



inevap
DURANGO
INSTITUTO DE EVALUACIÓN DE POLÍTICAS
PÚBLICAS DEL ESTADO DE DURANGO

Términos de Referencia para la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango

Programa Anual de Evaluación de las Políticas y Programas Públicos del
Estado de Durango (PAE) 2024

abril 2024

Contenido

Marco de la evaluación.....	3
Antecedentes	3
Características de la intervención	3
Objetivos.....	4
Hipótesis	4
Alcance	4
Usos y usuarios previstos	5
Principios.....	5
Consideraciones administrativas de la evaluación	6
Etapas, productos y plazos	6
Vías de comunicación.....	7
Roles y responsabilidades.....	7
Confidencialidad y propiedad intelectual.....	8
Consideraciones técnicas de la evaluación.....	8
Metodología.....	8
Preguntas de evaluación	9
Criterios para responder las preguntas	12
Criterios para redactar la propuesta de recomendaciones y observaciones.....	12
Configuración del informe de evaluación	13
Resumen ejecutivo	14
Introducción.....	14
Descripción de la intervención evaluada.....	15
Descripción del diseño y proceso de evaluación	15
Fortalezas, oportunidades, debilidades y amenazas	15
Conclusiones.....	16
Ficha de la evaluación.....	16
Anexos	17
Apéndices	17
Glosario.....	17
Deflactor del PIB.....	17
Referencias	18

Marco de la evaluación

Antecedentes

El Instituto de Evaluación de Políticas Públicas del Estado de Durango (Inevap) es el organismo constitucional autónomo encargado de coordinar y realizar las evaluaciones de las políticas y programas públicos que operan los poderes ejecutivo, legislativo y judicial, otros órganos constitucionales autónomos, los municipios, y las entidades paraestatales y paramunicipales del estado de Durango.

El objetivo del Inevap es generar información para mejorar las intervenciones públicas a partir del trabajo independiente, objetivo, transparente, colaborativo y de calidad. Las evaluaciones del Inevap reportan sobre asuntos de interés público, e incorporan análisis sobre los factores subyacentes de éxito de un programa o política pública, contribuyen a la cultura de transparencia y guían sobre referentes y buenas prácticas en el diseño y la implementación de las intervenciones públicas. El trabajo de Inevap ayuda a todo aquel que usa recursos públicos a reflexionar sobre los resultados obtenidos para mejorar los servicios públicos.

Los *Lineamientos Generales para la Evaluación de las Políticas Públicas y de los Programas Presupuestarios del Estado de Durango* vigentes establecen las bases para el cumplimiento de la función de evaluación del Inevap. Estos lineamientos definen que las evaluaciones deben apegarse a los modelos de Términos de Referencia (TdR) que emita el Inevap, los cuales especifican la intervención pública, instrumentos, acciones, resultados y alcances que sean objeto de evaluación.

Previo a la decisión de evaluar, el Inevap realizó una valoración de evaluabilidad que estimó la disposición y preparación de las intervenciones para ser evaluadas. Los resultados de dicha valoración fueron el principal insumo para configurar los objetivos, hipótesis y preguntas de evaluación que se documentan en estos TdR. De esta forma, el diseño de la evaluación responde a las características de las intervenciones y se sincroniza con las necesidades de información de sus responsables.

Una evaluación en el Colegio de Bachilleres del Estado de Durango (COBAED) permitiría conocer la operación y actividades llevadas a cabo por parte del Colegio, con el fin de identificar aquellas que aportan al mejoramiento de la educación media superior y del mismo modo, identificar aquellas que es necesario retomar o implementar. También, se pueden encontrar áreas de mejora que pueden ayudar a la consecución de objetivos prioritarios como la calidad educativa, la disminución de la deserción, el incremento de la eficiencia terminal, la mejora infraestructural y de los procesos administrativos.

Características de la intervención

El 27 de octubre de 1987 se publicó el Decreto No. 63 que creó al Colegio de Bachilleres del Estado de Durango (COBAED) como un organismo público descentralizado, dotado de autonomía técnica y operativa. El objetivo del COBAED es impartir e impulsar la educación a nivel bachillerato en Durango. Actualmente el COBAED cuenta con un programa de estudios y organización académica homologados al sistema de la Dirección General de Bachilleratos. El Colegio tiene presencia en 24 de los 39 municipios del Estado, en los que se distribuyen 33 planteles y una plantilla de alrededor de 1,600 trabajadores (900 docentes) que atiende a una matrícula de más de 16,000 estudiantes.

El COBAED cuenta con objetivos prioritarios encaminados a posicionarse como la mejor institución de educación media superior en Durango y una de las mejores a nivel nacional, en el ánimo de brindar a sus alumnos una educación media superior de calidad y un mejor servicio a docentes y trabajadores. El crecimiento de la institución y la mejora de sus instalaciones son también un eje prioritario en la visión de sus directivos.

El abandono escolar es un tema que llama la atención en todos los niveles e instituciones educativas, y COBAED no es la excepción, pues actualmente existe una tasa abandono de 9.4%, que si bien ha disminuido (16% durante la pandemia), se plantea alcanzar el porcentaje más bajo posible. Por otro lado, la eficiencia terminal es del 70%, el cual sigue siendo bajo y se tiene que contrarrestar con estrategias, particularmente en aquellos alumnos que abandonan la educación por razones meramente académicas.

La infraestructura es un factor crucial al momento de brindar y recibir educación, pues las condiciones físicas de los espacios permiten que esta se lleve a cabo de mejor forma. En el subsistema COBAED, la mayoría de los planteles cuenta con buena infraestructura y servicios básicos. Sin embargo, existe infraestructura que requiere atención o mejora, en especial algunos de los planteles ubicados en regiones lejanas a la mancha urbana de los municipios.

Objetivos

Los objetivos de la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango tienen un enfoque formativo y se alinean con el paradigma orientado al uso de los resultados del ejercicio evaluativo.

Objetivos de la evaluación

Principal	Analizar los desafíos del servicio de educación media superior en el subsistema COBAED y las estrategias para abordarlos.
Específicos	- Examinar las características y contexto del servicio de educación media superior en el subsistema COBAED. - Identificar los retos pedagógicos de la educación media superior en el subsistema COBAED y analizar las estrategias para superarlos. - Examinar la pertinencia de las estructuras organizativas para garantizar el servicio de educación media superior en el subsistema COBAED.

Hipótesis

Las hipótesis que guían la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango derivan del encuentro de las necesidades de información y los requerimientos de la intervención para alcanzar sus resultados esperados. Estas hipótesis se vinculan con los objetivos y secciones de la evaluación, y apoyan la configuración de su resumen ejecutivo.

Hipótesis de la evaluación

Principal	El subsistema COBAED reconoce los desafíos del servicio de educación media superior y los aborda con estrategias que aseguran la normalidad y calidad del servicio.
Adicionales	- El COBAED conoce las características y entorno del servicio de educación media superior. - El COBAED reconoce los retos de la enseñanza en la educación media superior e implementa estrategias para asegurar la calidad del servicio educativo. - El COBAED define estructuras organizacionales y cuenta con recursos para garantizar el servicio de educación media superior a la población.

