 53%. Los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 51.8%, 59.3% y 69.1% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Matemática** fue un 53.6%, mientras que en la admisión 2019 correspondió al 53.1%.
- Municipal: El porcentaje de inscritos en esta dependencia corresponde a 33.5%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 33.4%, 26.5% y 23% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Matemática** fue un 32%, mientras que en la admisión 2019 correspondió al 34.4%.
- Servicio Local de Educación: El porcentaje de inscritos en esta dependencia corresponde a 2.6%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 2%, 3.6% y 0.3% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Matemática** fue un 2.4%, mientras que en la admisión 2019 correspondió al 1.1%.

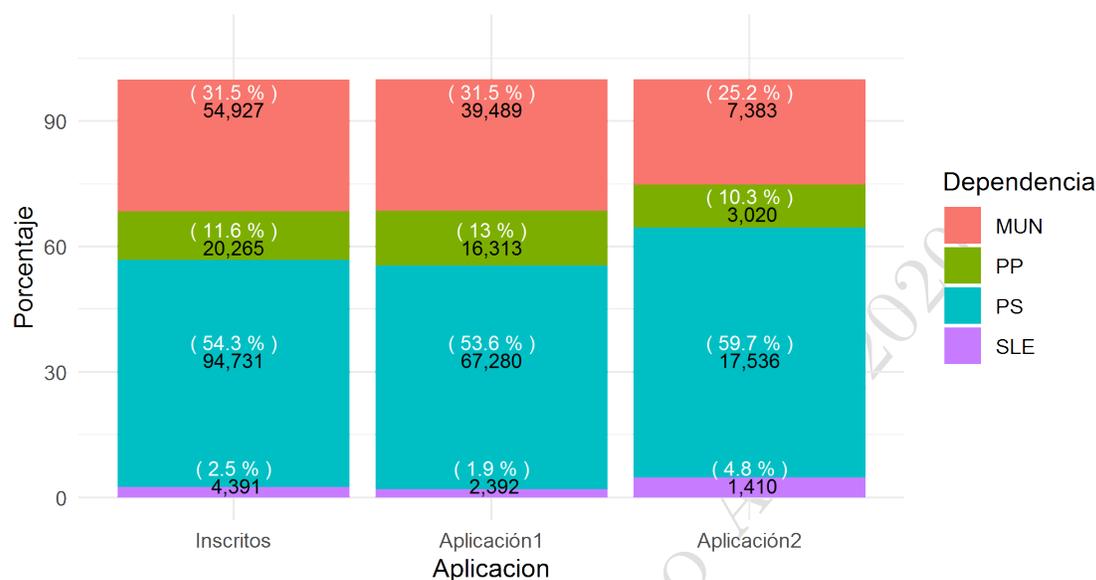
Figure 2.14: *Distribución por la dependencia del establecimiento de egreso de los postulantes en cada aplicación de Matemática*



2.5.3 Ciencias

- Particular Pagado: El porcentaje de inscritos en esta dependencia corresponde a 11.6%. Los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 13% y 10.3% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Ciencias** fue un 12.4%, mientras que en la admisión 2019 correspondió al 12%.
- Particular Subvencionado: El porcentaje de inscritos en esta dependencia corresponde a 54.3%. Los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 53.6% y 59.7% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Ciencias** fue un 54.6%, mientras que en la admisión 2019 correspondió al 54.2%.
- Municipal: El porcentaje de inscritos en esta dependencia corresponde a 31.5%. Los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 31.5% y 25.2% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Ciencias** fue un 30.6%, mientras que en la admisión 2019 correspondió al 32.8%.
- Servicio Local de Educación: El porcentaje de inscritos en esta dependencia corresponde a 2.5%. Los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 1.9% y 4.8% respectivamente. En general, el porcentaje de examinados de esta dependencia que rindió **Ciencias** fue un 2.4%, mientras que en la admisión 2019 correspondió al 1%.

Figure 2.15: *Distribución por la dependencia del establecimiento de egreso de los postulantes en cada aplicación de Ciencias*



2.6 Rama educacional del establecimiento educacional de egreso

2.6.1 Lenguaje

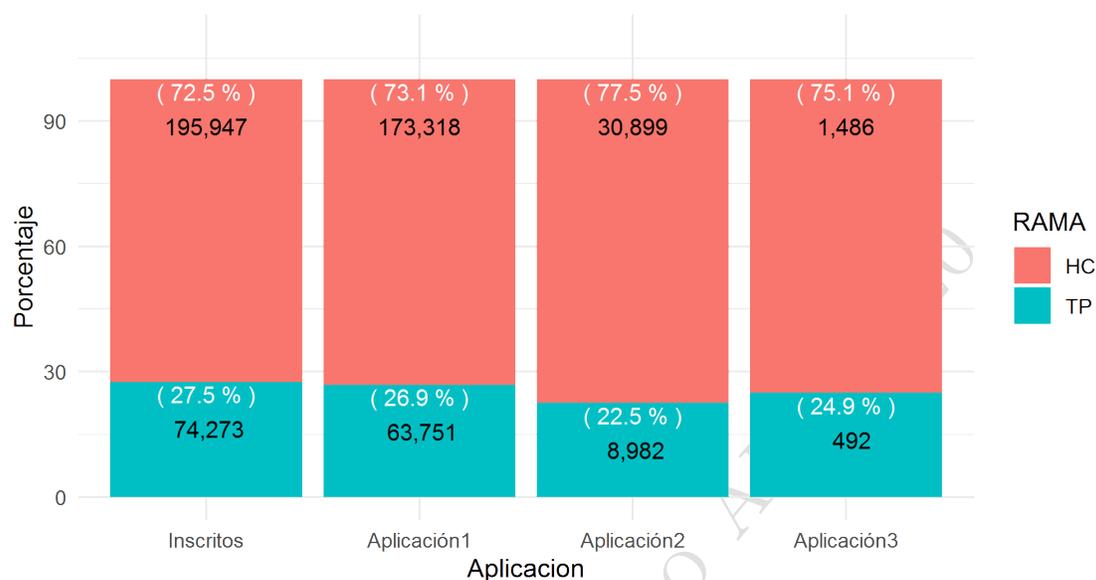
- Rama Educacional Humanista Científica (HC): El porcentaje de inscritos en esta rama corresponde a 77.8%. Mientras, los porcentajes en las aplicación 1, 2 y 3 de **Lenguaje** fueron de 73.1%, 77.5% y 75.1% respectivamente.

En general, el porcentaje de examinados de esta rama que rindió la prueba de **Lenguaje** fue de un 73.3%, mientras que en la admisión 2019 correspondió al 73.2%.

- Rama Educacional Técnico Profesional (TP): El porcentaje de inscritos en esta rama corresponde a 22.2%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Lenguaje** fueron de 26.9%, 22.5% y 24.9% respectivamente.

En general, el porcentaje de examinados de esta rama que rindió la prueba de **Lenguaje** fue de un 26.7%, mientras que en la admisión 2019 correspondió al 26.8%.

Figure 2.16: *Distribución por rama educacional del establecimiento de egreso de los postulantes en cada aplicación de Lenguaje*



2.6.2 Matemática

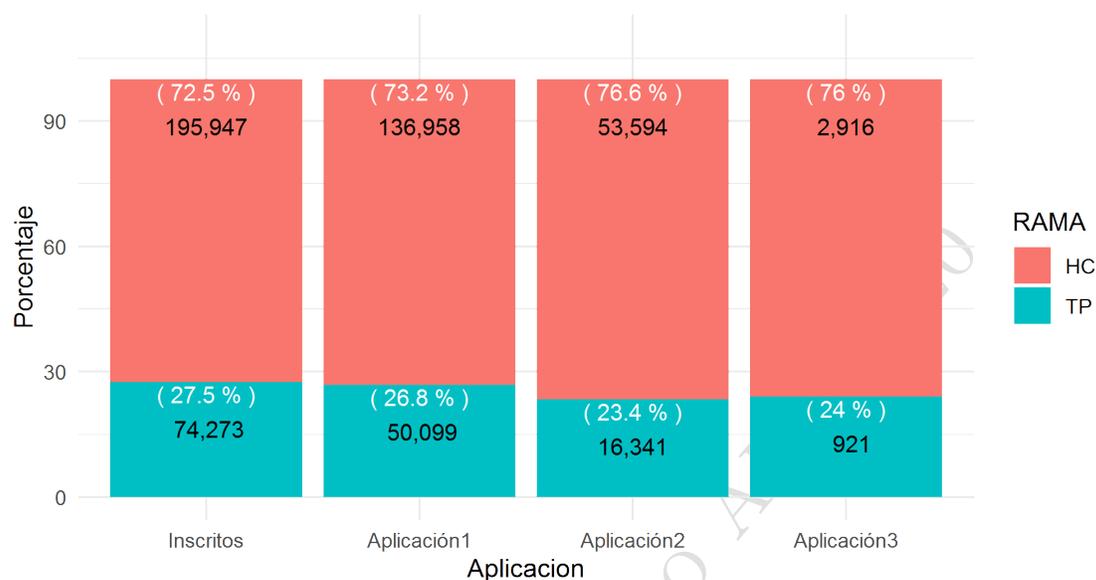
- Rama Educacional HC: El porcentaje de inscritos en esta rama corresponde a 77.8%. Mientras, los porcentajes en las aplicación 1, 2 y 3 de **Matemática** fueron de 73.2%, 76.6% y 76% respectivamente.

En general, el porcentaje de examinados de esta rama que rindió la prueba de **Matemática** fue de un 73.8%, mientras que en la admisión 2019 correspondió al 73.3%.

- Rama Educacional TP: El porcentaje de inscritos en esta rama corresponde a 22.2%. Mientras, los porcentajes en las aplicaciones 1, 2 y 3 de **Matemática** fueron de 26.8%, 23.4% y 24% respectivamente.

En general, el porcentaje de examinados de esta rama que rindió la prueba de **Matemática** fue un 26.2%, mientras que en la admisión 2019 correspondió al 26.7%.

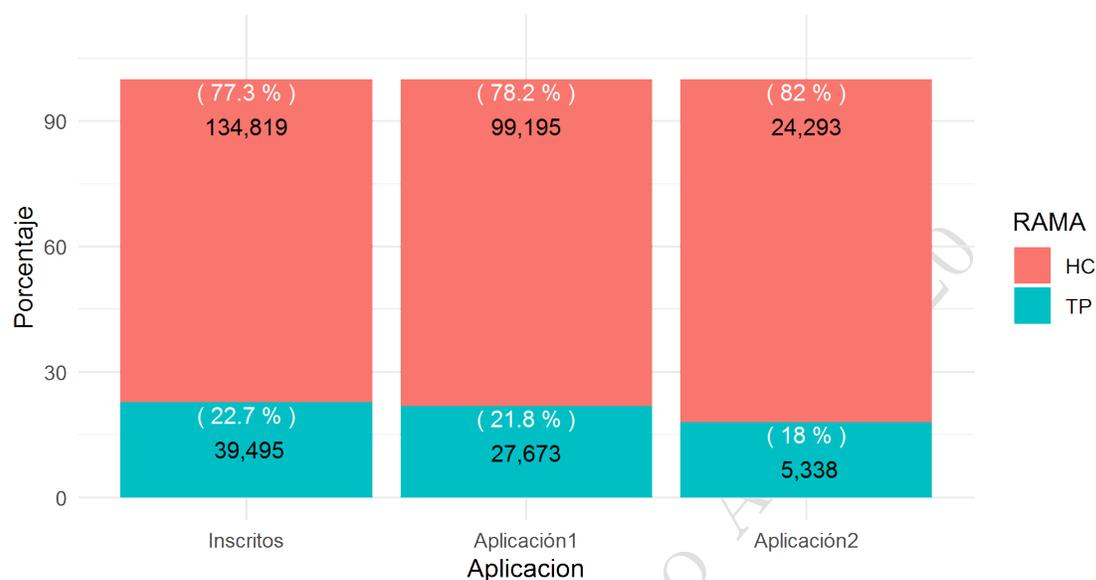
Figure 2.17: *Distribución por rama educacional del establecimiento de egreso de los postulantes en cada aplicación de Matemática*



2.6.3 Ciencias

- Rama Educacional HC: El porcentaje de inscritos en esta rama corresponde a 77.3%. Mientras, los porcentajes en las aplicación 1 y 2 de **Ciencias** fueron de 78.2% y 82% respectivamente. En general, el porcentaje de examinados de esta rama que rindió la prueba de **Ciencias** fue de un 78.7%, mientras que en la admisión 2019 correspondió al 78.1%.
- Rama Educacional TP: El porcentaje de inscritos en esta rama corresponde a 22.7%. Mientras, los porcentajes en las aplicaciones 1 y 2 de **Ciencias** fueron de 21.8% y 18% respectivamente. En general, el porcentaje de examinados de esta rama que rindió la prueba de **Ciencias** fue un 21.3%, mientras que en la admisión 2019 correspondió al 21.9%.

Figure 2.18: *Distribución por rama educacional del establecimiento de egreso de los postulantes en cada aplicación de Ciencias*



2.7 Conclusiones

Ante el complicado proceso de admisión vivido este año es importante destacar que a pesar de que la asistencia a las pruebas disminuyó con respecto al proceso anterior, aumentó el porcentaje de postulantes habilitados para postular a la universidad, lo cual indicaría que, frente a las complicaciones, aquellos postulantes que tenían más expectativas de continuar con el proceso de selección fueron más perseverantes.

Además, se muestra que las poblaciones que rindieron las distintas aplicaciones de cada prueba son distintas, donde es posible señalar que en la tercera aplicación estaba compuesta casi totalmente por postulantes de la zona central o que en cada aplicación el porcentaje de postulantes proveniente de colegios particulares pagados disminuía y el de particulares subvencionados aumentaba en cada una de las pruebas, y así otras diferencias descritas anteriormente, evidenciarían la imposibilidad de aplicar teoría clásica para el cálculo de puntajes al tener formas en cada prueba que dificultan la comparación entre sí.

Capítulo 3

Descripción de las pruebas

Las PSU de Lenguaje, Matemática, Ciencias e Historia, Geografía y Ciencias Sociales de este proceso fueron ensambladas del mismo modo que en años anteriores. Todas las pruebas tienen 80 preguntas y se construyen en cuatro formas, excepto las pruebas de Ciencias que solo tienen dos formas paralelas por cada una de las cuatro electivas: Biología, Física, Química y Técnico Profesional. Las pruebas de Lenguaje, Matemáticas e Historia tienen 75 preguntas operativas y 5 preguntas piloto.

Cuando el estallido social del mes de octubre obligó a postergar la fecha de aplicación de las pruebas y se contempló la posibilidad de realizar aplicaciones adicionales si no se lograba que todos los postulantes rindieran en condiciones adecuadas, se tomó la decisión de construir pruebas del mismo largo (80 preguntas), pero con igual número de preguntas operativas distintas para cada aplicación. Eso produjo que las pruebas de Ciencias no compartieran ningún ítem entre aplicaciones y las demás pruebas compartieran 5 ítems entre aplicaciones.

En la Tabla 3.1 se muestra la cantidad de formas utilizadas en cada prueba y en cada aplicación, con los códigos que las identifican.

Table 3.1: *Listado de códigos de formas utilizadas durante la aplicación de la Prueba de Selección Universitaria (PSU), Admisión 2020.*

	Primera aplicación	Segunda aplicación	Tercera aplicación
Prueba de Lenguaje	101, 102, 103 y 104	105 y 106	107
Prueba de Matemática	111, 112, 113 y 114	115 y 116	117
Prueba de Ciencias Biología	151 y 152	153 y 154	No Aplica
Prueba de Ciencias Física	161 y 162	163 y 164	No Aplica
Prueba de Ciencias Química	171 y 172	173 y 174	No Aplica
Prueba de Ciencias Técnico Profesional	181 y 182	183 y 184	No Aplica

3.1 Lenguaje

La prueba de Lenguaje se aplicó en tres ocasiones distintas y mediante siete formas, cada una de las cuales estaba compuesta por 80 preguntas en total (75 preguntas operativas y 5 preguntas piloto).

Para la primera aplicación, se utilizaron 4 formas de códigos 101, 102, 103 y 104 (ver Tabla 3.1). La segunda aplicación contempló dos formas, cuyos códigos fueron 105 y 106, y en la tercera aplicación se utilizó solamente una forma, la 107.

