

Análisis de Gaps

Verificación de Contenidos Mínimos

Grupo 4

Pontificia Universidad Católica de Chile

EYP2417 - Muestreo
Grupo 4

Noviembre 2025

1. Diseño Muestral - Estado General

Resumen de Completitud

Estado de verificación de los 4 componentes del diseño muestral

Componente	Estado	Completitud
Tipo de Diseño		95 %
Tamaño de Muestra		70 %
Marco Muestral		90 %
Nivel de Inferencia		65 %

Completitud General

80 % - Requiere ajustes en 2 componentes

1.1 Tipo de Diseño - Análisis Detallado

Información Presente (95 %)

Diseño probabilístico claramente explicado

Estratificación geográfica y socioeconómica detallada

Diseño bietápico con etapas bien definidas

PPT en primera etapa y MAS en segunda etapa

Gap Menor Identificado (5 %)

Falta: Justificación explícita de por qué se eligió diseño bietápico vs. otras alternativas

Falta: Mención del efecto de diseño (DEFF) estimado

1.2 Tamaño de Muestra - Análisis Detallado

Información Presente (70 %)

Tabla con tamaño de muestra por nivel (nacional, urbano, rural)

Errores absolutos y relativos presentados

Mención de sobremuestreo

Criterios generales para tamaño de muestra

Gaps Identificados (30 %)

Falta: Explicación de cómo se calcularon los errores (fórmulas)

Falta: Justificación del nivel de confianza usado (95 %?)

Falta: Tasa de respuesta esperada vs. obtenida

Falta: Impacto de no respuesta en precisión

1.3 Marco Muestral - Análisis Detallado

Información Presente (90 %)

Mención del MMV 2020 como base

335 comunas identificadas

Estratificación por geografía, área y NSE

Conexión con ENE 2020

Gaps Menores (10 %)

Falta: Fecha de actualización del marco (¿2020 está actualizado para 2022?)

Falta: Número total de UPM en el marco

Falta: Cobertura del marco (¿qué % de la población?)

1.4 Nivel de Inferencia - Análisis Detallado

Información Presente (65 %)

Objetivo: viviendas particulares ocupadas

Exclusión de viviendas no elegibles claramente listada

Mención de áreas especiales excluidas (introducción)

Gaps Importantes (35 %)

Falta: Niveles específicos de inferencia (¿nacional? ¿regional? ¿comunal?)

Falta: ¿Se puede inferir a nivel de zona urbana/rural por región?

Falta: Limitaciones de inferencia explícitas

Falta: ¿Qué % de la población queda excluida?

2. Plan de Análisis - Estado General

Resumen de Completitud

Estado de verificación de los 3 componentes del plan de análisis

Componente	Estado	Completitud
Objetivos Específicos		85 %
Variables Involucradas		90 %
Métodos Propuestos		70 %

Completitud General

82 % - Requiere ajustes en métodos propuestos

2.1 Objetivos Específicos - Análisis Detallado

Información Presente (85 %)

Objetivo 1: Brecha salarial de género claramente definido

Objetivo 2: Distribución de pobreza bien articulado

Ambos objetivos son medibles y específicos

Gaps Identificados (15 %)

Falta: Conexión explícita entre objetivos y diseño muestral

Falta: ¿Por qué estos objetivos son importantes para CASEN?

Falta: Hipótesis de investigación explícitas

2.2 Variables Involucradas - Análisis Detallado

Información Presente (90 %)

Objetivo 1 - Brecha Salarial:

Variable dependiente: `ytrabajocorh`

Variables principales: `sexo`, `esc`, `edad`, `oficio4_08`

Variable estructural: `tot_per_h`

Objetivo 2 - Pobreza:

Variable dependiente: `pobreza`

Variables principales: `zona`, `ytotcorh`, `esc`, `región`, `edad`

Ponderador: `expr`

Gap Menor (10 %)

Falta: Descripción de escalas de medición (continua, categórica, ordinal)

2.3 Métodos Propuestos - Análisis Detallado (1/2)

Información Presente (70 %)

Objetivo 1:

Análisis descriptivo mencionado

Regresión lineal ponderada propuesta

Uso de factores de expansión mencionado

Objetivo 2:

Cálculo de porcentajes de pobreza

Comparación de ingresos promedio

Pruebas de hipótesis ponderadas

Regresión logística propuesta

Gaps Importantes (30 %)

Falta: ¿Cómo se incorpora el diseño complejo en las regresiones?

Falta: ¿Qué paquetes de R/Python se usarán para cada método?

Falta: ¿Cómo se calculan errores estándar robustos?

Falta: Nivel de significancia para pruebas de hipótesis

Falta: ¿Cómo se manejan valores missing?

Falta: ¿Se harán ajustes por comparaciones múltiples?