Alcance

Por sus objetivos, la evaluación se alinea con los criterios de evaluación del Comité de Ayuda al Desarrollo de la Organización para la Cooperación y el Desarrollo Económicos.

Criterios de evaluación

Pertinencia	El grado en que los objetivos y el diseño de la intervención son coherentes con las necesidades, políticas y prioridades de los beneficiarios, asociados e instituciones, así como con las prioridades a pesar de los cambios del contexto.
Coherencia	La medida en la que la intervención es compatible con otras dentro de un mismo contexto, de manera que se complementan y armonizan con el objeto de agregar valor y evitar la duplicidad de esfuerzos.
Eficacia	El grado en que la intervención ha alcanzado, o se espera que alcance, sus objetivos teniendo en cuenta los resultados diferenciados entre grupos.
Eficiencia	La medida en que la intervención cumple, o es probable que cumpla, con los resultados de manera económica y oportuna.
Impacto	El grado en que la intervención ha producido, o se espera que produzca, efectos significativos, positivos o negativos, directos o indirectos, en el nivel más alto.
Sostenibilidad	La medida en que continúan, o es probable que continúen, los beneficios netos de la intervención a lo largo del tiempo.

Fuente: Organización para la Cooperación y el Desarrollo Económicos (2019). *Mejores criterios para una mejor evaluación. Definiciones revisadas de los criterios de evaluación y principios para su utilización*. OCDE. Adaptado por Inevap.

El equipo evaluador debe interpretar la información disponible, emitir juicios valorativos y proponer vías de mejora de los resultados de la intervención. Para ello es necesario analizar la información documental pública y provista por los responsables de la intervención, además de la evidencia relacionada con el logro de sus objetivos y otras referencias de buenas prácticas. Así mismo, el equipo evaluador debe realizar entrevistas personales o grupos focales con los responsables de la intervención para recoger su experiencia y conocimiento y propiciar su compromiso con el ejercicio evaluativo. De forma adicional, el equipo evaluador puede aportar elementos y herramientas de análisis que fortalezcan la evaluación, pero esto no debe representar un costo extra ni sustituir los requerimientos de estos TdR.

La evaluación se concentra en el ejercicio fiscal terminado, aunque pueden añadirse hallazgos de años anteriores si su pertinencia e importancia lo justifican. La cobertura geográfica de la evaluación depende de la intervención evaluada, pero se limita al territorio estatal.

Usos y usuarios previstos

Los principales usuarios previstos de la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango son los responsables directivos y de operación en el subsistema COBAED, la Subsecretaría de Educación Media Superior de la Secretaría de Educación del Estado de Durango (SEED) y la Dirección de Planeación y Evaluación de la SEED. El Inevap espera que los usuarios primarios utilicen los resultados de la evaluación para reflexionar sobre los desafíos que prevalecen en el servicio de educación media superior, especialmente, en el subsistema COBAED y tomar acciones enfocadas a mejorar el servicio. Un segundo grupo de usuarios previstos de esta evaluación son las autoridades hacendarias del estado, para quienes los resultados de la evaluación sirven para reflexionar sobre las necesidades de financiamiento del servicio de educación media superior.

Igualmente, toda evaluación es un ejercicio de transparencia y rendición de cuentas que fortalece la calidad democrática de la sociedad. Así, los resultados de esta evaluación contribuyen a elevar el nivel de la discusión pública sobre la pertinencia, relevancia, capacidad y resultados de la política de educación media superior, a través del subsistema COBAED y cómo se pueden mejorar.

Principios

La Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango debe conducirse por el equipo evaluador y ser validada por el Inevap según los principios de utilidad, factibilidad, propiedad y precisión.

Principios de la evaluación

Utilidad	La evaluación se sincroniza con las necesidades de información de sus usuarios previstos.
Factibilidad	La evaluación es realista, eficiente y efectiva de acuerdo con sus procesos, contexto y recursos.
Propiedad	La evaluación se conduce de manera legal, ética y transparente.
Precisión	La evaluación expone y justifica técnicamente la interpretación, valoración y recomendaciones sobre los factores que determinan el valor y mérito de la intervención evaluada.
Fuente: Yarbrough, D.B., Shula, L.M., Hopson, R.K., & Caruthers, F.A. (2010). <i>The Program Evaluation Standards: A guide for evaluators and evaluation users (3rd. ed)</i> . Thousand Oaks, CA: Corwin Press. Adaptado por Inevap. Traducción libre.	

El equipo evaluador debe propiciar el involucramiento y participación de los responsables de la intervención, de los usuarios previstos de la evaluación y de aquellos que pueden ser afectados por sus resultados. La independencia del equipo evaluador es indispensable para garantizar la objetividad del ejercicio evaluativo, por lo que cualquier conflicto de interés o situación potencial de afectar la imparcialidad de la evaluación debe ser evitada.

Consideraciones administrativas de la evaluación

Etapas, productos y plazos

El plan de evaluación del equipo evaluador debe considerar mínimamente **4 etapas**: 1) Apertura de la evaluación, 2) Desarrollo metodológico, 3) Informe preliminar final y, 4) Dictaminación de la evaluación. El desarrollo de la evaluación deriva en, al menos, **1 producto intermedio**: la versión preliminar de la evaluación, **1 producto final**: el informe de la evaluación y **1 producto complementario**: la presentación de los resultados de la evaluación.

Etapas, productos y plazos de la evaluación

Etapas y productos	Fecha límite y pago
1. Apertura de la evaluación Presentación del equipo evaluador y del instrumento y proceso de evaluación	[Cuando aplica, 30% del monto contratado]
2. Trabajo de evaluación Desarrollo metodológico del equipo evaluador	
3. Informe preliminar final - Configuración del informe preliminar de la evaluación <i>Producto intermedio: Versión preliminar de la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango</i>	
Periodo para comentarios de la dependencia o entidad responsable de la intervención y del Inevap al informe preliminar de la evaluación	Hasta 5 días hábiles
4. Dictaminación de la evaluación - Dictaminación de la evaluación por el Consejo General del Inevap <i>Producto final: Informe de la Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango</i> <i>Producto complementario: Presentación de la evaluación</i>	Último día del año [Cuando aplica, 70% del monto contratado]

El equipo evaluador debe entregar los productos de la evaluación a la dependencia o entidad responsable de la intervención y al Inevap según las condiciones y plazos de la convocatoria, propuesta técnica de trabajo y contrato de evaluación

correspondiente, pero debe atender las fechas y periodos límite que indican estos TdR. Por su parte, los mecanismos y plazos para el pago del equipo evaluador se especifican en el mismo documento contractual.

Al momento de entregar la versión preliminar de la evaluación, el equipo evaluador debe hacer una presentación ejecutiva de los principales resultados a la dependencia o entidad responsable de la intervención y al Inevap. Dicho producto de la evaluación debe revisarse por la dependencia o entidad responsable de la intervención y validarse técnicamente por el Inevap para emitir sus comentarios al equipo evaluador en un periodo de **hasta 5 días hábiles**. A su vez, el equipo evaluador cuenta con otros **5 días hábiles** para responder los comentarios y hacer las adecuaciones necesarias al producto de la evaluación, manteniendo siempre su independencia y objetividad.