Mientras que las cuatro formas de la prueba de Lenguaje aplicadas para la primera fecha comparten algunas de las preguntas entre ellas. Para la segunda aplicación las dos formas utilizadas estuvieron compuestas por las mismas 80 preguntas. El detalle de los ítems compartidos entre formas se entrega en la tabla a continuación:

Table 3.2: *Ítems compartidos entre formas de la prueba de Lenguaje*

	Formas contrastadas	Número de ítems compartidos
Primera aplicación	101 y 102	29
	101 y 103	31
	101 y 104	33
	102 y 103	33
	102 y 104	31
	103 y 104	29
	Segunda aplicación	105 y 106

Por otra parte, en lo que refiere a los ítems compartidos entre aplicaciones, es posible afirmar que la primera, la segunda y la tercera aplicación comparten 5 ítems, correspondientes a preguntas piloto.

Del total de ítems operativos, 25 presentaron anomalías de acuerdo a los criterios técnicos para determinar la aceptación, rechazo o revisión de los ítems que han sido ensamblados en las pruebas oficiales. De éstos, 11 pertenecen a la primera aplicación, 3 a la segunda y 11 a la tercera. En total, se determinó eliminar 9 y se mantuvieron 16.

Su distribución en las distintas formas aplicadas se presenta en la tabla a continuación:

Table 3.3: *Distribución de ítems fuera de rango en la prueba de Lenguaje*

Aplicación	Forma	Número de ítems fuera de rango			Número final de preguntas para el Puntaje
		Total	Permanecen	Eliminados	
Primera	101	6	3	3	72
	102	7	4	3	72
	103	7	3	4	71
	104	4	2	2	73
Segunda	105	3	3	0	75
	106	3	3	0	75
Tercera	107	11	7	4	71

3.2 Matemática

La prueba de Matemática se aplicó en tres fechas y mediante siete formas. Cada una de las formas estuvo compuesta por 80 preguntas, 75 preguntas que se utilizan para el cálculo de puntaje y 5 preguntas piloto.

Las formas correspondientes a la prueba de Matemática son la 111, 112, 113 y 114 para la primera aplicación; en la segunda aplicación se hizo uso de las formas 115 y 116 y la última aplicación contempló únicamente la forma 117 (ver Tabla 3.1).

Al igual que en el caso de Lenguaje, para la primera aplicación de la prueba de Matemáticas las cuatro formas presentan ítems compartidos (ver Tabla 3.4). En la segunda fecha de aplicación, en cambio, las formas 115 y 116 comparten la totalidad de sus preguntas.

En lo que respecta a los ítems compartidos entre aplicaciones, es necesario tener en cuenta que para el ensamblaje de las formas 115, 116 y 117 se seleccionaron 5 ítems operativos de la primera aplicación y se incorporaron en las nuevas formas como preguntas piloto. En consecuencia, la primera, la segunda y la tercera aplicación tienen 5 preguntas en común; el resto de los ítems difiere.

Table 3.4: *Ítems compartidos entre formas de la prueba de Matemática*

	Formas contrastadas	Número de ítems compartidos
Primera aplicación	111 y 112	29
	111 y 113	27
	111 y 114	19
	112 y 113	19
	112 y 114	27
	113 y 114	29
	Segunda aplicación	115 y 116

En el total de formas utilizadas, 10 ítems se encontraron fuera de rango de acuerdo a los criterios psicométricos establecidos para las pruebas oficiales: 3 de la primera aplicación, 4 de la segunda y 3 de la tercera. De éstos, se resolvió la mantención de 6 y la eliminación de 4. Su distribución en las distintas formas aplicadas se presenta en la tabla a continuación:

Table 3.5: *Distribución de ítems fuera de rango en la prueba de Matemática*

Aplicación	Forma	Número de ítems fuera de rango			Número final de preguntas para el Puntaje
		Total	Permanecen	Eliminados	
Primera	111	2	1	1	74
	112	2	1	1	74
	113	2	2	0	75
	114	0	-	-	75
Segunda	115	4	1	3	72
	116	4	1	3	72
Tercera	117	3	3	0	75

3.3 Ciencias

A diferencia de las pruebas anteriores, la prueba de Ciencias se aplicó en dos instancias, en cada una de las cuales se hizo uso de dos formas por cada módulo electivo. En total, se utilizaron 16 formas (ver Tabla 3.1).

Cada forma de la prueba de Ciencias está constituida por 80 preguntas, 54 de las cuales pertenecen al módulo común (18 de cada disciplina) y 26 correspondientes al módulo electivo.

Para la primera aplicación, el módulo común de las ocho formas aplicadas estaba compuesto por los mismos 54 ítems. A su vez, las dos formas pertenecientes a cada electivo comparten también los 26

ítems del módulo específico y, como tal, es posible afirmar que estas dos formas están compuestas por las mismas 80 preguntas. A modo de ejemplo: las formas 151 y 152 -ambas de un mismo electivo- comparten la totalidad de los ítems, mientras que las formas 151 y 161 -correspondientes a distintos electivos- solamente comparten las 54 preguntas pertenecientes al módulo común.

Esta situación se repitió para la segunda aplicación, no obstante, entre la primera y segunda aplicación no se compartió ningún ítem.

De la totalidad de preguntas aplicadas, 4 ítems se encontraron fuera de rango de acuerdo a los criterios psicométricos establecidos para determinar la aceptación, rechazo o revisión de los ítems que han sido ensamblados en las pruebas oficiales: 3 en la primera aplicación y 1 en la segunda. De éstos, se decidió eliminar 2 y se mantuvieron 5.

Su distribución en las distintas formas aplicadas se presenta en la tabla a continuación:

Table 3.6: *Distribución de ítems fuera de rango en la prueba de Ciencias*

Aplicación	Forma	Número de ítems fuera de rango			Número final de preguntas para el Puntaje
		Total	Permanecen	Eliminados	
Primera	151	3	1	2	78
	152	3	1	2	78
	161	1	1	0	80
	162	1	1	0	80
	171	1	1	0	80
	172	1	1	0	80
	181	1	1	0	80
	182	1	1	0	80
Segunda	153	1	1	0	80
	154	1	1	0	80
	163	1	1	0	80
	164	1	1	0	80
	173	1	1	0	80
	174	1	1	0	80
	183	1	1	0	80
	184	1	1	0	80

Capítulo 4

Análisis con Teoría Clásica del Test

El cálculo de puntajes en la batería de pruebas PSU hasta la admisión 2019 fue realizado mediante Teoría Clásica del Test (CTT en inglés), teoría que propone que toda medición observada X es generada a través de un puntaje verdadero T más un error E , es decir, $X = T + E$.

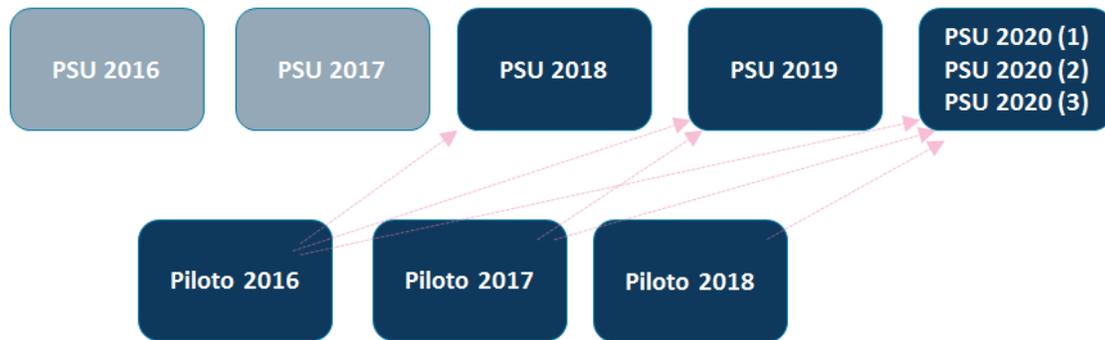
Bajo esta teoría el cálculo de puntaje se realiza contando el número de respuestas correctas que obtiene cada postulante. En todas las aplicaciones de las pruebas PSU se consideran distintas formas (Ver capítulo 3) que se construyen resguardando equivalencia de contenido. Para asegurar que las puntuaciones en las distintas formas representen un mismo desempeño, se debe realizar un procedimiento de equiparación entre éstas. Los procedimientos de equiparación que se utilizan bajo esta teoría requieren también que las poblaciones que rindieron las distintas formas sean equivalentes, adicionalmente al requisito de compartir al menos un 20% de ítems en común.

Uno de los problemas que se presenta en admisión 2020 radica en la equiparación de los puntajes, dado que hay que considerar no sólo las diferentes formas de las pruebas, sino también las diferentes aplicaciones de ésta. En este caso, las diferentes aplicaciones no compartieron la cantidad suficiente de ítems y, adicionalmente, las poblaciones que participaron en las distintas aplicaciones son diferentes (Ver capítulo 2). La alternativa para abordar este problema de equiparación es usar modelos de Teoría de Respuesta al Ítem considerando la información que se obtiene en los proceso de pilotaje, particularmente la dificultades de los ítems (Ver Capítulo 5).

DEMRE, todos los años realiza una aplicación piloto con el fin de probar ítems que serán ensamblados en futuras pruebas del proceso oficial. Se pilotean los ítems en una muestra intencionada de estudiantes, representativa de la población de recién egresados que rinde finalmente las pruebas, con el propósito de saber si los ítems son pertinentes a la población objetivo y estimar sus parámetros psicométricos.

El gráfico 4.1 muestra cómo son ensambladas (construidas) las pruebas oficiales a partir de varias aplicaciones pilotos. Por ejemplo, la PSU 2018 fue construida con ítems del piloto 2016 y de otros pilotos anteriores, y las pruebas de la admisión 2020 con ítems de pilotos 2016, 2017 y 2018.

Figure 4.1: *Esquema de pilotos*



La dificultad es una propiedad psicométrica del ítem que refiere a cuánto conocimiento o habilidad debe tener el examinado para responder correctamente. Cuanto mayor sea el conocimiento necesario para responder correctamente el ítem, mayor será su grado de dificultad. Existen varios métodos para cuantificar este estadístico, que dependerán del enfoque metodológico utilizado. En el caso de Teoría Clásica del Test, la dificultad se estima como la proporción de examinados que responden correctamente el ítem, por lo cual es muy sensible a la población que rinde el test y a las condiciones de rendición.

La PSU es una prueba de alta consecuencia que se rinde bajo estrés, mientras que las aplicaciones piloto no tiene consecuencia para los examinados, por lo que se espera que las dificultades de un mismo ítem aplicado en pilotos y prueba oficial varien. Por ello, se hace necesario estudiar su relación.

4.1 Comparación de Dificultades¹

Una forma de establecer una relación entre las dificultades obtenidas en el piloto y en la prueba oficial es probando el ajuste de un modelo de regresión lineal.

4.1.1 Modelo de Regresión Lineal

Un modelo de regresión lineal trata de explicar la relación que existe entre una variable dependiente (variable respuesta, y) y una variable independiente (x) o un conjunto de variables independientes (x_1, x_2, \dots, x_n).

En este contexto, se trata de explicar y predecir las dificultades de los ítems en la prueba oficial 2020 (variable dependiente y) a partir de las dificultades de los ítems obtenidas en los pilotos en que participaron (variable independiente x). Así, el modelo de regresión lineal tiene la siguiente expresión:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Donde,

- y_i corresponde a la dificultad del ítem i en la prueba oficial 2020.
- x_i corresponde a la dificultad del ítem i en el piloto.
- β_0 es el intercepto de la regresión, lo que corresponde al valor promedio que tienen las dificultades de los ítems de la prueba oficial 2020 si las dificultades de los ítems en los pilotos fueran 0.
- β_1 corresponde a la pendiente de la recta de regresión, que representa el aumento de la dificultad en la prueba oficial al aumentar en una unidad la dificultad en el piloto.
- ϵ_i representa el error aleatorio asociado a x_i , donde se supone que $\epsilon_i \sim N(0, \sigma^2)$, tiene una distribución normal con media 0 y varianza constante σ^2 y que ϵ_i con ϵ_j son independientes cuando $i \neq j$.

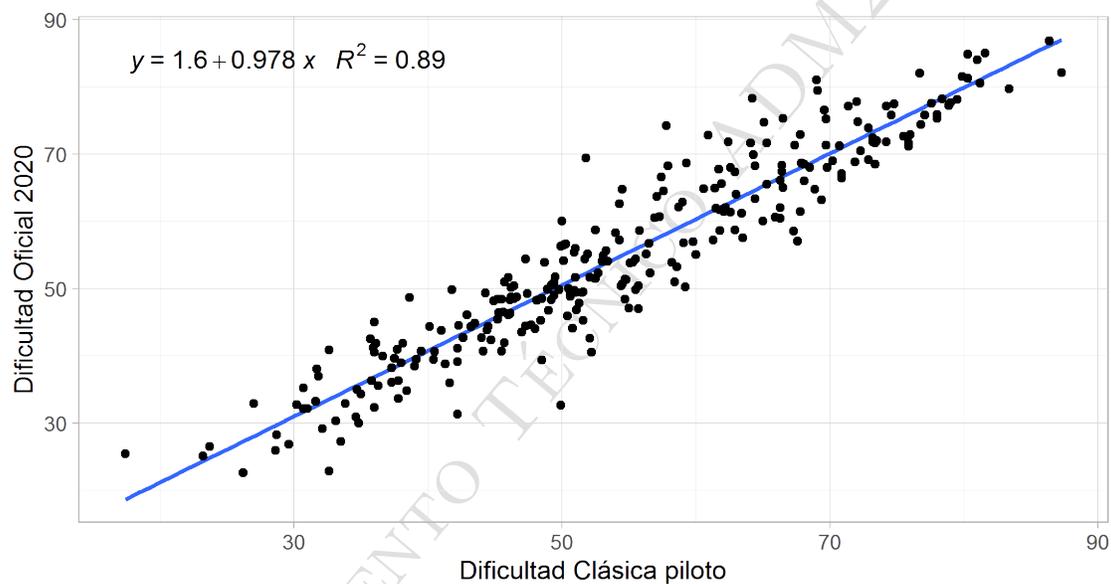
El ajuste del modelo se puede evaluar mediante el coeficiente de determinación R^2 , que se trata de una medida estandarizada que toma valores entre 0 y 1; un valor cercano a 1 indica un mejor ajuste del modelo.

¹Para la aplicación 1 se consideró sólo las respuestas de postulantes que rindieron sin complicación, es decir, que no fueron habilitados para las aplicaciones siguientes. Se aplicó este criterio porque existieron hojas de respuestas de postulantes que solo alcanzaron a poner el nombre, o se debió suspender durante la aplicación por diversos motivos de seguridad.

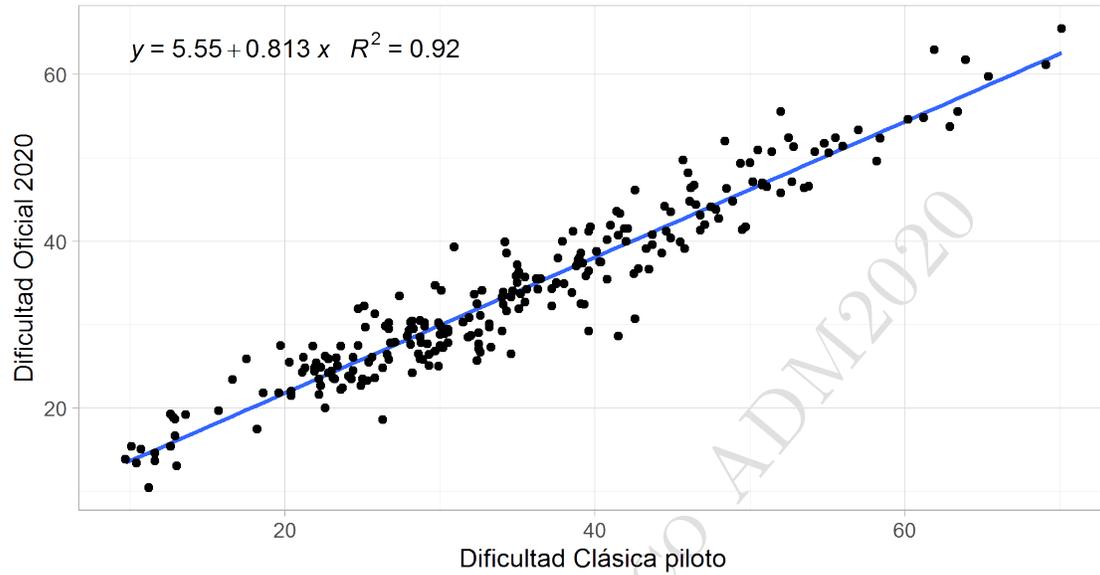
De esta manera, se analizó mediante modelos lineales la asociación de las dificultades obtenidas bajo Teoría Clásica del Test en los pilotos y en la prueba oficial. En las siguientes tres figuras se muestra la relación que presentaron las dificultades obtenidas en la prueba oficial 2020 versus las dificultades obtenidas en el proceso de pilotaje, donde en las tres pruebas observamos buen ajuste con $R^2 = 0.89$ en Lenguaje (Ver Figura 4.2), $R^2 = 0.92$ en Matemática (Ver Figura 4.3) y $R^2 = 0.72$ en Ciencias (Ver Figura 4.4).