3. Metodología Detallada - Estado General

Resumen de Completitud

Estado de verificación de los 5 componentes de metodología detallada

Componente	Estado	Completitud
Estimadores Usados		85 %
Varianzas		75 %
Ponderadores		90 %
Software		95 %
Limitaciones		0 %

Completitud General

69 % - ¡GAP CRÍTICO en limitaciones!

3.1 Estimadores Usados - Análisis Detallado

Información Presente (85 %)

Estimador de Horvitz-Thompson presentado

Fórmula matemática correcta

Notación bien definida (y_i , π_i , $expr_i$)

Aproximación con factor de expansión mostrada

Explicación de componentes incluida

Gaps Menores (15 %)

Falta: Relación entre π_i y $expr_i$ (¿son inversos?)

Falta: ¿Cómo se obtienen las probabilidades de inclusión?

Falta: Estimador para totales además de medias

3.2 Varianzas - Análisis Detallado

Información Presente (75 %)

Fórmula de varianza presentada

Consideración del efecto del diseño mencionada

Estratificación incluida en fórmula

Corrección por población finita ($1 - n/N$)

Mención de intervalos de confianza

Gaps Identificados (25 %)

Falta: ¿Cómo se incorpora la conglomeración en la varianza?

Falta: Método de estimación de varianza (linearización, bootstrap, jackknife?)

Falta: ¿Cómo se calculan grados de libertad?

3.3 Ponderadores - Análisis Detallado

Información Presente (90 %)

Factor de expansión **expr** claramente identificado

Justificación: corrección de probabilidades desiguales

Integración en estimador de Horvitz-Thompson

Uso en análisis de regresión mencionado

Alertblock destacando importancia de ponderadores

Gaps Menores (10 %)

Falta: ¿Los ponderadores incluyen ajustes por no respuesta?

Falta: Rango de valores del factor de expansión

3.4 Software - Análisis Detallado

Información Presente (95 %)

Paquetes de R:

- survey y `srvyr` - diseño complejo
- `ggplot2` - visualización
- `dplyr` - procesamiento

Paquetes de Python:

- pandas - manipulación
- numpy - cálculos numéricos
- matplotlib y `seaborn` - visualización
- statsmodels - análisis estadístico

Gap Menor (5 %)

Falta: Versiones específicas de software/paquetes

3.5 Limitaciones - Análisis Detallado

GAP CRÍTICO (0 % de información)

NO HAY NINGUNA DIAPOSITIVA SOBRE LIMITACIONES

Limitaciones que DEBEN ser agregadas:

Limitaciones del diseño muestral:

- Exclusión de áreas especiales
- No inferencia a nivel comunal en todos los casos
- Efecto del diseño aumenta varianza vs. MAS

Limitaciones de los datos:

- Sesgo de no respuesta
- Datos autoreportados (subjetividad)
- Missing values en variables clave

3.5 Limitaciones - Análisis Detallado (cont.)

Limitaciones que DEBEN ser agregadas (continuación):

Limitaciones metodológicas:

- Causalidad vs. asociación en análisis de regresión
- Variables confundidoras no incluidas
- Supuestos de los modelos estadísticos

Limitaciones de inferencia:

- Representatividad solo para viviendas particulares
- Datos de corte transversal (no se puede analizar cambios temporales)
- Intervalos de confianza pueden ser amplios en subgrupos pequeños

Acción Requerida

CREAR UNA O DOS DIAPOSITIVAS DE LIMITACIONES

4. Criterios de Evaluación - Estado General

Resumen de Completitud

Evaluación contra criterios del curso

Criterio	Estado	Evaluación
Claridad Exposición Oral		Bueno
Comprensión Diseño		Muy Bueno
Rigor Metodológico		Regular
Calidad Diapositivas		Muy Bueno

Evaluación General

Bueno con reservas - Rigor metodológico afectado por falta de limitaciones

4.1 Claridad en Exposición Oral

Fortalezas

- Texto conciso en diapositivas
- Puntos clave bien resaltados con colores
- Uso de bloques para organizar información
- Buena estructura con secciones claras
- Transiciones lógicas entre temas

Áreas de Mejora

- Algunas diapositivas tienen mucho texto (especialmente metodología)
- Falta estimación de tiempo por sección
- Considerar agregar más ejemplos visuales

4.2 Comprensión del Diseño Muestral

Fortalezas

- Explicación clara de diseño bietápico
- Buena descripción de estratificación
- Conexión entre marco muestral y UPM
- Presentación de tamaño de muestra con tabla
- Mención de métodos de selección (PPT, MAS)

Áreas de Mejora

- Falta justificación de elecciones de diseño
- No se explicita nivel de inferencia específico
- Falta discusión de implicaciones del diseño para análisis