- Versión preliminar en extenso: formato Word editable (.docx) según la *Guía de formato y redacción de los informes de evaluación*
- Presentación de resultados preliminares: formato Power Point (.pptx) según la plantilla proporcionada

Luego que el Inevap valide técnicamente el informe preliminar de la evaluación, el documento pasa a ser dictaminado por el Consejo General, de acuerdo con el artículo 13, fracción V de la Ley de Evaluación de Políticas Públicas del Estado de Durango. Hasta que la evaluación haya sido dictaminada aprobatoriamente, se genera el producto final de la evaluación y se considera finalizada.

- Versión final en extenso: formato Word editable (.docx) según la *Guía de formato y redacción de los informes de evaluación*
- Presentación de resultados: formato Power Point (.pptx) según la plantilla proporcionada

Vías de comunicación

La comunicación y encuentros entre el equipo evaluador, el Inevap y la dependencia o entidad responsable de la intervención pueden ser presenciales o a distancia vía telefónica, videoconferencia, correo electrónico o plataformas de mensajería instantánea, pero los productos de la evaluación deben entregarse mediante oficio.

El Inevap es un intermediario entre el equipo evaluador y la dependencia o entidad responsable de la intervención para la comunicación, aclaración de consultas, seguimiento de avances, retroalimentación, reuniones e intercambios de información, documentos o productos de la evaluación. Cualquier situación de este tipo debe ser compartida con el Inevap.

Roles y responsabilidades

En el proceso de evaluación participan **3 actores** con roles y responsabilidades propias para cumplir los objetivos del ejercicio evaluativo: 1) el responsable de la evaluación designado por el Inevap, 2) el equipo evaluador y, 3) los responsables de la intervención que se evalúa.

Roles y responsabilidades

Responsable de la evaluación del Inevap

- Entregar la información documental provista por los responsables de la intervención al equipo evaluador.
 - Facilitar la comunicación entre el equipo evaluador y los responsables de la intervención.
 - Coordinar la entrega y presentación de los productos de la evaluación y los demás encuentros del equipo evaluador con los responsables de la intervención.
 - Comentar y validar técnicamente los productos de la evaluación.
 - Integrar la propuesta de recomendaciones y observaciones.
 - Publicar y difundir los resultados de la evaluación.
-

Roles y responsabilidades

Equipo evaluador

- Notificar al Inevap sobre las reuniones, entrevistas y otros encuentros con los responsables de la intervención.
- Avisar al Inevap sobre cualquier intercambio de información, documentos o productos de la evaluación con la dependencia o entidad responsable de la intervención.
- Integrar y entregar los productos de la evaluación a los responsables de la intervención y al Inevap.
- Resolver las dudas y comentarios de los responsables de la intervención respecto de la evaluación.
- Atender los comentarios y requerimientos técnicos del Inevap.

Responsables de la intervención que se evalúa

- Proporcionar la información documental necesaria para la evaluación y aquella que se solicite durante el proceso.
 - Participar en las entrevistas, grupos focales, encuestas y otras técnicas de investigación que el equipo evaluador realice para la evaluación.
 - Compartir con el Inevap cualquier intercambio de información o documentos que se transmitan al equipo evaluador.
 - Revisar y comentar los productos de la evaluación.
-

Confidencialidad y propiedad intelectual

Toda la información a la que tenga acceso el equipo evaluador como parte del proceso de evaluación, ya sean documentos, registros administrativos, bases de datos o lo dicho durante entrevistas, encuestas, grupos focales u otras interacciones con la dependencia o entidad responsable de la intervención, es de uso estricto para la evaluación. El equipo evaluador no está autorizado para difundirla de ninguna manera ni utilizarla para otros fines. Los datos e información deben protegerse de acuerdo con la normatividad aplicable y evitar cualquier alteración, pérdida, tratamiento o acceso no permitido.

Igualmente, todos los derechos de propiedad sobre los resultados y productos de la evaluación son del Inevap, y de la dependencia o entidad responsable de la intervención en caso de haber pagado una parte o todo el ejercicio. Cualquier otra consideración al respecto se documenta en el contrato de evaluación correspondiente. De esta forma, el equipo evaluador no está autorizado para publicar, difundir, utilizar o reproducir total o parcialmente cualquier resultado o producto de la evaluación por sí solo sin permiso explícito del Inevap.

Consideraciones técnicas de la evaluación

Metodología

La Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango tiene un enfoque formativo y se alinea con el paradigma orientado al uso de los resultados del ejercicio evaluativo. Para su desarrollo, dicha evaluación es predominantemente cualitativa, con un alcance descriptivo y utiliza técnicas no experimentales como:

- Análisis documental
- Observación directa
- Análisis comparado
- Entrevistas a profundidad
- Grupos de enfoque

Otras técnicas de investigación pueden ser empleadas en la medida que aporten elementos sustanciales para responder las preguntas e hipótesis de evaluación, pero el equipo evaluador debe exponer los argumentos que justifiquen su utilización.

Dependiendo de la información, es posible utilizar herramientas estadísticas como medidas de tendencia central, variabilidad, identificar relaciones entre variables y modelos básicos de regresión.

En el caso de usar instrumentos de recolección como encuestas, ya sean en línea o mediante operativos presenciales, es importante referirse a la *Guía para el diseño estadístico de encuestas del Inevap*.

Preguntas de evaluación

La Evaluación Específica del Programa de Atención al Alumno del Colegio de Bachilleres del Estado de Durango debe responder **10 preguntas** distribuidas en **3 secciones** que se relacionan con las hipótesis del ejercicio.

Matriz de evaluación	
#	Pregunta de evaluación
Sección: I. Dimensión descriptiva-contextual	
Hipótesis: El COBAED conoce las características y entorno del servicio de educación media superior.	
<i>Fuentes mínimas de información: registros administrativos, bases de datos sobre las características de las escuelas, formatos 911 de inicio y cierre de los ciclos escolares 2021-2022, 2022-2023 y 2023-2024, estadística educativa, datos del Sistema de Información y Gestión Educativa, documentos de evaluación o investigación educativa, resultados de encuestas de egresados.</i>	
1	¿Cuál es la situación y contexto del servicio de educación media superior que presta el COBAED? <i>Pertinencia</i> - Se describen las características básicas y la evolución temporal del servicio de educación media superior que presta el COBAED según: - Municipios - Grupos - Alumnos - Sexo - Condición de repetición - Capacidades y aptitudes sobresalientes - Población indígena - Personal directivo y docente - Sexo - Categorías válidas de personal - Formación profesional - Se analizan las estrategias de captación de estudiantes de educación secundaria al subsistema COBAED. - Se describen las condiciones de infraestructura de los planteles COBAED según: - Acceso a servicios básicos - Aulas - Materiales de construcción - Disponibilidad y calidad del equipamiento - Se analiza el contexto de los planteles COBAED: - Ámbito rural y urbano - Tamaño de la localidad - Grado de marginación - Condición de ubicación geográfica - Grado de accesibilidad a carretera pavimentada (GACP) - Se analiza la estadística de egresados del subsistema COBAED y si se conoce el nivel de absorción en educación superior respecto de otros subsistemas.
2	¿Cuáles son las características educativas de la población de educación - Se analizan las características educativas de la población de estudiantes de media superior en el subsistema COBAED: - Eficiencia terminal - Tasa de aprobación - Aprovechamiento por asignatura