Lenguaje

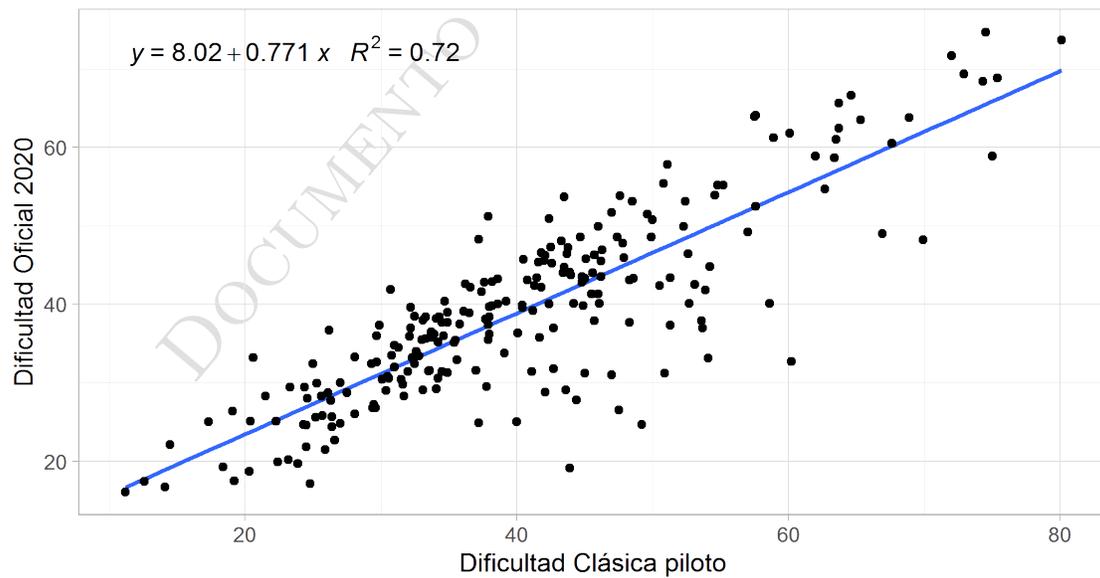
Figure 4.2: *Relación de las dificultades obtenidas con Teoría Clásica en Lenguaje*



Matemática

Figure 4.3: *Relación de las dificultades obtenidas con Teoría Clásica en Matemática*

Ciencias

Figure 4.4: *Relación de las dificultades obtenidas con Teoría Clásica en Ciencias*

Si bien las dificultades obtenidas por Teoría Clásica del Test se pueden ajustar bien entre el piloto y la prueba oficial, esto **no es suficiente para resolver la complejidad de este año en particular, que radica en la necesidad de equiparar las formas en las distintas aplicaciones del proceso de admisión 2020.**

4.2 Confiabilidad

La Teoría Clásica conceptualiza las puntuaciones como la suma del “puntaje verdadero” más un error de medida. Uno de los objetivos respecto a la asignación de puntajes en la PSU, es determinar con la mayor precisión posible el puntaje verdadero de cada individuo y, por lo tanto, minimizar el error.

La confiabilidad de una prueba es un concepto que indica el grado en que las puntuaciones observadas pueden ser consideradas como puntuaciones verdaderas. El indicador comúnmente utilizado se denomina coeficiente Alfa de Cronbach, que es un indicador del grado en que fluctúan los puntajes entre formas equivalentes de una misma prueba. La fórmula para la obtención del Alfa de Cronbach se muestra en la siguiente ecuación,

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_T^2} \right)$$

Donde

- n : es el número de ítems
- $\sum_{i=1}^n \sigma_i^2$: es la suma de las varianzas de los n ítems.
- σ_T^2 : es la varianza de las puntuaciones de test.

Valores superiores a 0.9 para el indicador Alfa de Cronbach son considerados, tanto a nivel nacional como internacional, como altamente satisfactorios. Es decir, el puntaje obtenido en el test por los individuos se acerca con precisión a su puntaje verdadero.

En las siguientes secciones se muestra una tabla por prueba de la confiabilidad de cada instrumento, α , de cada una de las formas, donde $\alpha_{2.5}$ y $\alpha_{97.5}$ corresponde a el percentil 2.5 y 97.5 de las estimaciones de confiabilidad obtenidas a través del método Bootstrap.

4.2.1 Lenguaje

Table 4.1: *Confiabilidad de las formas de Lenguaje*

Formas	Aplicación	α	$\alpha_{2.5}$	$\alpha_{97.5}$
101	Aplicación 1	0.920	0.917	0.918
102	Aplicación 1	0.910	0.913	0.915
103	Aplicación 1	0.910	0.911	0.913
104	Aplicación 1	0.920	0.922	0.924
105	Aplicación 2	0.910	0.904	0.907
106	Aplicación 2	0.910	0.908	0.911
107	Aplicación 3	0.910	0.901	0.911

4.2.2 Matemática

Table 4.2: *Confiabilidad de las formas de Matemática*

Formas	Aplicación	α	$\alpha_{2.5}$	$\alpha_{97.5}$
111	Aplicación 1	0.920	0.920	0.924
112	Aplicación 1	0.920	0.916	0.920
113	Aplicación 1	0.920	0.916	0.919
114	Aplicación 1	0.930	0.925	0.928
115	Aplicación 2	0.920	0.917	0.921
116	Aplicación 2	0.920	0.915	0.919
117	Aplicación 3	0.920	0.916	0.925

4.2.3 Ciencias

Table 4.3: *Confiabilidad de las formas de Ciencias*

Formas	Módulo	Aplicación	α	$\alpha_{2.5}$	$\alpha_{97.5}$
151	BIO	Aplicación 1	0.940	0.935	0.938
152	BIO	Aplicación 1	0.940	0.936	0.938
153	BIO	Aplicación 2	0.920	0.917	0.923
154	BIO	Aplicación 2	0.920	0.917	0.924
161	FIS	Aplicación 1	0.940	0.943	0.946
162	FIS	Aplicación 1	0.940	0.942	0.945
163	FIS	Aplicación 2	0.920	0.919	0.928
164	FIS	Aplicación 2	0.930	0.922	0.930
171	QUI	Aplicación 1	0.950	0.947	0.950
172	QUI	Aplicación 1	0.950	0.946	0.949
173	QUI	Aplicación 2	0.940	0.935	0.942
174	QUI	Aplicación 2	0.940	0.930	0.938
181	TP	Aplicación 1	0.840	0.837	0.851
182	TP	Aplicación 1	0.840	0.832	0.846
183	TP	Aplicación 2	0.770	0.744	0.788
184	TP	Aplicación 2	0.790	0.766	0.809

Las Confiabilidades obtenidas en cada una de las formas de las pruebas de este año se mantienen en rangos similares a las obtenidas en la de admisión 2019, donde la confiabilidad de la prueba de Matemática varió entre 0.919 y 0.921 en Lenguaje entre 0.908 y 0.912. Las confiabilidades de las pruebas de ciencias Biología, Física y Química variaron entre 0.901 y 0.925, mientras que en ciencias TP su variación estuvo entre 0.746 y 0.752.

Capítulo 5

Análisis con Teoría de Respuesta al Ítem

Como se mencionó en el capítulo anterior, el problema que no se puede resolver mediante Teoría Clásica del Test es el de equiparación entre las distintas aplicaciones de la PSU 2020, por lo cual se hace necesario utilizar un modelo de Teoría de Respuesta al Ítem.

Antes se explicó la imposibilidad de usar Teoría Clásica de Test en términos técnicos, referidos al incumplimiento de supuestos subyacentes. Sin embargo, la principal particularidad del problema que se enfrenta es que la población que rinde dos y hasta tres aplicaciones no solo es distinta de la que rinde solo una, sino que es una población que enfrenta mayores dificultades en su rendición. Una de las propiedades de este tipo de modelos es que tanto las habilidades de los examinados como las dificultades de los ítems son estimados en la misma escala y que la estimación de las habilidades de los examinados depende de la estimación de las dificultades de los ítems. De este modo, si se logra realizar una estimación de las dificultades de los ítems que use información histórica de ellos, que los libere del efecto de las particulares circunstancias de estas aplicaciones, se podrá tener una estimación más justa y precisa de las habilidades de los postulantes.

El modelo que se decidió usar fue el modelo de Rasch por su simplicidad, por ser parsimonioso ya que considera de igual manera cada uno de los ítems (preguntas) para el cálculo del puntaje, y es uno de los modelos de Teoría de Respuesta al Ítem cuyos resultados se asemejan en mayor medida a los obtenidos con el método usado en los procesos anteriores.

5.1 Modelo de Rasch

El modelo de Rasch al igual que otros modelos, se enmarca en la Teoría de Respuesta al Ítem, TRI en español o IRT en inglés. En éste la probabilidad de responder correctamente una pregunta se modela a través de una función de los parámetros de persona y de ítems. De esta manera, si X_{ij} es una variable aleatoria definida como 1 si el i -ésimo examinado contesta correctamente el j -ésimo

ítem y 0 si no, entonces, se define la probabilidad, p_{ij} , de que una persona i responda correctamente el ítem j , como:

$$p_{ij} = p(X_{ij} = 1 | \theta_i, \delta_j) = \frac{e^{(\theta_i - \delta_j)}}{1 + e^{(\theta_i - \delta_j)}} \quad (5.1)$$

Donde,

- $i \in \{1, 2, \dots, M\}$, donde M es la cantidad de postulantes.
- $j \in \{1, 2, \dots, L\}$, donde L es la cantidad de ítems.
- θ_i es la habilidad de la i -ésima persona.
- δ_j es la dificultad del j -ésimo ítem.

Comunmente, en las pruebas educativas los parámetros de los ítems representan la dificultad de estos, mientras que los parámetros de las personas representan la capacidad o el nivel de logro de las personas evaluadas. Cuanto mayor sea la capacidad de una persona en relación con la dificultad de un ítem, mayor será la probabilidad de una respuesta correcta en ese ítem. Dicho de otra manera, si $\theta_i = \delta_j$ en ecuación (5.1), entonces $p_{ij} = \frac{1}{2}$, mientras mayor sea el valor del parámetro de dificultad del ítem, mayor habilidad es requerida para poder contestarlo de manera correcta. Es decir, para cierta habilidad θ ,

$$\text{si } \delta_1 > \delta_2 \Rightarrow p(X_{.1} = 1 | \delta_1, \theta) < p(X_{.2} = 1 | \delta_2, \theta) \quad (5.2)$$

5.1.1 Método de Estimación

Estimación de las dificultades

Sea \tilde{x}_i el vector de respuesta de un postulante, entonces tenemos que la probabilidad de que se haya observado ese patrón de respuestas correctas dada la habilidad (θ_i) de un estudiante i y el vector de dificultad de los items, $\tilde{\Delta} = (\delta_1, \delta_2, \dots, \delta_L)$ es:

$$p(\tilde{x}_i|\theta_i, \tilde{\Delta}) = \prod_{j=1}^L p_{ij}^{x_{ij}} (1 - p_{ij})^{1-x_{ij}} \quad (5.3)$$

De esta forma, podemos obtener la probabilidad marginal de obtener el vector de respuesta que viene dada por:

$$p(\tilde{x}) = \int_{-\infty}^{\infty} p(\tilde{x}|\theta, \tilde{\Delta})g(\theta|\tilde{v})d\theta$$

donde $p(\tilde{x})$ es la función de verosimilitud del vector de respuestas definido en (5.3) y $g(\theta|\tilde{v})$ representa la función de probabilidad de las habilidades, siendo \tilde{v} el vector de parámetros de las habilidades (localización y escala) que se consideraron libre en la estimación.

Utilizando un método iterativo que combina sucesivas aproximaciones de la habilidad es posible obtener las estimaciones para el vector de dificultades.

Estimación de las habilidades

La estimación de la habilidad a través de EAP (Expected a Posteriori) esta dada por la esperanza a posteriori:

$$\hat{\theta}_i = E(\theta|\tilde{x}) = \frac{\int_{-\infty}^{\infty} \theta p(\tilde{x}|\theta)g(\theta)d\theta}{\int_{-\infty}^{\infty} p(\tilde{x}|\theta)g(\theta)d\theta} \quad (5.4)$$

donde, $p(\tilde{x}|\theta)$ es la función de verosimilitud (ver ecuación 5.1) y $g(\theta)$ es la densidad a priori de las habilidades. En otras palabras, se busca el valor de θ que maximiza la probabilidad de haber observado el vector de respuestas que cada postulante obtuvo.

5.1.2 Supuestos del modelo Rasch

1. Monotonicidad: Este supuesto corresponde a que la probabilidad de dar una respuesta correcta a un ítem es una función no decreciente del fenómeno latente que se está midiendo, es decir, de la habilidad en nuestro caso.
2. Independencia local del ítem: Este supuesto hace referencia a que la probabilidad de que un examinado responda correctamente a un ítem determinado dada la habilidad es independiente de los demás ítems del test.
3. Unidimensionalidad: Este supuesto se refiere a que existe un único factor que explica la respuestas de los examinados, o en otras palabras, una única dimensión o variable latente que se pretende “cuantificar” a partir del conjunto de ítems seleccionados en el test.

DOCUMENTO TÉCNICO ADM2020

5.2 Resultado de la estimación de dificultades

Como se mencionó en el capítulo anterior, la complicación de este proceso es la de medir las habilidades de los postulantes que rindieron las pruebas de modo que no dependa de cual aplicación rindió. Una manera de que las habilidades que se estimen en las tres aplicaciones sean equivalentes es a través de una calibración común de las dificultades de los ítems que permite obtener sus estimaciones en la misma escala, lo que implica que las habilidades de los postulantes sean comparables. Así, mientras más estables sea la estimación de las dificultades de los ítems mejor será la estimación de las habilidades.

5.2.1 Calibración sobre las dificultades de los ítems

Una calibración es el proceso por el cual se obtienen estimaciones de parámetros de ítems, en este caso dificultad, en una escala común mediante un proceso de equiparación. La información disponible de los ítems que fueron ensamblados en las pruebas oficial 2020 es la obtenida en los procesos de pilotaje, pero sabemos que estos son de baja consecuencia y no consideran a toda la población que rinde la prueba oficial. Por este motivo, en esta calibración se incluyó la información de ítems que estuvieron tanto en procesos de pilotajes como en pruebas oficiales. Es importante señalar que las pruebas de los procesos oficiales no comparten ítems entre años (Ver Figura 5.1)¹

Esta calibración se llevó a cabo con la información de los procesos de admisión 2019, 2018, 2017 y 2016, más la información de los pilotos 2018, 2017 y 2016².

Para la calibración se eliminaron algunos ítems que no tenían un buen ajuste, por lo cual existen ítems ensamblados en las pruebas oficiales 2020 cuyas dificultades no fueron estimados en esta calibración. Aquellos ítems cuya dificultades fueron estimadas en esta calibración se denominan ítems anclas.

El anclaje inicial para Matemática fueron 217 ítems, en Lenguaje 276, mientras que en Ciencias fue de 261 ítems. De esta manera, el porcentaje de anclas inicial más bajo corresponde a Matemática con un 73.3% del total, seguido por Ciencias con un 83.1% y en Lenguaje se tiene el mayor porcentaje de anclaje con un 97.9% (Ver Tabla 5.1).

5.2.2 Análisis de regresión sobre las dificultades de los ítems

Se realizó un análisis de regresión lineal para relacionar las dificultades estimadas de los ítems desde la calibración a las estimaciones obtenidas sólo con la información de la prueba oficial 2020.

¹Los colores blancos del esquema representa ítems que perteneces sólo al proceso.

²Este trabajo fue parcialmente apoyado por la infraestructura de supercómputo del NLHPC (ECM-02)

Figure 5.1: Esquema de Calibración**Table 5.1:** Ítems ensamblados por prueba

Prueba	Número de Formas inicial	Número de Ítem Total	Anclaje	Porcentaje de anclaje
Matemática	7	296	217	73.3%
Lenguaje	7	282	276	97.9%
Ciencias	16	314	261	83.1%

Para cada una de las tres pruebas se buscó establecer los parámetros β_0 y β_1 de la relación lineal $\hat{\delta}_i^{(o)} = \beta_0 + \beta_1 \hat{\delta}_i^{(c)}$ que mejor ajuste a los datos $\hat{\delta}_i^{(c)}, \hat{\delta}_i^{(o)}$, donde $\hat{\delta}_i^{(c)}$ es la dificultad del ítem i en la calibración concurrente y $\hat{\delta}_i^{(o)}$ es la dificultad del ítem i en la prueba oficial 2020.