4.3 Rigurosidad y Coherencia Metodológica

Fortalezas

- Estimador de Horvitz-Thompson bien presentado
- Uso apropiado de ponderadores
- Métodos apropiados para objetivos
- Coherencia entre objetivos y variables
- Software apropiado identificado

Debilidades Importantes

FALTA COMPLETAMENTE: Sección de limitaciones

- Falta explicación de cómo se incorpora diseño complejo en regresiones
- Falta detalles sobre estimación de varianza con conglomeración
- No se discuten supuestos de los modelos

4.4 Calidad de las Diapositivas

Fortalezas

Diseño consistente con colores institucionales PUC

Excelente legibilidad de texto

Buen uso de ecuaciones matemáticas

Tablas bien formateadas con colores

Balance apropiado entre texto y espacio blanco

Uso efectivo de bloques de color

Logo institucional bien posicionado en portada

Mejora Menor

Considerar agregar más gráficos/diagramas visuales

Resumen Ejecutivo - Completitud por Sección

Sección Principal	Completitud	Estado
1. Diseño Muestral	80 %	Regular
1.1 Tipo de Diseño	95 %	Excelente
1.2 Tamaño de Muestra	70 %	Regular
1.3 Marco Muestral	90 %	Muy Bueno
1.4 Nivel de Inferencia	65 %	Regular
2. Plan de Análisis	82 %	Regular
2.1 Objetivos	85 %	Muy Bueno
2.2 Variables	90 %	Muy Bueno
2.3 Métodos	70 %	Regular
3. Metodología Detallada	69 %	Deficiente
3.1 Estimadores	85 %	Muy Bueno
3.2 Varianzas	75 %	Regular
3.3 Ponderadores	90 %	Muy Bueno
3.4 Software	95 %	Excelente
3.5 Limitaciones	0 %	Ausente

PRIORIDAD CRÍTICA (Requiere atención inmediata)

1. CREAR DIAPOSITIVA(S) DE LIMITACIONES

- Del diseño muestral
- De los datos
- Metodológicas
- De inferencia

2. Especificar niveles de inferencia exactos

- Nacional, regional, zona urbana/rural
- Limitaciones por tamaño de muestra

3. Explicar incorporación de diseño complejo en análisis

- ¿Cómo se usa **survey** en regresiones?
- Estimación de errores estándar robustos

Prioridades de Acción (continuación)

PRIORIDAD ALTA (Debe completarse antes de presentación)

4. Agregar explicación de cálculo de errores muestrales
5. Incluir tasa de respuesta y su impacto
6. Detallar método de estimación de varianza con conglomeración
7. Especificar nivel de significancia y manejo de missing values

PRIORIDAD MEDIA (Mejoras recomendadas)

8. Agregar justificación de elección de diseño bietápico
9. Incluir escalas de medición de variables
10. Agregar más visualizaciones (gráficos, diagramas)

Para mejorar la calidad de la presentación:

1. **Agregar 1-2 diapositivas de Limitaciones** después de Software
 - Es un requisito explícito de contenidos mínimos
 - Demuestra comprensión profunda del análisis
 - Aumenta credibilidad académica
2. **Mejorar conexiones explícitas** entre secciones
 - "Debido al diseño bietápico, usaremos..."
 - "Para incorporar la estratificación, el software..."
3. **Preparar respuestas a preguntas potenciales**
 - ¿Por qué Horvitz-Thompson y no otro estimador?
 - ¿Cómo afecta la conglomeración a la precisión?
 - ¿Qué limitaciones tiene el marco muestral?

Cronograma Sugerido

Prioridad	Tarea	Tiempo Estimado
Crítica	Crear diapositivas de limitaciones	2-3 horas
Crítica	Especificar niveles de inferencia	1 hora
Crítica	Explicar diseño complejo en análisis	1-2 horas
Alta	Agregar cálculo de errores	1 hora
Alta	Incluir tasa de respuesta	30 min
Alta	Detallar varianza con conglomeración	1 hora
Alta	Nivel de significancia y missing values	30 min
Media	Justificación de diseño	1 hora
Media	Escalas de medición	30 min
Media	Agregar visualizaciones	2 horas

Tiempo Total Estimado

11-13 horas para completar todos los gaps

Conclusión

Estado Actual

La presentación tiene una **base sólida** (77 % de completitud) con:

- + Excelente calidad visual
- + Buena explicación del diseño muestral
- + Software bien identificado
- + Objetivos claros

Gap Crítico

FALTA SECCIÓN DE LIMITACIONES - Requisito obligatorio

Con los ajustes sugeridos:

Se puede alcanzar **90-95 % de completitud** y obtener una excelente evaluación

¿Preguntas?

Grupo 4

Alexander Pinto, Esteban Román,
Julián Vargas, Francisca Sepúlveda

Este análisis de gaps fue generado automáticamente
basado en la verificación de contenidos mínimos