Matriz de evaluación

#	Pregunta de evaluación	Elementos esperados en la respuesta
media superior del COBAED?		
<i>Eficacia</i>		
Sección: II. Dimensión pedagógica-didáctica Hipótesis: El COBAED reconoce los retos de la enseñanza en la educación media superior e implementa estrategias para asegurar la calidad del servicio educativo. <i>Fuentes mínimas de información: normatividad, documentación oficial, mapa curricular, Nuevo Marco Curricular Común de la Educación Media Superior, planes de estudio de educación media superior, programación regular, libros de texto, materiales audiovisuales, materiales informáticos, programas de unidades de aprendizaje curricular de área básica, específica y laboral, matrices de correspondencia, resultados de exámenes o pruebas estandarizadas, actas, minutas o evidencia de reuniones de las Academias Estatales, proyectos integradores de las Academias por plantel, documentos de planeación didáctica (metodologías activas, dosificación de temas, progresiones), ejemplos de planes de clase, proyectos transversales, materiales de cursos de inducción para estudiantes de nuevo ingreso, información del monitoreo disciplinar para los semestres, formatos 911, convenios de colaboración con instituciones de educación superior, registros de estrategia de tutorías (estadística, tests de estilos de aprendizaje y modelos de entrevistas), estadística de canalizaciones sobre Primeros Auxilios Psicológicos y otras intervenciones (educación sexual, prevención de adicciones, salud mental, escuela para padres).</i>		
3	¿Cuáles son los retos del subsistema COBAED para implementar el servicio educativo de media superior y con respecto del marco curricular común de la educación media superior?	- Se describe el Marco Curricular Común de la Educación Media Superior y los aprendizajes esperados en el subsistema COBAED. - Se exponen las estrategias y herramientas para implementar el currículo oficial en el subsistema COBAED. - Se expone la organización del tiempo escolar en los planteles COBAED. - Se identifican los eventos o condiciones que impiden el cumplimiento de las horas lectivas obligatorias para la educación media superior y las estrategias para evitarlas o superarlas. - Se exponen las estrategias del subsistema COBAED para garantizar la normalidad mínima de operación escolar de los planteles.
<i>Eficiencia</i>		
4	¿Cuál es la disponibilidad de materiales educativos del subsistema COBAED?	- Se indica la disponibilidad de materiales educativos y didácticos en los planteles COBAED: - Libros de texto - Materiales como cuadernillos, guías, y otros. - Se describen los mecanismos de coordinación para crear y/o adquirir y distribuir los materiales educativos y didácticos.
<i>Eficiencia</i>		
5	¿Cuáles son los resultados del aprendizaje en el subsistema COBAED?	- Se analizan los resultados de la evaluación estandarizada implementada por la SEED en la educación media superior de Durango. - Se realiza una comparación de medias de porcentajes de aciertos entre los subsistemas de educación media superior. - En función de los datos, aplicar un marco estadístico robusto: análisis de muestras desbalanceadas, tipo de variables, distribución normal e igualdad de varianzas. Prever ANOVA¹
<i>Eficacia</i>		

¹ Analysis of Variance en inglés. Es una colección de modelos estadísticos en la que se asume que la varianza está particionada debido a la presencia de diferentes variables explicativas.

Matriz de evaluación

#	Pregunta de evaluación	Elementos esperados en la respuesta
		debido a la presencia de más de dos tipos de subsistema de educación media superior.
		○ En función de los datos, se sugiere la elaboración de mapas de Wright² para representar los datos politómicos de asignaturas clave para los grupos semestrales. ○ En función de los datos, se sugiere realizar un análisis sobre la probabilidad de responder adecuadamente un reactivo y determinar los reactivos más problemáticos para los estudiantes, para las asignaturas evaluadas. Ver metodología sobre modelos logísticos basados en la Teoría de Respuesta al Ítem (TRI).
Sección: III. Dimensión institucional-organizacional		
Hipótesis: El COBAED define estructuras organizacionales y cuenta con recursos para garantizar el servicio de educación media superior a la población.		
<i>Fuentes mínimas de información: normatividad del COBAED, documento de creación, reglamento interior, manuales de organización, manuales de procesos, actas e informes de las sesiones de Junta Directiva, registros y documentos sobre la supervisión, documentos diagnósticos, documento ADN (existencia y necesidad) de los recursos humanos, documentos programáticos y financieros, informes trimestrales, estados analíticos del ejercicio del presupuesto, cuenta pública e información contable y presupuestaria, reportes del Sistema de Recursos Federales Transferidos, Matriz de Indicadores para Resultados, convocatorias y resultados para la admisión, promoción y reconocimiento docente, programas de formación académica y de capacitación.</i>		
6	¿Cuáles son los retos de la estructura organizacional y de gestión escolar en el subsistema COBAED? <i>Eficiencia</i>	• Se valora el nivel de coordinación, colaboración y comunicación entre las figuras educativas de la educación media superior en el COBAED. ○ Se valora el nivel de integración entre las estrategias del subsistema COBAED para intercambiar materiales, buenas prácticas y experiencias para fortalecer el trabajo y formación docente. ○ Se analiza la estrategia de capacitación realizada al personal docente y administrativo del subsistema COBAED. • Se analiza la estructura de la organización escolar de los planteles COBAED. • Se identifican las estrategias de las figuras educativas del subsistema COBAED para cumplir con las funciones de gestión escolar. • Se identifican y valoran las oportunidades para el trabajo colegiado en las reuniones de academia. • Se identifican y valoran las estrategias de asesoría y acompañamiento pedagógico y técnico del COBAED para las figuras educativas del subsistema.
7	¿Cuáles son los retos de infraestructura y equipamiento de los planteles del subsistema COBAED? <i>Eficiencia</i>	• Se valora la forma en la que el COBAED realiza y actualiza su diagnóstico o inventario de la infraestructura física de los planteles. • Se valora la forma en la que COBAED realiza y actualiza su diagnóstico o inventario de equipamiento de los planteles. • Se valora la existencia de estrategias, programas o actividades encaminadas al fortalecimiento de la infraestructura y equipamiento de los planteles COBAED.
8	¿El subsistema COBAED cuenta con un diagnóstico de	• Se indica si el COBAED cuenta con un diagnóstico de déficit y superávit del personal educativo y administrativo, con base en la demanda escolar y el análisis de vacancias.

² Conocido también como mapa de elementos o mapa de elemento-persona que representa la ubicación de respondientes y reactivos en una misma escala.