La calidad del ajuste se evaluó mediante el coeficiente de determinación (R^2) y adicionalmente se analizaron los residuos del modelo para verificar que sus supuestos se cumplan y que no existan valores atípicos que pueden influir de manera no deseada a la estimación final de las habilidades. A continuación se describen los resultados por prueba.

Lenguaje

El ajuste inicial de Lenguaje fue de un $R^2 = 0.88$. Al estudiar los residuos se identificaron 15 valores atípico (Ver Figura 5.2). Se decidió desanclar esos 15 ítems obteniendo un $R^2 = 0.91$ (Ver Figura 5.3), con un porcentaje de anclaje final del 92.6% (Ver Tabla 5.5).

Posteriormente se analizaron los supuestos del modelo de regresión lineal, cumpliendo con supuestos de independencia, normalidad y homocedasticidad de los residuos (Ver Tabla 5.6).

Figure 5.2: *Dificultades de Lenguaje, con valores atípicos*

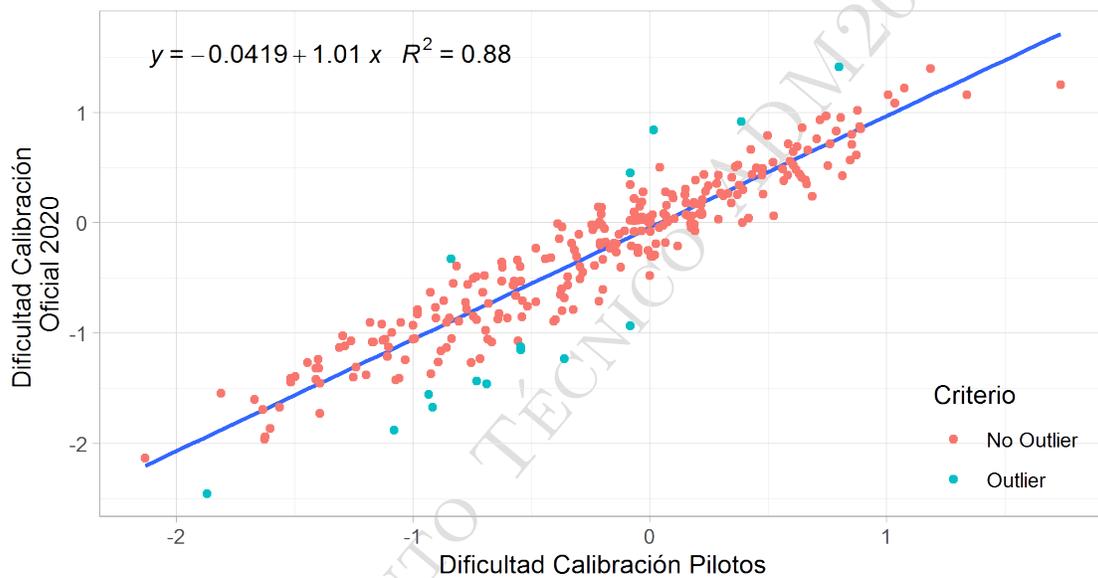
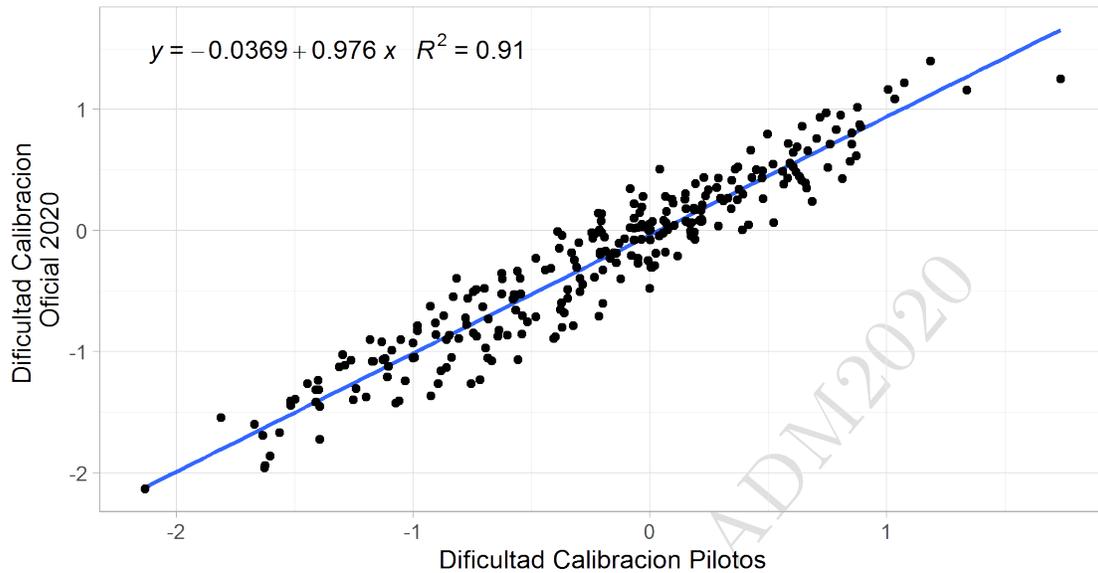
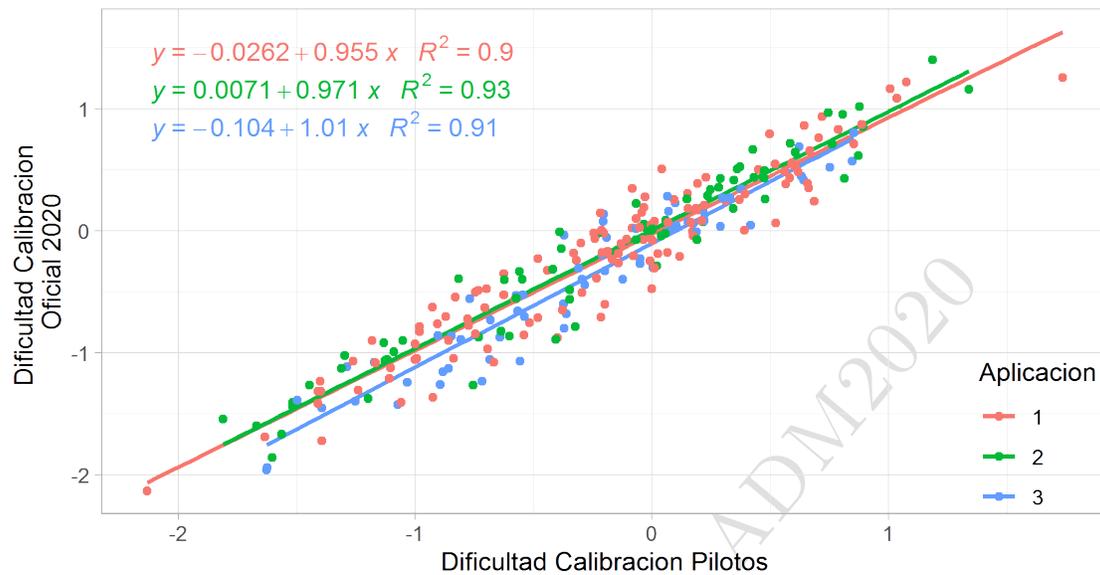


Figure 5.3: *Dificultades de Lenguaje, sin valores atípicos*

Por otra parte, se estudió la pendiente por aplicación (Ver Figura 5.4) calculando sus respectivos intervalos de confianza (Ver Tabla 5.2) para verificar que el estimador de la pendiente de la regresión general ($\hat{y} = -0.0369 + 0.976x$) pertenece a los intervalos de confianza de cada aplicación, concluyendo que se ha logrado un ajuste aceptable en Lenguaje con una pendiente única, lo que implica que las tres aplicaciones estarán en la misma escala.

Table 5.2: *Intervalo de confianza de los estimadores de regresión por aplicación en Lenguaje*

Parámetro	Aplicación 1	Aplicación 2	Aplicación 3
$\hat{\beta}_0$	-0.026	0.007	-0.104
$IC(\beta_0, 95\%)$	(-0.065 ; 0.013)	(-0.043 ; 0.057)	(-0.159 ; -0.049)
$\hat{\beta}_1$	0.955	0.971	1.015
$IC(\beta_1, 95\%)$	(0.899 ; 1.011)	(0.908 ; 1.033)	(0.935 ; 1.095)

Figure 5.4: *Dificultades en Lenguaje por aplicación, sin valores atípicos*

Matemática

El ajuste inicial de Matemática fue de un $R^2 = 0.92$. Al estudiar los residuos se identificaron 14 valores atípicos (Ver Figura 5.5). Se decidió desanclar esos 14 ítems obteniendo un $R^2 = 0.94$ (Ver Figura 5.6), con un porcentaje de anclaje final del 68.6% (Ver Tabla 5.5).

Posteriormente se analizaron los supuestos del modelo de regresión lineal, cumpliendo los supuestos de independencia, normalidad y homocedasticidad de los residuos (Ver Tabla 5.6).

Figure 5.5: *Dificultades de Matemática, con valores atípicos*

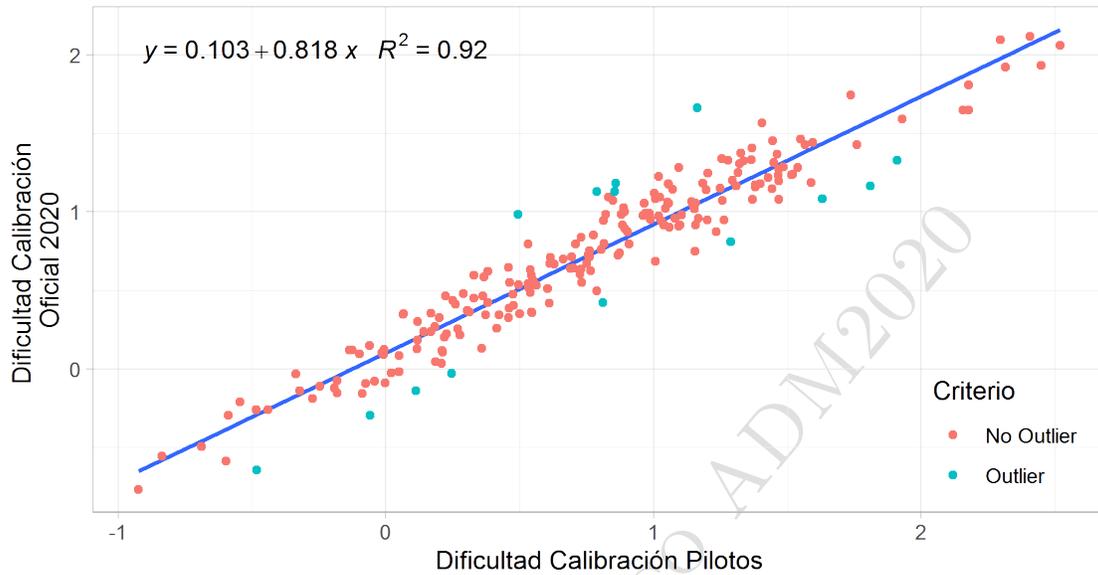
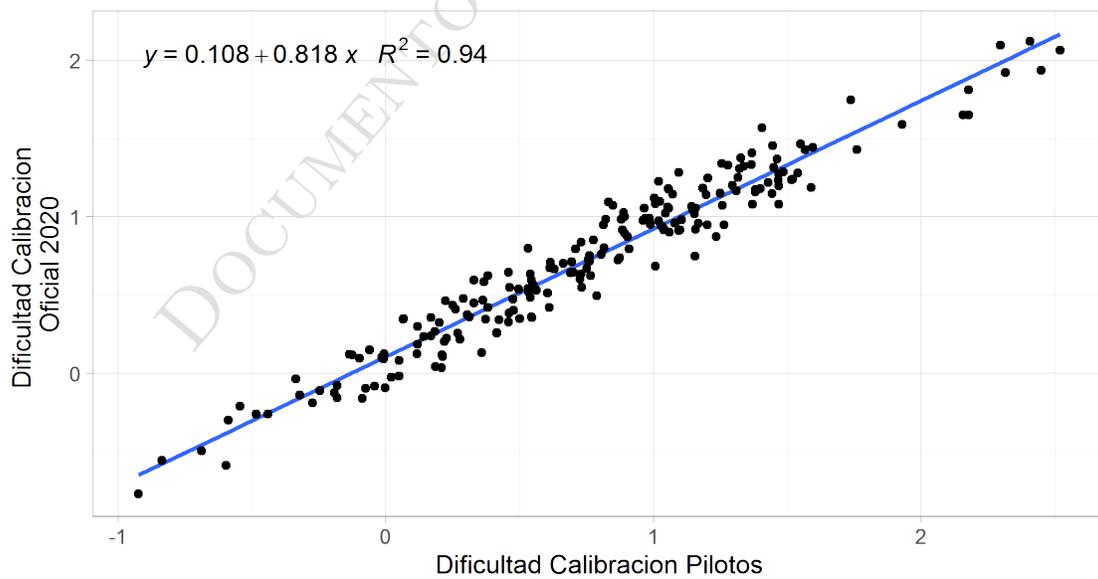


Figure 5.6: *Dificultades de Matemática, sin valores atípicos*

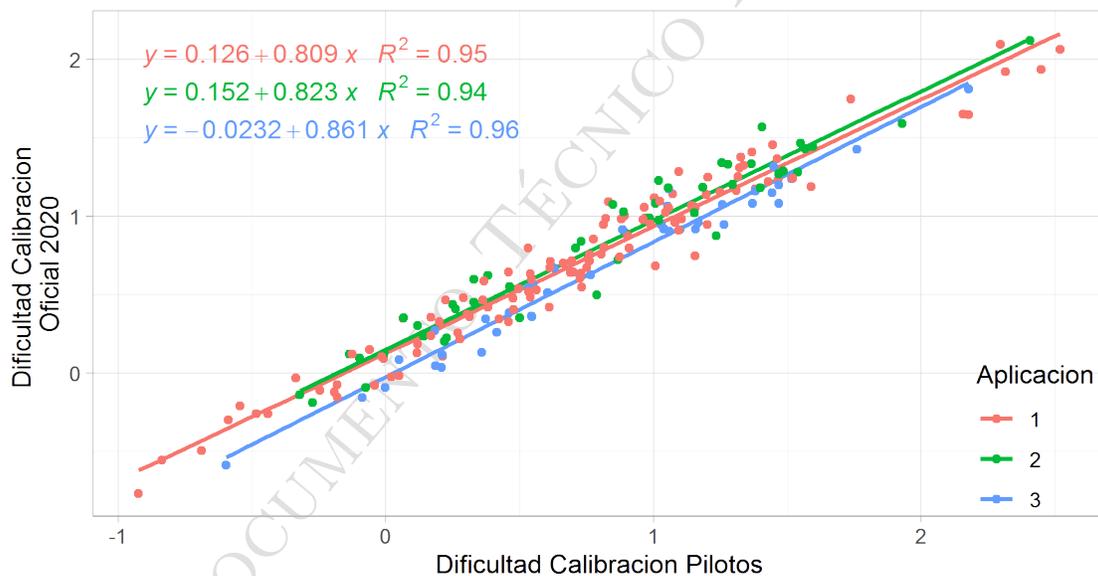


Por otra parte, se estudió la pendiente por aplicación (Ver Figura 5.7) calculando sus respectivos intervalos de confianza (Ver Tabla 5.3) para verificar que el estimador de la pendiente de la regresión general ($\hat{y} = 0.108 + 0.818x$) pertenece a los intervalos de confianza de cada aplicación, concluyendo que se ha logrado un ajuste aceptable en Matemática con una pendiente única, lo que implica que las tres aplicaciones están en la misma escala.