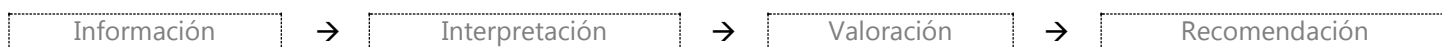
Matriz de evaluación

#	Pregunta de evaluación	Elementos esperados en la respuesta
	necesidades de recursos humanos para la prestación del servicio de educación media superior?	- Se analiza si el COBAED cuenta con procesos documentados y formatos establecidos para la gestión de los recursos humanos del subsistema. - Se describe cómo el COBAED actualiza su diagnóstico de necesidades de recursos humanos.
<i>Pertinencia</i>		
9	¿El subsistema COBAED cuenta con mecanismos y criterios formales para la admisión, promoción y reconocimiento del personal docente?	- Se valora la existencia de mecanismos, procedimientos y criterios para la admisión de nuevo personal docente al subsistema COBAED de educación media superior. - Se valora la existencia de mecanismos, procedimientos y criterios de promoción a funciones directivas y de supervisión del personal docente del subsistema COBAED. - Se valora la existencia de mecanismos y procedimientos de reconocimiento al desarrollo profesional del personal docente del subsistema COBAED.
<i>Eficiencia</i>		
10	¿El COBAED documenta el destino de las aportaciones federales y estatales para la operación del subsistema?	- Se indica la existencia de mecanismos de coordinación para promover y prestar el servicio educativo media superior entre el nivel federal y el estatal, que comprometan recursos públicos para financiar el subsistema COBAED en Durango. - Se analiza el esfuerzo de gasto de aportación federal y estatal para garantizar la prestación del servicio de educación media superior. - Servicios personales - Gasto de operación - Se analiza si el COBAED registra el destino del gasto por: - Municipio - Planteles - Categorías válidas de personal
<i>Eficiencia</i>		

Criterios para responder las preguntas

Las preguntas de evaluación deben responderse con sustento en la información disponible y hacer explícitos los argumentos que sostienen la interpretación, valoración y recomendaciones del equipo evaluador. Para mayor sofisticación, se recomienda ampliamente el marco de argumentación de Toulmin.³

Trayectoria analítica para responder las preguntas de evaluación



Cada respuesta debe incluir la pregunta de evaluación, el análisis del equipo evaluador, las fuentes de información utilizadas y las recomendaciones u observaciones propuestas. Para todas las preguntas de evaluación, estos TdR señalan los elementos esperados en la respuesta y las fuentes mínimas de información que deben ser empleadas, el equipo evaluador puede añadir conceptos que considere importantes a las respuestas de las preguntas de evaluación, siempre y cuando se respalden con evidencia y la calidad de los argumentos demuestren su importancia.

Criterios para redactar la propuesta de recomendaciones y observaciones

³ Toulmin, S. (2007). *Los usos de la argumentación*. Barcelona. Península.

La propuesta de recomendaciones y observaciones del equipo evaluador debe propiciar la mejora de la intervención y respetar los criterios de claridad, relevancia, justificación y factibilidad.

Criterios para redactar la propuesta de recomendaciones y observaciones

Claridad	Expresión precisa e inequívoca
Relevancia	Aportación específica y significativa para el logro de los objetivos de la intervención
Justificación	Sustentación estricta en los hallazgos de la evaluación
Factibilidad	Viable según las capacidades institucionales de la dependencia o entidad responsable de la intervención
Oportunidad	Pertinente de acuerdo con el contexto y momento que vive la intervención

Cada recomendación u observación propuesta debe clasificarse según una temática de actuación. Esta clasificación no busca limitar el número y tipo de recomendaciones u observaciones que se propongan, ni forzar al equipo evaluador a proponer una recomendación u observación de cada temática.

Temáticas de las recomendaciones y observaciones propuestas

Diseño	Hace referencia a la justificación y consistencia de la lógica causal de la intervención
Cobertura	Busca el establecimiento de estrategia de cobertura para ampliarla o limitarla
Indicadores	Mejora la pertinencia, consistencia y alcance de los instrumentos para monitorear el avance y logro de los resultados
Ejecución	Estandariza o documenta los principales procesos necesarios para la implementación
Capacitación	Busca fortalecer el capital humano capacidades institucionales para mejorar la implementación de la intervención
Evaluación	Hace referencia a la necesidad de realizar algún tipo de evaluación
Planeación	Busca fortalecer los instrumentos de planeación y orientación hacia resultados
Productos	Mejora las estrategias para definir o entregar los bienes y servicios de la intervención
Transparencia	Hace referencia a la difusión de los objetivos, avance y resultados
Resultados	Mejora la entrega del servicio público de la intervención o el avance en la atención del problema o necesidad pública que justifica la intervención

La propuesta de recomendaciones y observaciones debe presentarse en el informe de evaluación según la **Tabla #. Propuesta de recomendaciones y observaciones**. Adicionalmente, se sugiere que la redacción de cada recomendación u observación propuesta inicie con un verbo en infinitivo.

Tabla #.

Propuesta de recomendaciones y observaciones

#	Recomendación u observación	Temática	*	Acciones propuestas	Resultados esperados
	<i>Acción de mejora derivada de los hallazgos del informe de evaluación</i>	<i>Clasificación de la recomendación u observación</i>		<i>Actividades específicas para alcanzar la recomendación u observación, cuyo objetivo es lograr el resultado esperado</i>	<i>Mejora esperada y tangible de la intervención derivada de la institucionalización de la recomendación u observación</i>

Nota: el símbolo (*) señala que en la columna debe incluirse el número de la pregunta de evaluación que sustenta la recomendación u observación propuesta.

Contenido del informe de evaluación

La redacción del informe de evaluación debe exponer la interpretación, valoración y recomendaciones del equipo evaluador de modo que fomente el uso de sus resultados. El informe de evaluación debe contener las siguientes secciones:

- Resumen ejecutivo (máximo 4 páginas)

- Marco de la evaluación
- Descripción de la intervención evaluada
- Principales resultados de la evaluación
- Propuesta de recomendaciones y observaciones
- Conclusiones
- Contenido
- Glosario (incluye terminología técnica sobre la evaluación y la intervención evaluada)
- Siglas y acrónimos
- Introducción (máximo 1 página)
- Descripción de la intervención evaluada (máximo 2 páginas)
- Descripción del diseño y proceso de la evaluación (máximo 2 páginas)
- Resultados de la evaluación
 - Secciones
 - Preguntas de evaluación
 - Fortalezas, oportunidades, debilidades y amenazas
- Propuesta de recomendaciones y observaciones
- Conclusiones
- Ficha de la evaluación
- Referencias
- Anexos

En cuanto a la presentación del informe de evaluación, el equipo evaluador debe seguir las indicaciones de la *Guía de formato y redacción de los informes de evaluación* vigente del Inevap.

Resumen ejecutivo

El resumen ejecutivo es una sinopsis de la información más importante del informe de evaluación. El resumen ejecutivo consta de cinco partes que sintetizan:

- Marco de la evaluación, es decir, sus antecedentes, objetivos y métodos
- Descripción de la intervención evaluada
- Principales resultados de la evaluación según las hipótesis del ejercicio.
- Propuesta de recomendaciones y observaciones
- Conclusiones que incluyen también las lecciones aprendidas y las posibles rutas de la evaluación en el futuro.

El resumen ejecutivo no debe añadir información, sino reflejar los resultados más importantes del informe de evaluación. La extensión máxima recomendada del resumen ejecutivo es **4 páginas**.

Introducción

La introducción del informe de evaluación orienta sobre el contenido del informe de evaluación. En especial, la introducción identifica la intervención que está siendo evaluada, las razones para evaluarla, quién realiza la evaluación y la estructura del informe de evaluación. La extensión recomendada de la introducción es **1 página**.

Descripción de la intervención evaluada

En esta sección se describe la intervención evaluada para que los lectores conozcan los antecedentes, características y contexto de la intervención que es objeto de evaluación en un máximo de **2 páginas**. La descripción de la intervención evaluada debe incluir:

- Problema o necesidad que pretende atender y sus objetivos
- Nombre, dependencia o entidad responsable y otras organizaciones involucradas en su diseño o implementación
- Fuentes y comportamiento de su financiamiento
- Historia de la intervención, es decir, año de creación y modificaciones importantes de su diseño o implementación
- Lógica causal o teoría del cambio (actividades, productos, resultados e impacto esperado)
- Población a la que se dirige (potencial) y pretende atender (objetivo)
- Factores importantes del contexto económico, político, social o cultural que influyen la operación de la intervención.
- Otros elementos necesarios para comprender el diseño e implementación de la intervención.

Descripción del diseño y proceso de evaluación

La descripción de la evaluación define sus principales características de diseño y explica cómo sucedió su implementación. La extensión máxima de esta sección de **2 páginas**, donde es necesario exponer:

- Objetivos e hipótesis
- Alcance, en términos conceptuales, geográficos, temporales y de componentes de la intervención
- Fuentes de financiamiento
- Características del equipo evaluador (solo si la evaluación es hecha por un equipo externo al Inevap)
- Principales resultados de evaluaciones pasadas, avances en la atención de las áreas de mejora identificadas y posibles implicaciones para la evaluación actual (solo si la intervención fue evaluada previamente)
- Usos y usuarios principales y secundarios previstos
- Principales enfoques y paradigmas teóricos
- Estrategia, instrumentos y herramientas metodológicas para recopilar información, responder las hipótesis y preguntas de evaluación y generar la propuesta de recomendaciones y observaciones
- Principales fuentes de información utilizadas
- Cronograma de la evaluación, es decir, las fechas de inicio, duración del trabajo y conclusión
- Limitantes, que incluyen los factores que interrumpieron el curso de la evaluación y aquellos que afectaron el alcance o precisión de los resultados previstos de la evaluación.

Fortalezas, oportunidades, debilidades y amenazas

El equipo evaluador debe identificar las fortalezas, oportunidades, debilidades y amenazas de la intervención a partir de los hallazgos de la evaluación. Esta información se presenta en el informe de evaluación según la **Tabla #. Matriz de fortalezas, oportunidades, debilidades y amenazas**.

Tabla #.	
Fortalezas, oportunidades, debilidades y amenazas	
Fortalezas internas	*
Debilidades internas	*

Oportunidades externas

*

Amenazas externas

*

Nota: el símbolo (*) señala que en la columna debe incluirse el número de la pregunta de evaluación que sustenta la fortaleza, oportunidad, debilidad o amenaza mencionada.

Las fortalezas son los elementos, recursos y capacidades internas que contribuyen al logro de los objetivos de la intervención. Al contrario, las debilidades son las limitaciones, fallas o defectos internos que lo obstaculizan.

Las oportunidades son los factores externos no controlables que representan elementos potenciales de crecimiento o mejora. Mientras que las amenazas identifican las condiciones del entorno que, directa o indirectamente, afectan o impiden el logro de los objetivos de la intervención.

Conclusiones

Las conclusiones de la evaluación deben ser claras, precisas y consistentes con los hallazgos de la evaluación. Las conclusiones resumen los hallazgos según las hipótesis de la evaluación, reiteran las recomendaciones y observaciones propuestas e insisten en la importancia de su atención, señalan los aspectos que pueden estudiarse en próximos ejercicios de evaluación y exponen las lecciones aprendidas por la intervención y por la evaluación. La extensión máxima de las conclusiones es de **2 páginas**.

Ficha de la evaluación

El informe de la evaluación debe incluir una ficha que resuma los aspectos administrativos y técnicos de la evaluación y sus resultados.

Ficha de la evaluación

Aspectos administrativos	- Responsable de la evaluación - Principales colaboradores (equipo evaluador) - Organización evaluadora (si aplica) - Unidad administrativa de la dependencia o entidad responsable de la intervención evaluada - Titular de la unidad administrativa de la dependencia o entidad responsable de la intervención evaluada - Unidad administrativa de la dependencia o entidad responsable de la intervención encargada de dar seguimiento a la evaluación - Forma de contratación del equipo u organización evaluadora (si aplica) - Costo total de la evaluación (si aplica) - Fuente de financiamiento de la evaluación (si aplica) - Fecha de inicio de la evaluación (reunión de apertura) - Fecha de conclusión de la evaluación (dictaminación por el Consejo General del Inevap)
Aspectos técnicos	- Objetivo de la intervención evaluada - Siglas de la intervención evaluada - Términos de Referencia de la evaluación - Tipo de evaluación - Objetivo general de la evaluación - Objetivos específicos de la evaluación - Palabras clave de la evaluación (de 3 a 6 palabras)

Ficha de la evaluación

Resultados • Síntesis de las conclusiones de la evaluación (de 3 a 6 enunciados)

Anexos

1. Categoría de plazas del personal docente, no docente, administrativo y manual

Tabla #.

Plazas y costo total de la nómina por categoría válida de personal en el COBAED, 2023

Categoría	Tipo de personal	Número de plazas	Costo total
Director General	Directivo		
[...]			
[...]			
Total			

Apéndices

Glosario

Enfoque formativo	Se refiere a una evaluación que busca mejorar el desempeño de la intervención, más que decidir sobre su valor y continuidad.
Evaluación	Análisis sistemático y objetivo de una intervención cuya finalidad es determinar su pertinencia, eficiencia, eficacia, impacto, sostenibilidad y coherencia, así como su valor y mérito.
Hallazgos	Evidencias obtenidas de una o más evaluaciones para realizar afirmaciones basadas en hechos.
Intervención	Iniciativas de un gobierno o ente público para producir algún cambio o resolver un problema, tales como proyectos, programas, políticas, planes y estrategias.
Propuesta de recomendaciones y observaciones	Sugerencias emitidas por el equipo evaluador derivadas de los hallazgos identificados en evaluaciones cuyo propósito es contribuir a la mejora.
Términos de Referencia	Documento donde se especifica el marco y las consideraciones técnicas y administrativas para realizar una evaluación.

Deflactor del PIB

Para el análisis de datos presupuestarios debe eliminarse el efecto de la inflación, dado que el valor del dinero cambia en el tiempo. Para hacer comparaciones válidas en el tiempo de datos presupuestarios, las cifras deben deflactarse conforme la siguiente fórmula:

$$x_t^r = \left(\frac{x_t^n}{Deflactor_t} \right) \times 100$$

Donde x_t^r es la cifra real en el periodo de alguna variable presupuestaria; x_t^n es la cifra nominal en el periodo de la variable presupuestaria; y $Deflactor_t$ es el valor del deflactor con base 2018 en el año de referencia. Se utiliza base 2018 porque el PIB real en México se calcula con precios constantes de 2018. Los datos del deflactor para 2023 y 2024 se obtuvieron con base en las estimaciones de la Secretaría de Hacienda y Crédito Público (5% y 4.2% de crecimiento, respectivamente), según los Criterios Generales de Política Económica 2024 (SHCP, 2023, pp. 44). En caso de que la información disponible no permita identificar datos presupuestarios para 2018 o antes, y para hacer pertinente el análisis, es posible cambiar la base del deflactor a otro año que se considere relevante. El análisis debe mencionar como nota que se utilizó el índice del deflactor del PIB base 2018 = 100.