Table 5.3: Intervalo de confianza de los estimadores de regresión por aplicación en Matemática

Parámetro	Aplicación 1	Aplicación 2	Aplicación 3
$\hat{\beta}_0$	0.126	0.152	-0.023
$IC(\beta_0, 95\%)$	(0.094 ; 0.159)	(0.087 ; 0.216)	(-0.080 ; 0.034)
$\hat{\beta}_1$	0.809	0.823	0.861
$IC(\beta_1, 95\%)$	(0.776 ; 0.842)	(0.760 ; 0.886)	(0.804 ; 0.917)

Figure 5.7: Dificultades en Matemática por aplicación, sin valores atípicos



Ciencias

El ajuste inicial de Ciencias fue de un $R^2 = 0.86$. Al estudiar los residuos se identificaron 16 valores atípicos (Ver Figura 5.8). Se decidió desanclar esos 16 ítems obteniendo un $R^2 = 0.91$ (Ver Figura 5.9), con un porcentaje de anclaje final del 78.0% (Ver Tabla 5.5).

Posteriormente se analizaron los supuestos del modelo de regresión lineal, cumpliendo los supuestos de independencia, normalidad y homocedasticidad de los residuos (Ver Tabla 5.6).

Figure 5.8: *Dificultades de Ciencias, con datos atípicos*

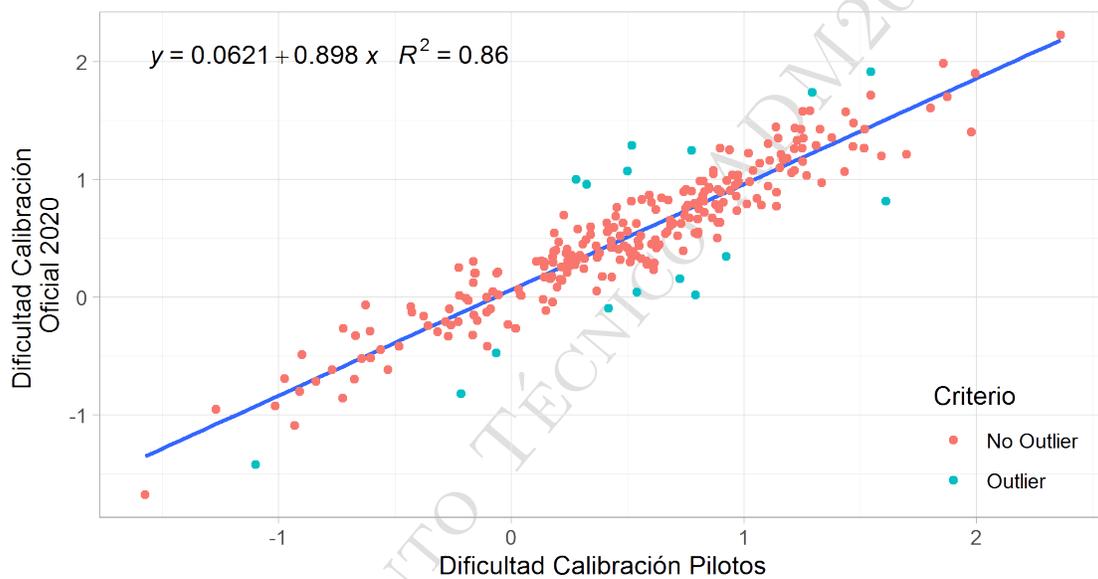
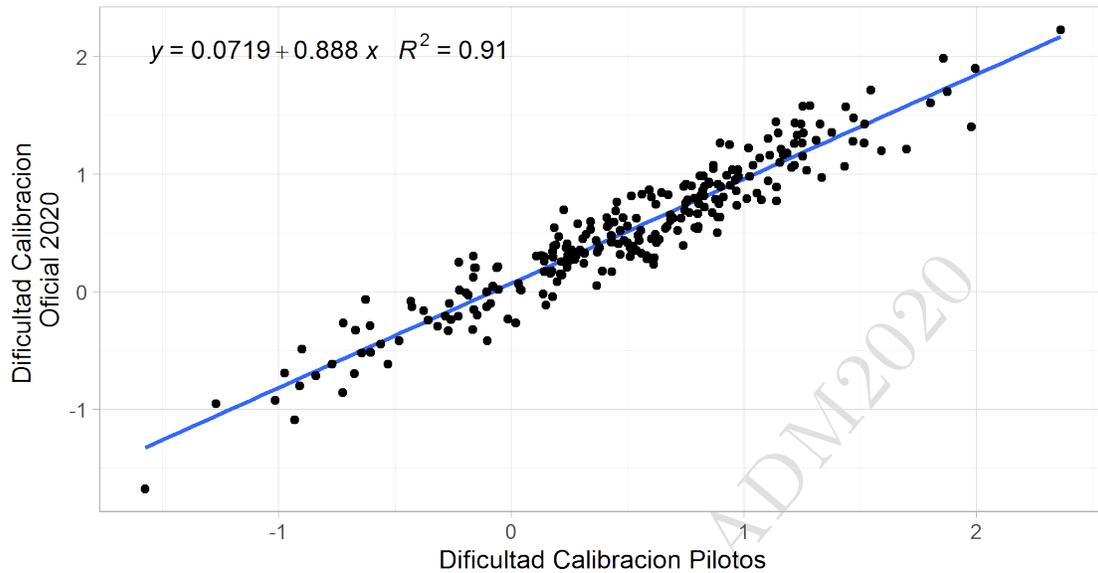


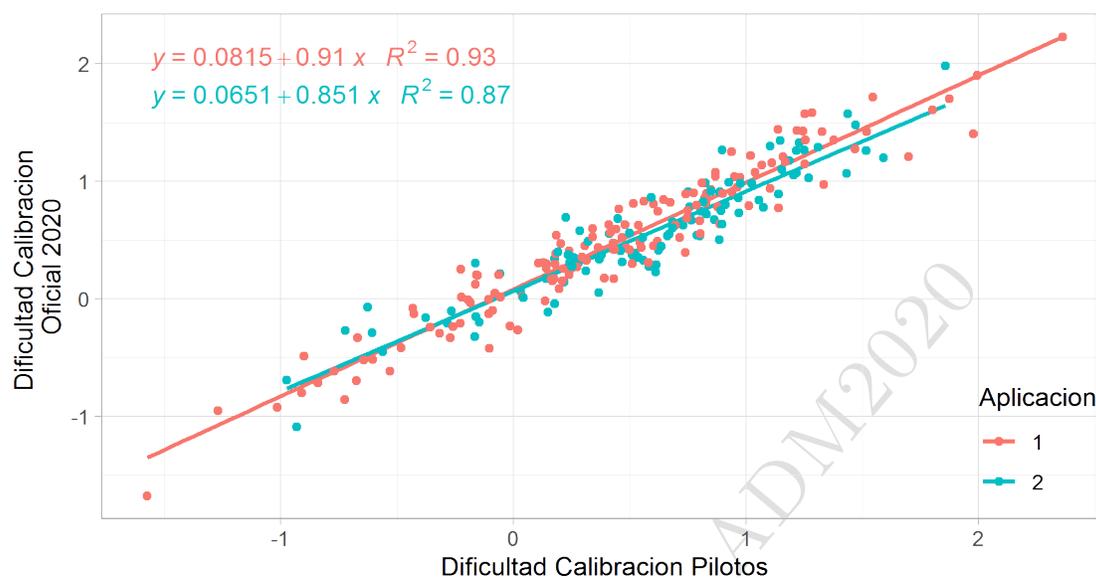
Figure 5.9: *Dificultades de Ciencias, sin datos atípicos*



Por otra parte, se estudió la pendiente por aplicación (Ver Figura 5.10) calculando sus respectivos intervalos de confianza (Ver Tabla 5.4) para verificar que el estimador de la pendiente de la regresión general ($\hat{y} = 0.0719 + 0.888x$) pertenece a los intervalos de confianza de cada aplicación, concluyendo que se ha logrado un ajuste aceptable en Ciencias con una pendiente única, lo que implica que las tres aplicaciones están en la misma escala.

Table 5.4: *Intervalo de confianza de los estimadores de regresión por aplicación en Ciencias*

Parámetro	Aplicación 1	Aplicación 2
$\hat{\beta}_0$	0.081	0.065
$IC(\beta_0, 95\%)$	(0.047 ;0.116)	(0.0086 ;0.1129)
$\hat{\beta}_1$	0.910	0.851
$IC(\beta_1, 95\%)$	(0.868 ;0.953)	(0.786 ;0.916)

Figure 5.10: *Dificultades en Ciencias por aplicación, sin valores atípicos*

Resumen

A continuación se presentan la tabla resumen del anclaje (Tabla 5.5) y el de ajuste de los modelos (Tabla 5.6).

Table 5.5: *Resumen del análisis de anclas*

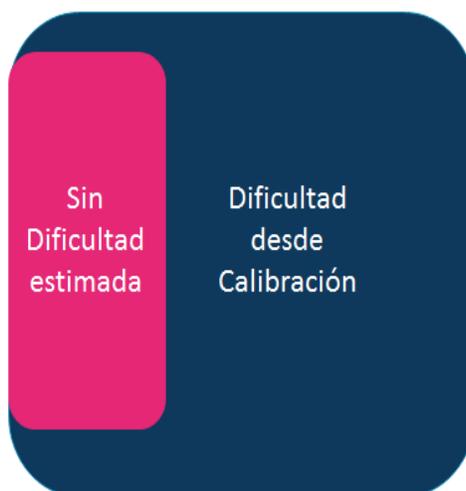
Prueba	Número de Formas	Número de Items	Numero de Anclaje	Porcentaje de anclaje inicial	Número de ítems desanclados	Número final anclas	Porcentaje de anclaje final	R^2
Matemática	7	296	217	73.3%	14	203	68.6%	0.94
Lenguaje	7	282	276	97.9%	15	261	92.6%	0.93
Ciencias	16	314	261	83.1%	16	245	78.0%	0.91

Table 5.6: *Resumen de verificación de supuestos de un modelo de regresión lineal*

Prueba	R^2	SCM	SCE	Independencia	Kolmogorov	Breusch
Matemática	0.94	56.57	3.41 (p=0.109)	DW=1.78 (p=0.91)	D=0.04 (p=0.15)	BP=2.07
Lenguaje	0.91	116.38	11.36 (p=0.95)	DW=2.00 (p=0.26)	D=0.06 (p=0.36)	BP=0.83
Ciencias	0.91	78.46	8.04 (p=0.77)	DW=1.97 (p=0.39)	D=0.05 (p=0.72)	BP=0.13

5.2.3 Estimación final de Dificultades de los ítems

Como resultado de estos procesos, existen ítems que no tienen su dificultad estimada a partir de la información histórica, las que deberán ser estimadas con otra metodología que utiliza en calidad de anclas las dificultades, y por ende ya establecidas.

Figure 5.11: *Dificultad Estimada Oficial 2020*

Para obtener las estimaciones de las dificultades faltantes se ajustó un modelo Rasch fijando las dificultades que se obtuvieron luego del análisis de regresión. Como resultado de este modelo se obtienen las estimaciones faltantes de los ítems que están en la misma escala que aquellos ítems cuya dificultad fue fijada en el procedimiento anterior, asegurando de esta manera que las habilidades estimadas en las tres aplicaciones están en la misma escala y que considera toda la información histórica disponible.

5.3 Obtención de Puntajes

5.3.1 Estimación de Habilidad

Con todas las dificultades estimadas de los ítems previamente, se procedió a estimar las habilidades de los examinados por medio de EAP (Ver 5.1.1) a todos los patrones de respuestas leídos. Dado que algunos postulantes tuvieron hasta 3 aplicaciones rendidas y de acuerdo a lo comunicado previamente se les asignó la mayor de las habilidades estimadas en cada una de las aplicaciones.

Además, a aquellos postulantes que alcanzaron el máximo posible en la forma de la prueba que rindieron, por ejemplo en Matemática el máximo se podía obtener con 72, 74 o 75 puntos dependiendo la forma (Ver Tabla 3.5), se les asignó a todos ellos una única habilidad máxima, habilidad que fue determinada según el promedio de habilidades de esos postulantes que alcanzaron el máximo posible de su forma rendida..

De esta misma forma se le asignó una única habilidad mínima a aquellos estudiantes que no obtuvieron respuestas correctas.

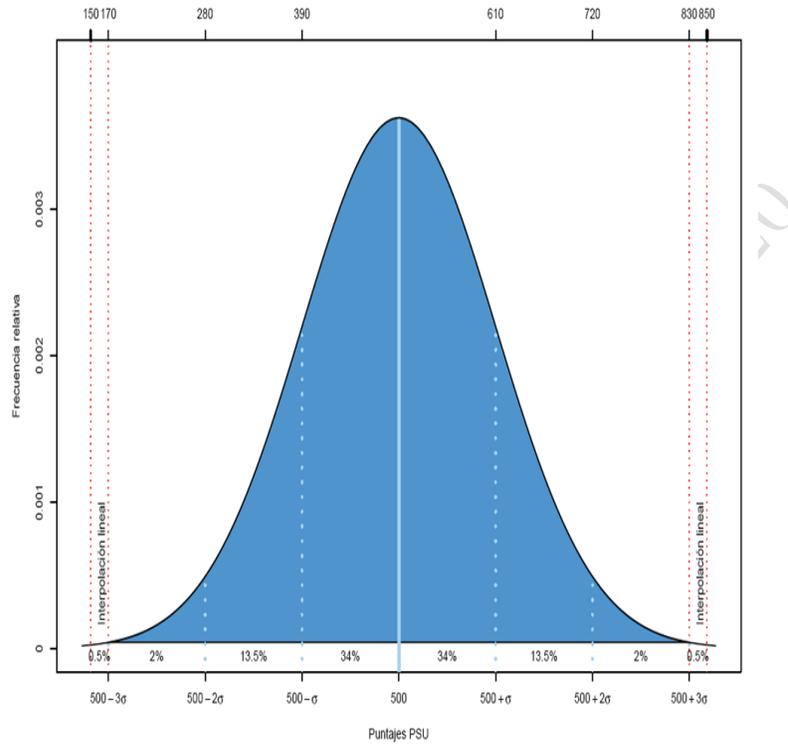
5.3.2 Transformación de Habilidad a Puntaje PSU

Se realizó una transformación de la distribución de las habilidades asignadas, que mantiene la posición relativa de los individuos, mediante un proceso de normalización que consiste en:

- Para cada habilidad se debe encontrar el percentil correspondiente.
- Realizar la corrección por continuidad que corresponde a sumar la mitad de la frecuencia relativa absoluta de la habilidad a la frecuencia relativa acumulada.
- Dado estos valores de frecuencias relativas acumuladas post corrección de continuidad (y), encontrar el valor de x tal que: $F(x) = y$, donde F es la función de densidad acumulada de una Normal($\mu=500, \sigma=110$).
- Aquellos puntajes corregidos cuyas frecuencias relativas acumuladas post corrección de continuidad se encuentren en el 0.5% más extremo se ajustan a través de una interpolación lineal. Esta interpolación considera la menor habilidad asignada como 150 puntos y la mayor como 850 puntos. Luego se considera el primer (o último) valor que no se interpola para realizar una recta con la cantidad de puntos que corresponda.
- Generar la tabla que lleva cada habilidad a puntaje en escala PSU (150-850) para luego asignarlos a cada individuo según corresponda.

La siguiente figura muestra un ejemplo de la distribución de puntajes normalizados.