Año	PIB Nominal (millones de pesos)	PIB Real (millones de pesos)	Deflactor del PIB Base 2018
2013	16,954,005.70	21,722,561.40	78.0479124
2014	18,137,650.63	22,266,442.95	81.4573332
2015	19,228,614.68	22,868,154.28	84.0846814
2016	20,758,790.50	23,273,490.75	89.1950018
2017	22,536,210.25	23,709,107.30	95.0529683
2018	24,176,670.40	24,176,670.38	100
2019	25,143,108.30	24,115,905.78	104.25944
2020	24,081,765.83	22,036,016.03	109.283665
2021	26,619,085.98	23,300,708.90	114.241528
2022	29,452,832.08	24,220,853.88	121.60113
2023 e/	ND	ND	127.681187
2024 e/	ND	ND	133.809884

Teoría de Respuesta al Ítem

La Teoría de Respuesta al Ítem (TRI) es un marco de análisis para medir rasgos latentes a través de modelos matemáticos. Estos rasgos latentes no pueden ser medidos directamente porque no son observables, pero pueden estimarse con un instrumento. Este instrumento puede ser una prueba estandarizada de varios ítems o reactivos. El principal concepto de la TRI es determinar la probabilidad de una persona en contestar correctamente un reactivo en particular, dado la habilidad particular de la persona. Esta probabilidad de éxito en un reactivo es una función, tanto del nivel del rasgo latente de la persona como de las propiedades del reactivo.

Al rasgo latente usualmente se le identifica como θ ; el valor de θ de una persona se llama ubicación de la persona. Por su parte, las propiedades del reactivo son parámetros comúnmente conocidos como discriminación y dificultad. El parámetro de dificultad usualmente se le conoce como b y representa la ubicación de un reactivo en una escala de habilidad. El parámetro de discriminación se le conoce como a y muestra qué tan rápido la probabilidad de éxito cambia con la habilidad cercana a la dificultad del reactivo. Un ítem con un valor alto de discriminación tiene una alta correlación entre el rasgo latente y la probabilidad de éxito en ese ítem. En otras palabras, un parámetro grande de discriminación puede distinguir mejor entre aquellas personas con bajos o altos niveles de rasgos latentes.

A través de un modelo TRI, se pueden estimar los parámetros de discriminación y dificultad, para cada reactivo de un instrumento diseñado para medir alguna habilidad latente. Para matemáticas, se asumirá un solo rasgo latente para explicar el comportamiento de la respuesta de una persona a los reactivos e independencia local de los reactivos (las respuestas de un reactivo son independientes a las respuestas de otros reactivos). Para lectura, se asumirá un diseño de reactivos agrupados en super ítems en dos niveles.

Para matemáticas, se expresa una forma funcional de la probabilidad de éxito:

$$\Pr(\text{éxito}|a, b, \theta) = F\{a(\theta - b)\}$$

La diferencia $(\theta - b)$ indica que la probabilidad de éxito en un reactivo es una función de la distancia entre la ubicación del reactivo y la ubicación de la persona. Cuando $\theta = b$, el individuo tiene la misma probabilidad de acertar o no acertar el reactivo. Cuando $\theta > b$ el individuo tiene una mayor probabilidad de acertar el reactivo. Como se puede obtener la misma distancia con diferentes valores de θ y b , se asume una métrica para θ para identificar el modelo. Entonces, se asume que $\theta \sim N(0,1)$, lo cual pone al parámetro de dificultad en la misma escala que una distribución normal. Así, los reactivos con

dificultades negativas se consideran relativamente fácil, y los reactivos con dificultades positivas se consideran relativamente difíciles.

Modelo Logístico de Un Parámetro (ML1P)

Sea Y_{ij} el resultado (por verse) del reactivo i del estudiante j , y sea y_{ij} el valor observado de Y_{ij} . Los resultados de Y_{ij} pueden ser correctos o incorrectos. También, $y_{ij} = 1$ se considera como correcto; $y_{ij} = 0$ como incorrecto.

Según la Teoría de Respuesta al Ítem (IRT en inglés), la probabilidad de un estudiante j con un nivel de característica latente θ_j que provee una respuesta correcta al reactivo i está dada por:

$$\Pr(Y_{ij} = 1 | a, b_i, \theta_j) = \frac{\exp\{a(\theta_j - b_i)\}}{1 + \exp\{a(\theta_j - b_i)\}}$$

Donde a representa la discriminación y b_i la dificultad del reactivo i . El modelo logístico de un parámetro ajusta el modelo mediante la forma pendiente-intercepto, por lo que la probabilidad para responder correctamente está parametrizada como:

$$\Pr(Y_{ij} = 1 | \alpha, \beta_i, \theta_j) = \frac{\exp\{\alpha(\theta_j - \beta_i)\}}{1 + \exp\{\alpha(\theta_j - \beta_i)\}}$$

La transformación entre estas parametrizaciones es:

$$a = \alpha; b_i = -\frac{\beta_i}{\alpha}$$

Sea $p_{ij} = \Pr(Y_{ij} = 1 | \alpha, \beta_i, \theta_j)$ y $q_{ij} = 1 - p_{ij}$. Se asume que las respuestas a los reactivos son independientes, condicional a θ_j , entonces, la densidad condicional del estudiante j está dada por:

$$f(y_j | \mathbf{B}, \theta_j) = \prod_{i=1}^I p_{ij}^{y_{ij}} q_{ij}^{1-y_{ij}}$$

Donde $y_j = (y_{1j}, \dots, y_{Ij})$, $\mathbf{B} = (\alpha, \beta_1, \dots, \beta_I)$, e I es el número de reactivos, y no se consideran valores faltantes en la ecuación de producto. La verosimilitud para el estudiante j se calcula mediante la integral de la variable latente desde la densidad conjunta:

$$L_j(\mathbf{B}) = \int_{-\infty}^{\infty} f(y_j | \mathbf{B}, \theta_j) \phi(\theta_j) d\theta_j$$

Donde $\phi(\cdot)$ es la función de densidad de probabilidad para la distribución normal. La log-verosimilitud para la estimación es la suma de las log-verosimilitudes de las N personas en la muestra.

$$\log L(\mathbf{B}) = \sum_{j=1}^N \log L_j(\mathbf{B})$$

La integral para $L_j(\mathbf{B})$ no es manejable, por lo que se utilizan métodos numéricos.

Formulación bayesiana del ML1P

Sea Y_{ij} el resultado (por verse) del reactivo i del estudiante j , y sea y_{ij} el valor observado de Y_{ij} . Los resultados de Y_{ij} pueden ser correctos o incorrectos. También, $y_{ij} = 1$ se considera como correcto; $y_{ij} = 0$ como incorrecto.