Figure 5.12: *Ejemplo de distribución de puntaje normalizados*



Los gráficos de transformación de habilidad asignada a puntaje estándar PSU se presentan en los siguientes gráficos:

Figure 5.13: *Puntaje PSU v/s estimación de habilidad en Lenguaje*

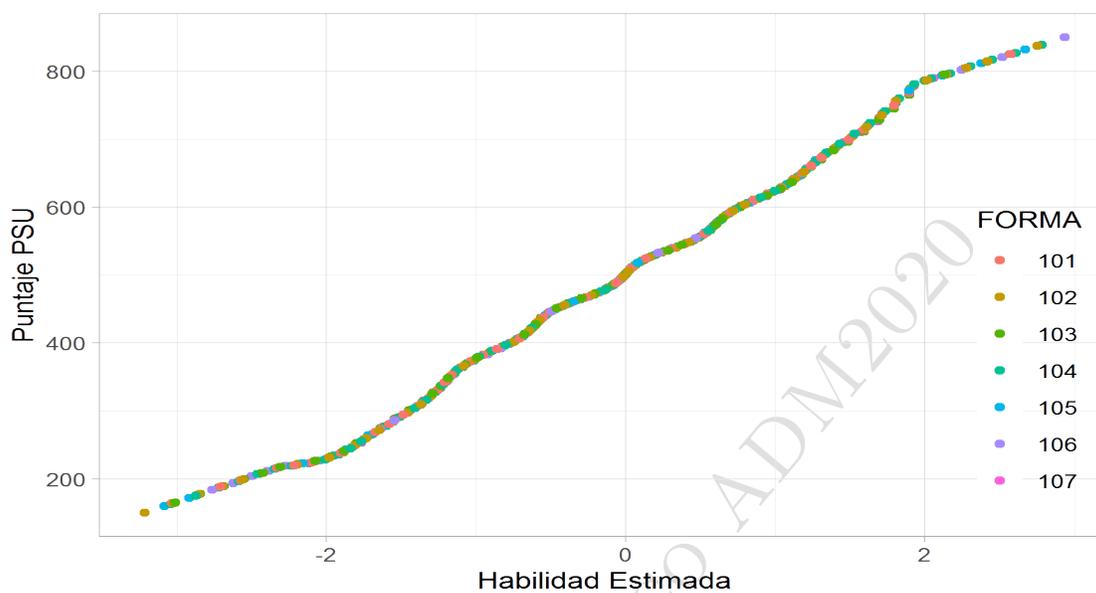


Figure 5.14: *Puntaje PSU v/s estimación habilidad en Matemática*

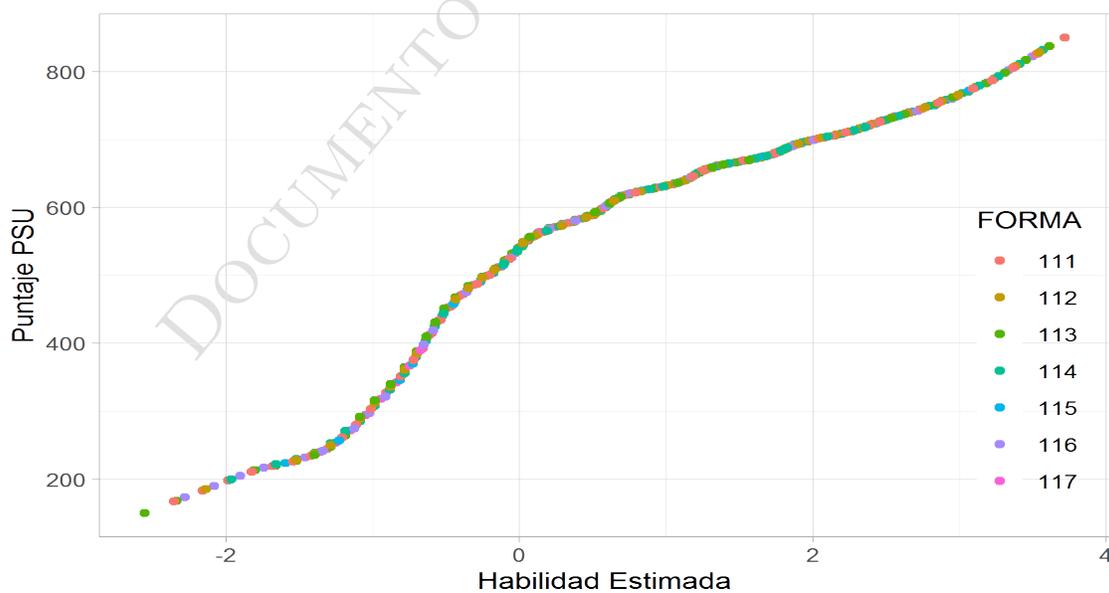
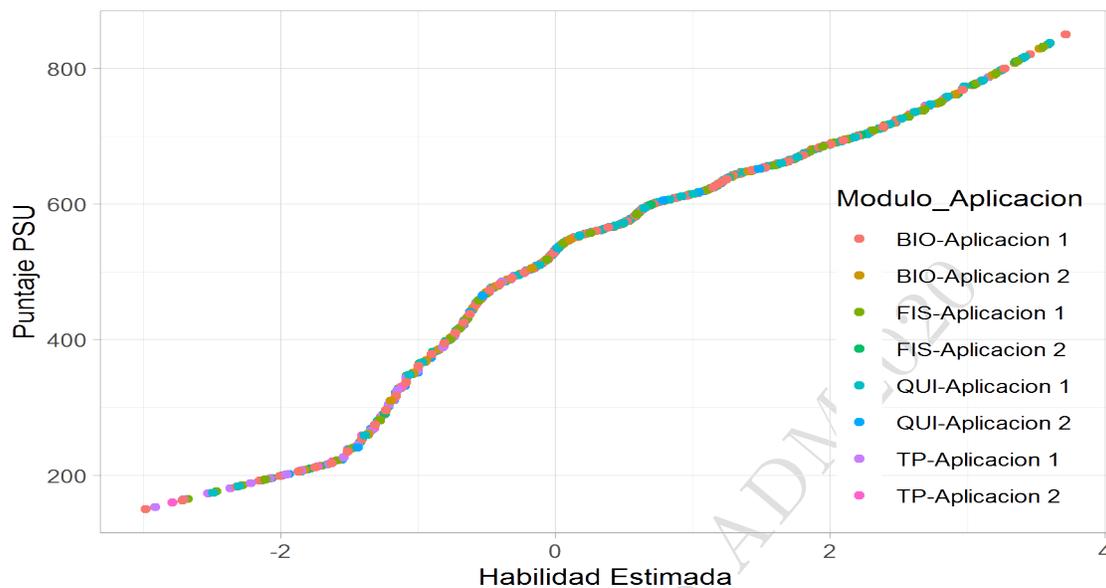


Figure 5.15: *Puntaje PSU v/s estimación habilidad en Ciencias*

5.4 Ajuste del modelo

Un aspecto importante a considerar es evaluar el ajuste de los modelos antes presentados. Si bien existen muchas pruebas estadísticas que permiten realizar esto, la mayoría son basadas en el estadístico chi-cuadrado que resulta ser muy sensible a la cantidad de datos con que se trabaja y, por lo tanto para este caso no permitiría evaluar adecuadamente el ajuste de los modelos.

Dado lo anterior, la manera en que se decidió evaluar el ajuste fue mediante el estadístico Infit de las habilidades estimadas para los sujetos, que a grandes rasgos corresponde a un estadístico de ajuste con información ponderada que se enfoca al comportamiento general de un ítem o persona, se calcula con el promedio de las desviaciones cuadráticas estandarizadas entre el desempeño observado y el esperado.

En general, se considera que un valor Infit entre 0.5 y 1.5 indica un buen ajuste, esto se debe a que un valor de este estadístico igual a 1 (uno) indica que el 100% de la varianza de los datos empíricos son explicados por el modelo. Por otro lado, valores menores a 0.5 y entre 1.5 y 2 son considerados como que no ajustan tan bien al modelo pero a su vez no distorsionan la medición, finalmente un Infit superior a 2 se considera un valor que podría distorsionar el análisis.

Considerando lo expresado en el párrafo anterior es esperable entonces que para asegurar un buen ajuste del modelo, todos los valores Infit se encuentren entre 0.5 y 1.5, esto ocurre en casi el 100%

de los casos para los 3 modelos ajustados (Lenguaje, Matemática y Ciencias), siendo la excepción los valores "extremos", es decir, aquellos que obtuvieron 0 respuestas correctas y quienes obtuvieron todo correcto.

En las tablas Tablas 5.7, 5.8 y 5.9 se muestra lo anteriormente expresado para las pruebas de Lenguaje, Matemática y Ciencias respectivamente.

Table 5.7: *Tabla Resumen Infit prueba de Lenguaje*

Rango	Total	Total Sin Extremos
<0.5	640	0
0.5-1.5	253189	253189

Table 5.8: *Tabla Resumen Infit prueba de Matemática*

Rango	Total	Total Sin Extremos
<0.5	740	0
0.5-1.5	243878	243878

Table 5.9: *Tabla Resumen Infit prueba de Ciencias*

Rango	Total	Total Sin Extremos
<0.5	193	0
0.5-1.5	151808	151808

5.5 Confiabilidad EAP

Las confiabilidades obtenidas por EAP en las tres pruebas son altamente satisfactorias, lo que implica que las predicciones individuales son confiables.

- Lenguaje 0.931
- Matemática 0.915
- Ciencias 0.934

DOCUMENTO TÉCNICO ADM2020

Capítulo 6

Resultados

6.1 Distribución de Puntaje por Prueba

Los gráficos 6.1, 6.2 y 6.3, muestran la distribución de puntajes en cada prueba rendida en los dos últimos procesos de admisión. En general, el histograma de puntaje en las pruebas rendidas en admisión 2020 se asemeja a la densidad estimada a diferencia del proceso de admisión 2019.

Figure 6.1: *Puntaje de Lenguaje Admisión 2020 y Admisión 2019*

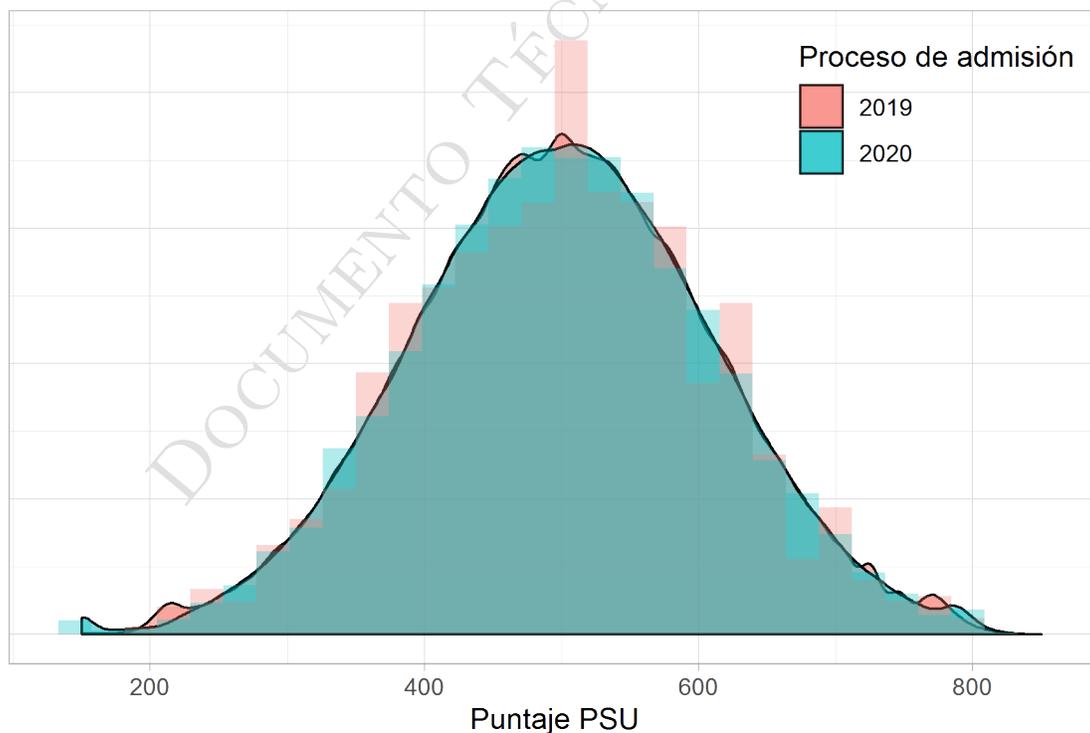


Figure 6.2: *Puntaje de Matemática Admisión 2020 y Admisión 2019*

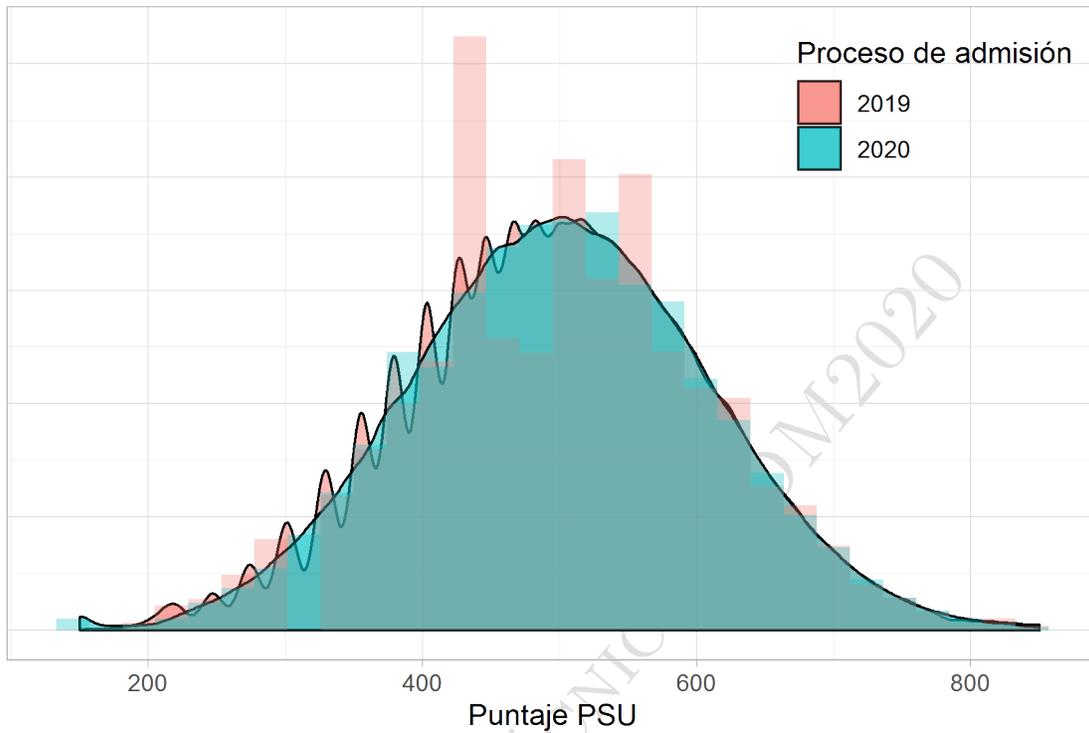
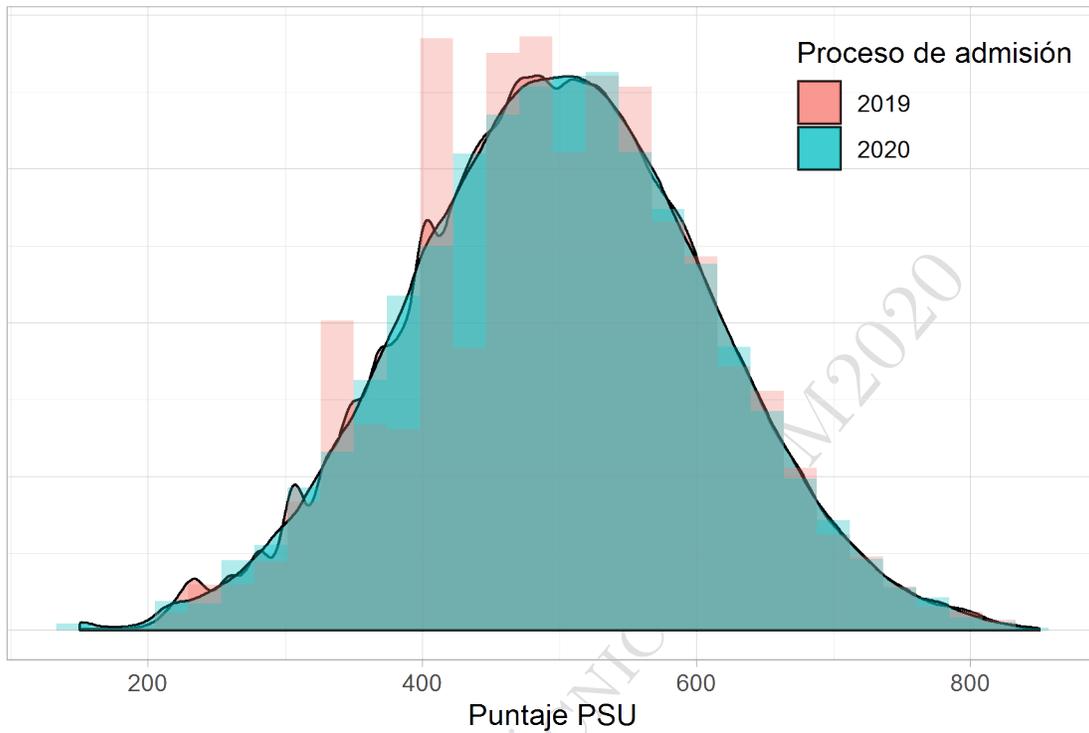


Figure 6.3: *Puntaje de Ciencias Admisión 2020 y Admisión 2019*



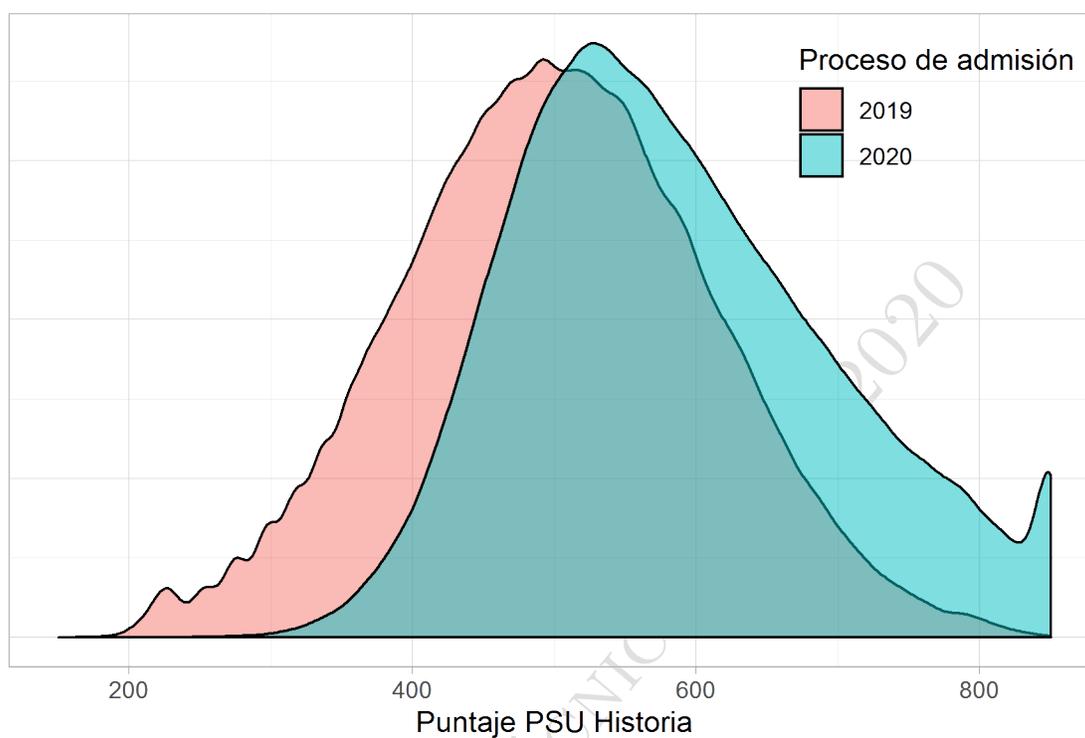
6.2 Puntaje de Historia

La PSU de Historia fue suspendida el día 7 de Enero por una filtración invalidante. Ante esta situación, el Consejo de Rectores (CRUCH) decidió que a los alumnos inscritos para esta prueba, por asignación automática, se les aplique el mejor puntaje obtenido en cualquiera de los otros factores, es decir, los mejores puntajes obtenidos en las otras pruebas, las notas de enseñanza media (NEM) o el Ranking. Las correlaciones de los puntajes de historia con estos factores se presentan en la tabla 6.1, donde el mejor predictor es Lenguaje con correlaciones cercanas al 0.8.

Table 6.1: *Correlaciones históricas de Historia*

Prueba	ADM2019	ADM2018	ADM2017	ADM2016
LENG	0.77	0.78	0.77	0.78
NEM	0.38	0.38	0.38	0.40
RANK	0.34	0.33	0.33	0.35

A 201.262 inscritos se le asignó el mejor factor a la prueba de historia. De este total, el 86% se presentó a rendir alguna prueba, es decir, 173.158 inscritos. Al 67.94% de los inscritos en Historia se les asignó el puntaje NEM o Ranking, y la distribución de Historia de este año no sigue una distribución Normal y se corre hacia la derecha (Ver Figura 6.4).

Figure 6.4: *Distribución de Historia Admisión 2020 y Admisión 2019*

6.3 Puntajes Faltantes de Ciencias

La prueba de Ciencias fue aplicada en dos oportunidades. Existe un número de postulantes (74 postulantes¹) que por diversas razones no pudieron rendir esta prueba en ninguna de las aplicaciones. Como el mejor predictor del puntaje de la prueba de Ciencias, según los resultados de admisiones pasadas, es el puntaje de la prueba de Matemática, se le asignó a estos 74 postulantes el resultado de Matemática en Ciencias.

Table 6.2: *Correlaciones históricas de Ciencias*

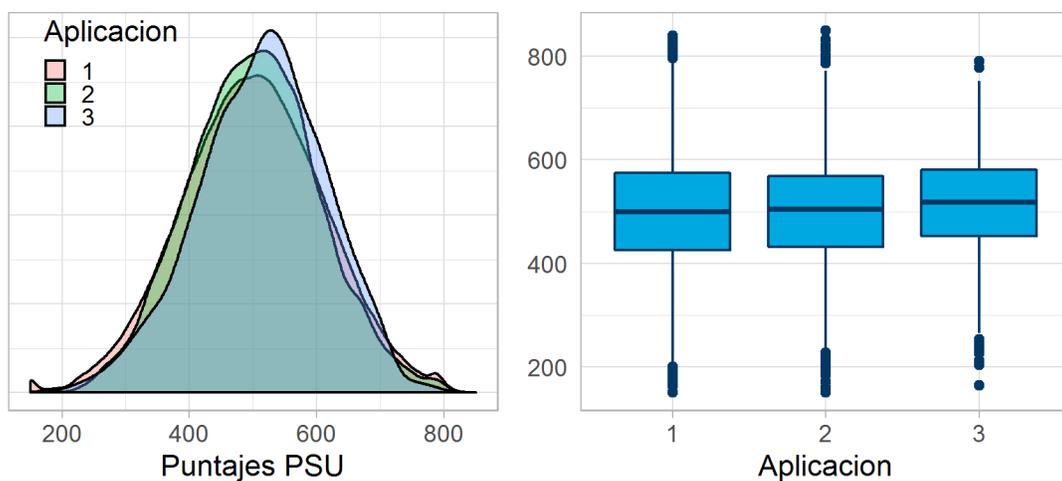
Prueba	ADM2019	ADM2018	ADM2017	ADM2016
MAT	0.71	0.73	0.77	0.78
NEM	0.52	0.53	0.53	0.55
RANK	0.48	0.48	0.47	0.50

¹Número validado al 24 de Febrero 2020

6.4 Distribución de puntajes por aplicación

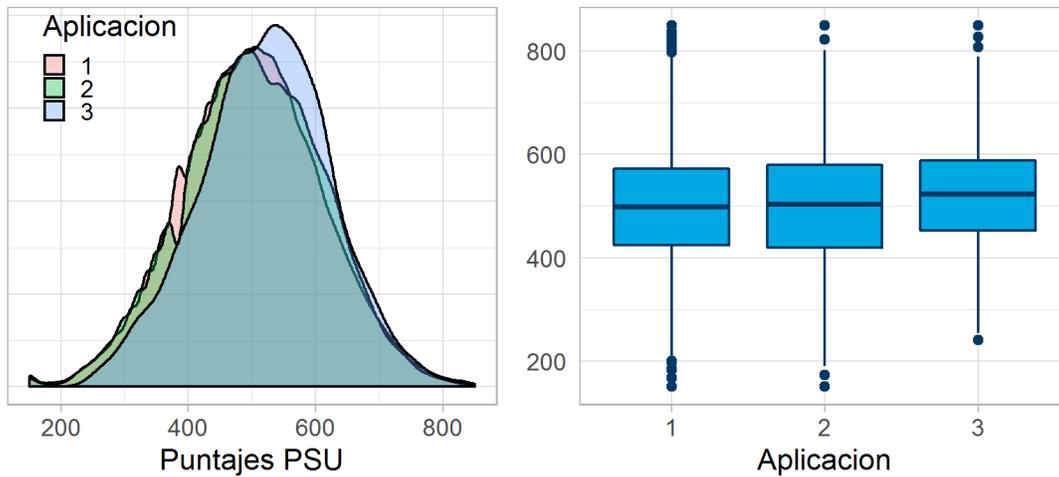
Al obtener los puntajes por aplicación de la prueba, donde para aquellos estudiantes que rindieron más de una aplicación se les considera la mejor prueba, se observa una distribución de puntajes similar en las tres aplicaciones en las pruebas de Lenguaje y Matemática, a excepción de la tercera en la cual se observa que sus puntajes son levemente superiores a las otras dos aplicaciones, tal como muestran la Figura 6.5 y la Figura 6.6 , respectivamente.

Figure 6.5: *Boxplot puntajes por aplicación Lenguaje*



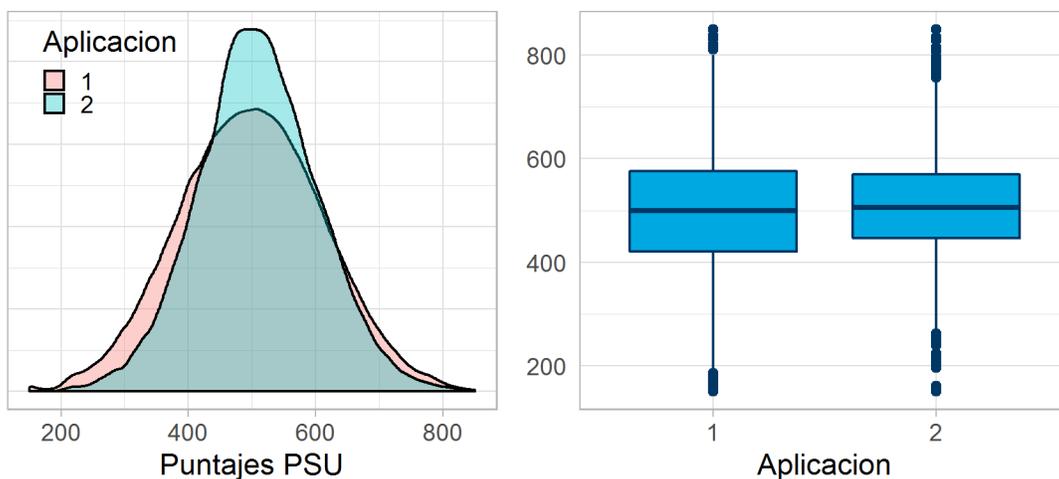
DOCUMENTO

Figure 6.6: *Boxplot puntajes por aplicación Matemática*



Al observar la Figura 6.7 se puede apreciar una distribución similar en ambas aplicaciones de la prueba de Ciencias, con la leve diferencias que los puntajes de la segunda aplicación se encuentran levemente más concentrados que los puntajes de aquellos que rindieron en la primera aplicación.

Figure 6.7: *Boxplot puntajes por aplicación Ciencias*



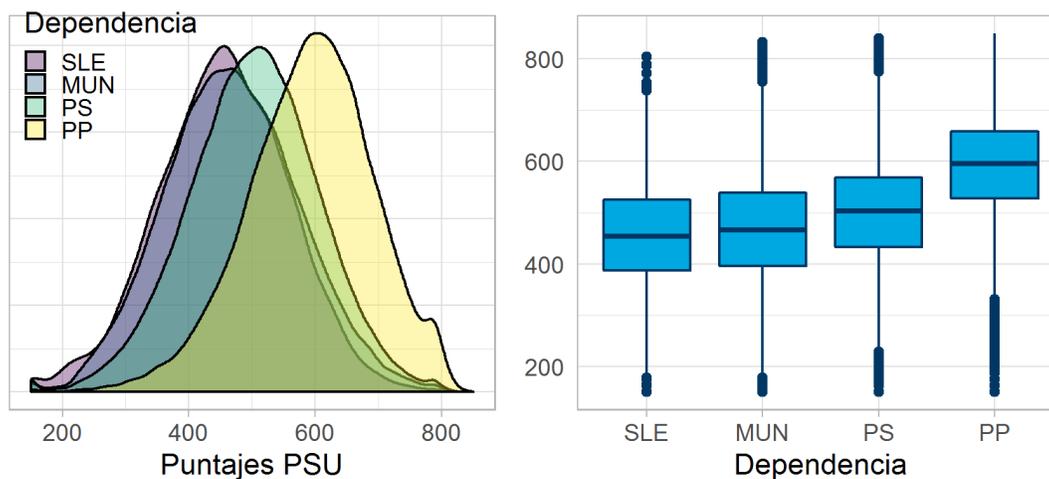
6.5 Densidades de distribución de puntajes por grupo

6.5.1 Dependencia

De los segmentos de dependencia económica establecidos por el Mineduc se observa una diferencia notoria en los puntajes de aquellos postulantes que provienen de una unidad educativa particular pagada con puntajes centrados en aproximadamente en 600 puntos para las tres pruebas, mientras que para unidades educativas particulares subvencionadas sus puntajes se centran alrededor de los 500 puntos y las unidades educativas de dependencia económica municipal o Servicio local de aplicación se centran en los 470 y 455 puntos respectivamente. Estas diferencias son similares a las de años anteriores para cada una de las pruebas.

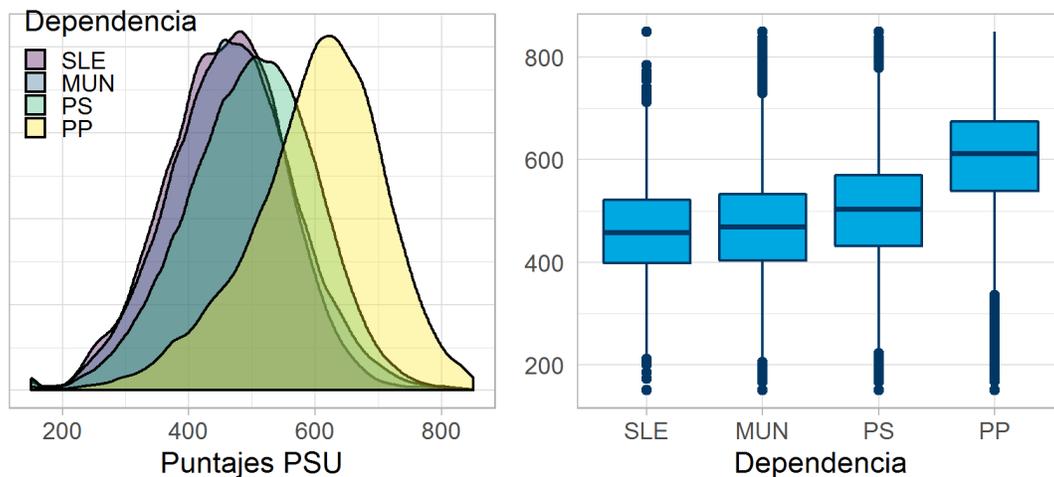
Para Lenguaje, tal como se ve en la Figura 6.8, se observa una concentración en puntajes considerablemente más altos en el grupo particular pagado que en los otros tres grupos, con una diferencia de medias de puntajes entre particulares pagados y municipales de aproximadamente de 123 puntos. En el año 2019, dicha diferencia corresponde a 120 puntos, por lo que no habría grandes diferencias en relación al año anterior.

Figure 6.8: *Densidad puntajes de lenguaje por dependencia*

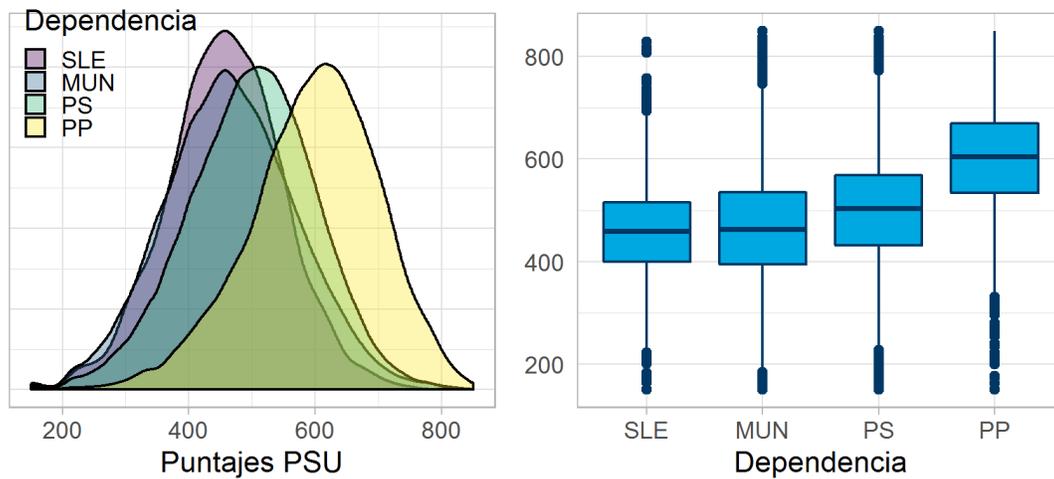


Para Matemática, tal como muestra la Figura 6.9, se observa la concentración en puntajes considerablemente más altos en el grupo particular pagado que en los otros tres grupos, con una diferencia de medias de puntajes entre particulares pagados y municipales de aproximadamente de 133 puntos. En el año 2019 dicha diferencia corresponde a 137 puntos, por lo que no habría grandes diferencias en relación al proceso de admisión anterior.