Según la Teoría de Respuesta al Ítem (TRI), la probabilidad de un estudiante j con un nivel de característica latente θ_j que provee una respuesta correcta al reactivo i está dada por:

$$\Pr(Y_{ij} = 1 | a, b_i, \theta_j) = \frac{\exp\{a(\theta_j - b_i)\}}{1 + \exp\{a(\theta_j - b_i)\}}$$

Las habilidades del estudiante θ_j se asumen estandarizadas, de tal forma que:

$$\theta_j \sim N(0,1)$$

Recordando el Teorema de Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_A P(B|A)P(A)}$$

Donde $P(A|B)$ es la probabilidad posterior de que un evento A, dado un evento B (la probabilidad de tener un reactivo correcto, dado el resultado del examen); $P(A)$ como la probabilidad previa de A (la probabilidad de tener un reactivo correcto); y $P(B|A)$ como la probabilidad condicional de B dado A (la probabilidad de un examen en particular, dado la presencia un reactivo correcto o incorrecto). El operador de suma representa la acumulación a lo largo de todos los posibles resultados del evento A, que puede ser tomado como la probabilidad de B. Al final, el Teorema de Bayes ofrece una representación de la probabilidad condicional de un evento dado otro.

Para Kim & Bolt (2007):

La formulación bayesiana [de un modelo de TRI] permite describir la probabilidad condicional de una respuesta a unos reactivos (B) dado unos parámetros del modelo A (discriminación, dificultad y habilidad latente). Tras algunas transformaciones que no permiten considerar a la probabilidad como valores continuos, entonces, conviene considerar al Teorema de Bayes en una forma continua: las funciones de densidad de probabilidad, las cuales representan la relativa verosimilitud de cada resultado. En términos de una función normal de densidad de probabilidad:

$$f(\Omega|X) = \frac{f(\Omega|X) \times f(\Omega)}{\left[\int_{\Omega} f(X|\Omega)f(\Omega)d\Omega \right]}$$

Donde X denota todos los datos de respuesta para cada reactivo; Ω representa a todos los parámetros desconocidos (discriminación del examen, dificultad de cada reactivo y habilidad latente de la persona). El uso de $f(\cdot)$ da cuenta de la naturaleza continua de los valores de los parámetros. La densidad conjunta posterior (de los parámetros del modelo dado los datos) se usa para determinar los estimadores de los parámetros del modelo. Para evaluar, se requiere conocimiento sobre las cantidades. La cantidad $f(X|\Omega)$, la cual expresa la verosimilitud de la información de las respuestas de los reactivos dado todos los parámetros del modelo es el modelo logístico de un parámetro, asumiendo supuestos de independencia local entre los reactivos. La cantidad $f(\Omega)$ es la densidad previa de los parámetros del modelo; se puede pensar como un indicador de la relativa verosimilitud de valores de parámetros en particular, previo a la recolección de información. La cantidad en el denominador se escribe en términos de integración⁴ de la distribución condicional de la información dados los parámetros sobre

⁴ Este tipo de integrales puede resultar incalculable, debido a la forma del dominio Ω a las dimensiones de Ω , a la complejidad de las funciones de verosimilitud y la densidad previa de Ω . El extremo Ω sobre su dominio se vuelve también complejo (Robert, 2016).

el espacio de los parámetros, y es una constante para un conjunto de datos fijos. Como el valor a menudo no se conoce o no es fácil de determinarlo, se escribe que:

$$f(\Omega|X) \propto f(X|\Omega) \times f(\Omega)$$

para indicar que la densidad conjunta posterior es proporcional al producto de las cantidades del lado derecho de la ecuación.

Para implementar, se requieren especificaciones previas para σ^2 y b_i . El parámetro de discriminación a se asume como positivo y modelado en una escala logística. Como no existe información previa sobre la discriminación o parámetros de dificultad, se asumen distribuciones previas de $\ln(a)$ y b_i que tienen soporte en la línea real, simétricas y centradas en 0. Entonces, hace sentido usar una distribución normal como distribución previa.

$$\ln(a) \sim N(0,1)$$

$$b_i \sim N(0,1)$$

La implementación de ML1P Bayesiano utiliza una especificación de una ecuación no lineal para la variable de respuesta y , en términos logit. La forma funcional para la probabilidad de éxito:

$$\Pr(Y_{ij} = 1) = y = f\{a(\theta_i - b_i)\}$$

Para hacer la especificación computacionalmente más eficiente, conviene tratar a b_i como parámetro de efectos aleatorios.

Luego, se utiliza el algoritmo adaptativo de caminata aleatoria Metropolis-Hastings para obtener la distribución posterior bajo el método de Montecarlo basado en cadenas de Markov (MCMC en inglés). MCMC crea muestras de una variable aleatoria con una densidad de probabilidad. El objetivo de MCMC es reproducir la distribución $f(\Omega|X)$ a partir de simulaciones suficientes como para obtener la media y la varianza. Una cadena de Markov es un método sistemático para generar una secuencia de valores en los que el valor actual depende en el valor anterior. El algoritmo Metropolis-Hastings utiliza un criterio de aceptación en el que el valor propuesto se acepta o se rechaza en la cadena de Markov (Brooks, Gelman, Jones, & Meng, 2011).

Para la implementación de MCMC, dado el número de observaciones, por el número de estudiantes alrededor de 16,000, se prevén entre 50,000 y 60,000 simulaciones y un periodo de quemado (*burn-in*) de 10,000 para permitir un mayor tiempo de adaptación del muestreo por MCMC, y dado el número de parámetros del modelo. Para el ajuste del modelo, debe permitirse una eficiencia de muestro por encima del 1%, como señal de convergencia del modelo. Esta implementación no genera el indicador de log-marginal verosimilitud debido a que el estimador Laplace-Metropolis no es posible calcularlo.

Referencias útiles para la evaluación

- Brooks, S., Gelman, A., Jones, G., & Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. Boca Raton: CRC Press. Taylor & Francis Group.
- Cao, Y., Lu, R., & Tao, W. (2014). Effect of Item Response Theory (IRT) Model Selection on Testlet-Based Test Equating. *ETS Research Report Series RR 14-19*. doi:10.1002/ets2.12017
- De Boeck, P., & Wilson, M. (2004). *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. New York: Springer.

- Hall, M., & Hampden-Thompson, G. (2022). The teacher as street-level bureaucrat: science teacher's discretionary decision-making in a time of reform. *International Journal of Science Education*, 44(6), 980-999.
- Kim, J., & Bolt, M. (2007). Estimating item response theory models using Markov chain Monte Carlo methods. *Educational Measurement: Issues and Practice* (26), 38-51.
- Lipsky, M. (2010). *Street-Level Bureaucracy. Dilemmas of the Individual in Public Services*. Nueva York: Russell Sage Foundation.
- Loyens, K., & Maesschalck, J. (2010). Toward a theoretical framework for ethical decision making of street-level bureaucracy: Existing models reconsidered. *Administration & Society*, 1(42), 66-100.
- Maynard-Moody, S., & Musheno, M. (2003). *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service*. Ann Arbor: University of Michigan Press.
- Muylaert, N. (2019). Contribuições da Ciência Política para o Campo Educacional: o burocrata de nível de rua. *Educação Online*, 14(31), 1-9.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Raykov, T., & Marcoulides, G. (2018). *A Course in Item Response Theory and Modeling with Stata*. College Station, Texas: Stata Press.