Figure 6.9: *Distribución de puntajes de matemática por dependencia*



Para Ciencias, tal como se aprecia en la Figura 6.10, se observa la concentración en puntajes considerablemente más altos en el grupo particular pagado que en los otros tres grupos, con una diferencia de medias de puntajes entre particulares pagados y municipales de aproximadamente de 131 puntos. En el año 2019 dicha diferencia corresponde a 124 puntos, siendo esta prueba donde se obtuvieron cambios más grandes en comparación al año anterior.

Figure 6.10: *Distribución de puntajes de ciencias por dependencia*

6.5.2 Año de egreso de la enseñanza media

De acuerdo al año de egreso, aquellos postulantes que egresaron de su enseñanza media antes del 2019 poseen, en las tres pruebas, puntajes levemente más altos que la generación del año, tal como se muestra en los gráficos 6.11 , 6.12 , 6.13, respectivamente. En comparación al proceso de admisión anterior, se presenta una diferencia de medias de puntajes entre la generación anterior y la generación actual de 45, 20 y 30 puntos para las pruebas de Lenguaje, Matemática y Ciencias, respectivamente. En el proceso anterior se presentan diferencias de 42, 21 y 32 puntos para las pruebas, por lo que las brechas en este segmento son similares a las anteriores.

Figure 6.11: *Densidad puntajes de lenguaje por año de egreso*

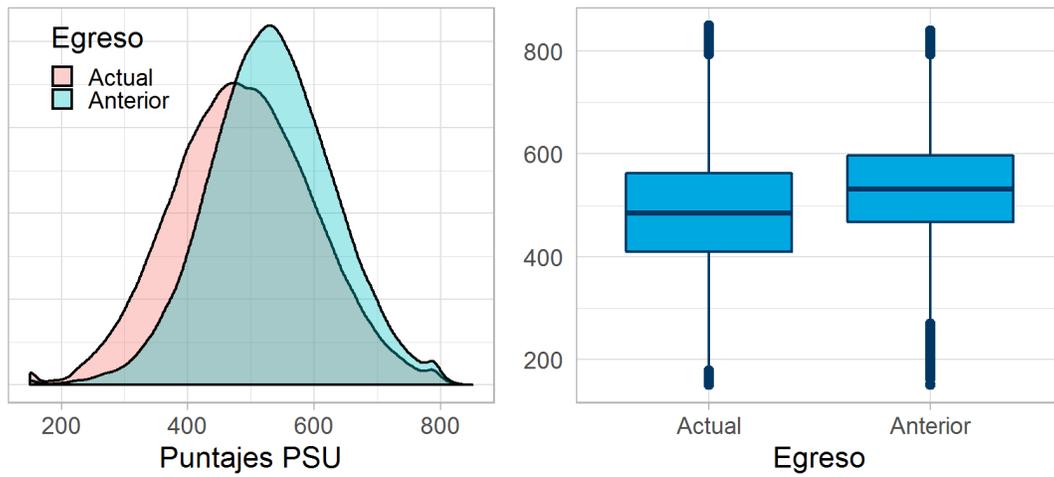


Figure 6.12: *Distribución de puntajes de matemática por año de egreso*

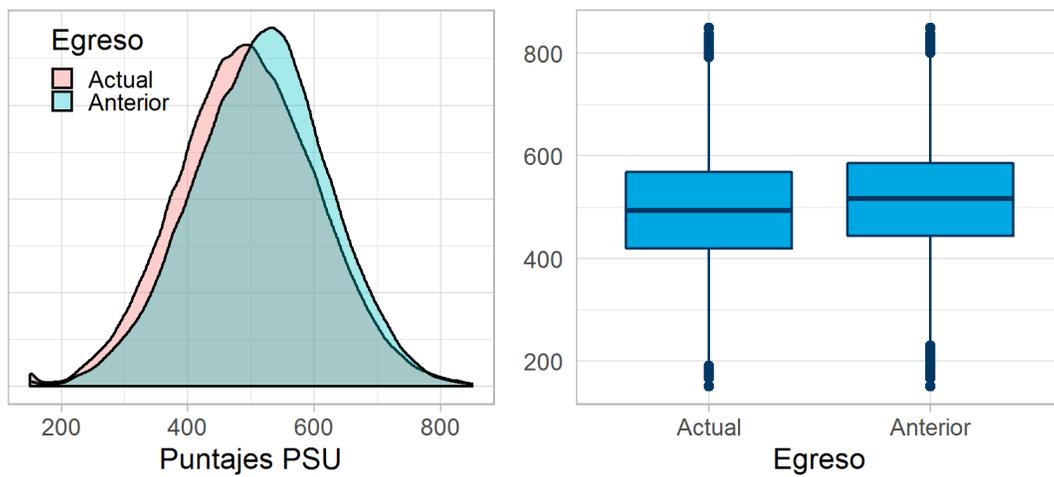
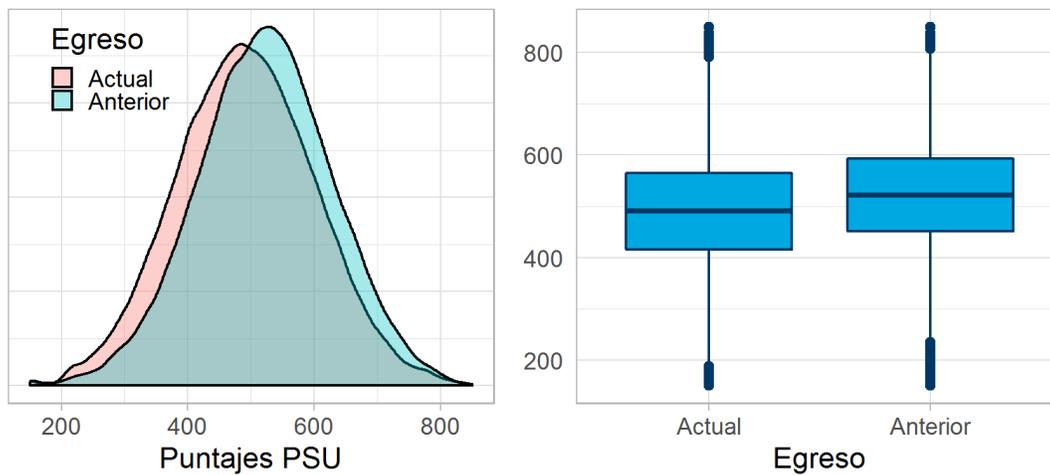


Figure 6.13: *Distribución de puntajes de ciencias por año de egreso*

6.5.3 Rama Educacional

Al realizar un contraste de los puntajes obtenidos en cada una de las pruebas según la rama educacional de la unidad educativa (UED) de egreso de los postulantes, se presenta una gran diferencia entre los egresados de unidades educativas científico humanista y técnico profesional, lo cual se aprecia en las figuras 6.14, 6.15 y 6.16, correspondientes a las pruebas de Lenguaje, Matemática y Ciencias. En comparación al proceso de admisión anterior, existe una diferencia de medias de puntajes entre HC y TP de 85, 77 y 92 puntos para las pruebas de Lenguaje, Matemática y Ciencias, respectivamente. En el proceso anterior se presentan diferencias de 80, 72 y 73 puntos para las pruebas, por lo que en ciencias se presentaría un brecha mayor en comparación al año anterior.

Figure 6.14: *Densidad puntajes de lenguaje por rama*

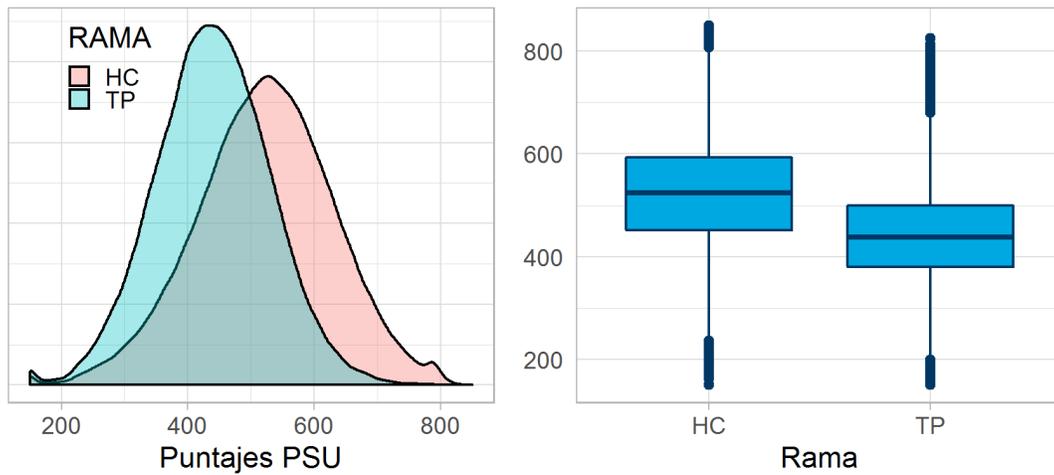


Figure 6.15: *Distribución de puntajes de matemática por rama*

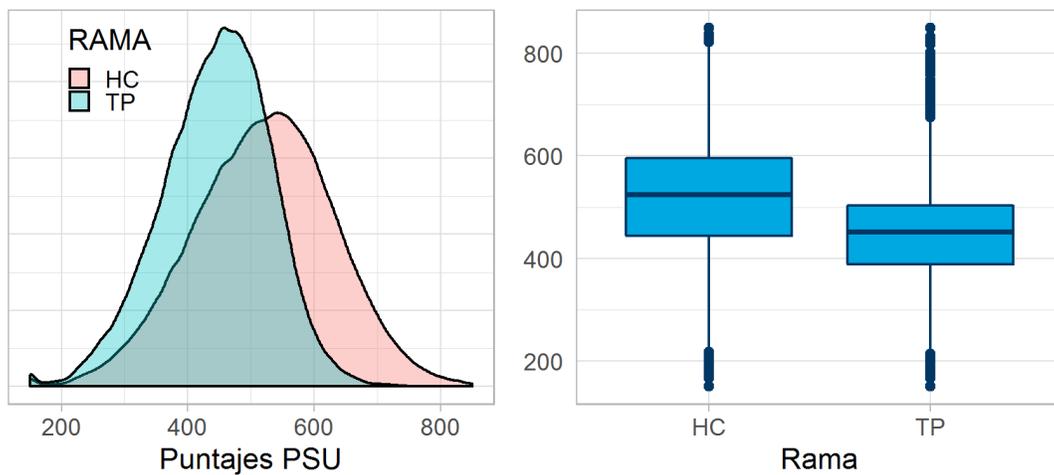
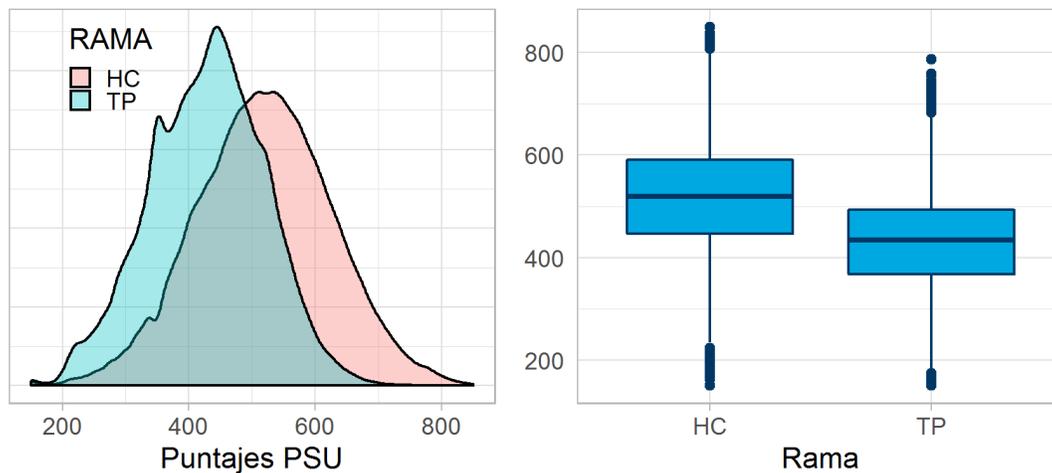


Figure 6.16: *Distribución de puntajes de ciencias por rama*

6.5.4 Sexo

Analizando la distribución de los puntajes de los postulantes según sexo, se aprecia una leve diferencia en el comportamiento de las tres pruebas, lo cual se aprecia en las figuras 6.17, 6.18 y 6.19, correspondientes a las pruebas de Lenguaje, Matemática y Ciencias. En comparación al proceso de admisión anterior, existe una diferencia aproximada de medias de puntajes entre sexo femenino y sexo masculino de 5, -17 y -23 puntos para las pruebas de Lenguaje, Matemática y Ciencias, respectivamente. En el proceso anterior se presentan diferencias aproximadas de 5, -17 y -18 puntos para las pruebas, por lo que en ciencias se presentaría un brecha mayor en comparación al año anterior.

Figure 6.17: *Distribución de puntajes de lenguaje por sexo*

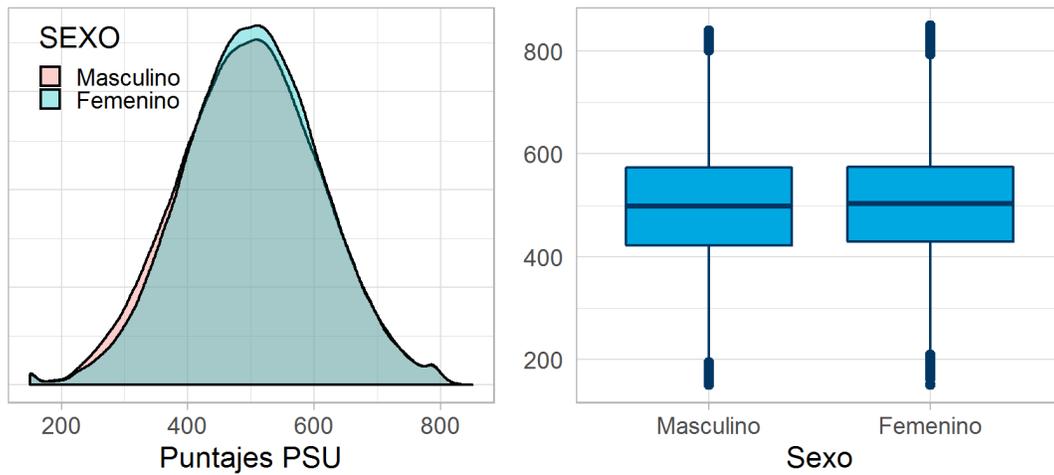


Figure 6.18: *Distribución de puntajes de matemática por sexo*

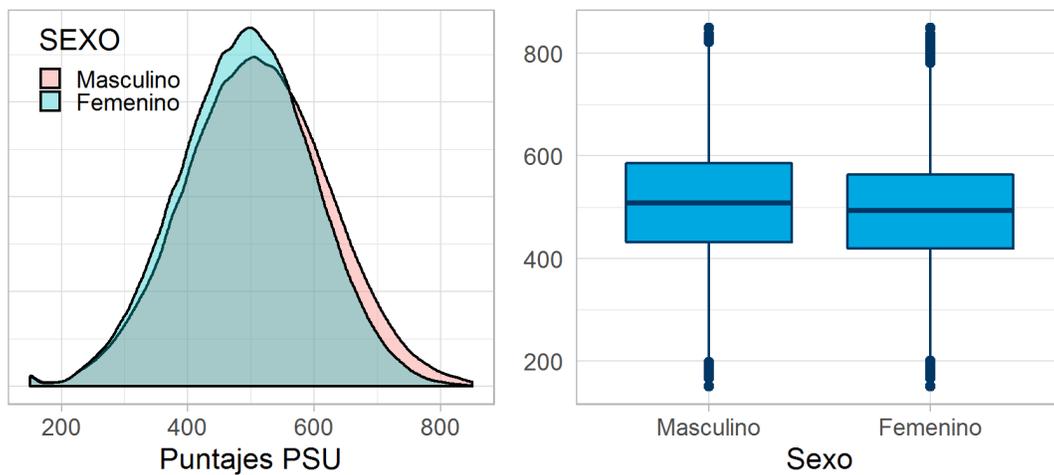


Figure 6.19: *Distribución de puntajes de ciencias por sexo*

