



parte 2

Diseño Metodológico

Antecedentes

La encuesta se aplicó entre el 2 de Agosto y el 30 de Septiembre de 2004 en todas las regiones del país, en virtud del “Convenio de Colaboración Interinstitucional

entre el Fondo Nacional de la Discapacidad y el Instituto Nacional de Estadísticas.

Actividades preparatorias

Elaboración conjunta de la primera versión del cuestionario.

Capacitación a los encuestadores que participaron en la prueba piloto del cuestionario.

Realización de una prueba piloto con el objeto de utilizar su experiencia en la preparación de la encuesta definitiva y del instrumento de recolección.

La prueba piloto se realizó en terreno, con 8 encuestadores durante dos días. Previo al trabajo de terreno se realizaron dos jornadas completas de instrucciones en

gabinete y posteriormente se realizó un focus group para evaluar la experiencia con el grupo de trabajo y obtener las conclusiones de ella. Estas últimas se centraron fundamentalmente en el cambio en la ubicación de algunas preguntas del cuestionario y en el mejoramiento del diseño y redacción de otras.

Elaboración del diseño definitivo del cuestionario.

Elaboración del manual del encuestador y supervisor.

Impresión de cuestionarios para la encuesta y de manuales del encuestador y supervisor.

Metodología de la encuesta

La encuesta utilizó la metodología de la entrevista a través de un informante válido, para recabar la información de todos los miembros del hogar principal de la vivienda seleccionada, a quienes se les aplicó el Registro de Personas del Hogar, (RPH) y el Módulo I.

El Módulo II se aplicó a una persona con discapacidad,

si la había, la que se determinó por el cumplimiento de ciertas características detectadas en la primera parte de la encuesta y seleccionada aleatoriamente dentro de las que cumplían estas condiciones. Esta parte se hacía cara a cara con el informante, solo si estaba en condiciones de responder por sus medios el Módulo II, en caso contrario las respuestas se solicitaban de un informante válido.

Diseño muestral

El diseño propuesto corresponde a una muestra probabilística, trietápica, estratificada en el área urbana y rural del país, correspondiente a una muestra seleccionada en tres etapas.

MARCO MUESTRAL

Para este estudio se tomó como base muestral el marco de áreas definido para el Programa Integrado de Encuesta de Hogares (PIDEH) donde la información utilizada, correspondió a las viviendas y población recogida en el Censo de Población y Viviendas 1992, con actualización permanente de secciones.

Considerando que la muestra consta de tres etapas, las unidades de muestreo son: los conglomerados geográficos o secciones, las viviendas particulares ocupadas y todas las personas que formen parte de las viviendas seleccionadas.

Unidad Primera Etapa: Secciones o conglomerados

Unidad Segunda Etapa: Viviendas particulares ocupadas

Unidad Tercera Etapa: Todos los miembros del hogar (Módulo I) y personas de 15 años y más (Módulo II)

POBLACIÓN EN ESTUDIO

La constituye las personas hombres y mujeres, residentes en los hogares que habitan en las viviendas particulares ocupadas, siendo seleccionada una de ellas para responder el módulo II de esta encuesta.

ÁMBITO GEOGRÁFICO Y ESTRATIFICACIÓN

La cobertura geográfica del estudio es nacional, urbano-rural. Para este estudio se consideran como estrato muestral el área urbana y rural en cada región. De esta forma fueron definidos dos estratos:

Estrato 1: Ciudades con población del área urbana

Estrato 2: Ciudades con población del área rural

POST ESTRATIFICACIÓN

Una vez realizado el levantamiento y elaborada la base datos se procedió a realizar una post estratificación o clasificación socioeconómica de los hogares que consiste en tomar algunas características de la vivienda, del hogar, y del jefe de hogar para generar un puntaje o índice, que corresponde a una medida ordenada del bienestar socio-económico relacionado con el hogar clasificado.

NIVEL DE ESTIMACIÓN

Dada las características del diseño y la validez estadística de la información a estimar, se tomó en consideración la importancia relativa de la población,

que habita el área urbana y rural de la ciudades, que presenten discapacidad. Los resultados serán entregados desagregados según variables más importantes en el área urbana a nivel nacional, regional, nacional urbano y nacional rural.

SELECCIÓN DE LAS UNIDADES MUESTRALES

La selección de la muestra se realizó en tres etapas:

- La selección de unidades de primera etapa (UPE) se efectuó en forma aleatoria y con probabilidad de selección proporcional al tamaño de cada sección de acuerdo al número de viviendas que tenían las mismas al Censo 1992.

La probabilidad de inclusión de la i -ésima unidad en la primera etapa es igual a:

$$p_{hi} = n_h \frac{M_{hi}}{M_h}, \text{ donde,}$$

h = representa el índice del estrato, $h = 1, 2$.

n_h = número de secciones en la muestra del estrato h .

M_{hi} = número total de viviendas en la sección i del estrato h al Censo 1992.

M_h = total de viviendas en el estrato h al Censo 1992.

- Luego, la selección de segunda etapa (USE), viviendas dentro de cada sección seleccionada, fue sistemática con probabilidad igual para las viviendas de cada sección.

$$p_{hij} = \frac{m_{hi}}{M'_{hi}}, \text{ siendo,}$$

m_{hi} = número de viviendas en la muestra de la sección i del estrato h .

M'_{hi} = número de viviendas actualizadas de la sección i del estrato h .

- Unidades de tercera etapa (UTE): las personas de 15 años y más con el método de selección aleatoria, tabla de Kish.

En el punto siguiente se presenta un listado de las 13 regiones que conforman el universo del estudio, con información de la muestra por sección según detalle.

TAMAÑO DE LA MUESTRA

Con base en la información del marco muestral del PIDEH y la composición de los nuevos estratos se seleccionó una muestra de 13.350 viviendas, correspondiendo a 11.000 unidades del área urbana y 2.350 del área rural.

$$n = \frac{k^2 * s^2}{e^2} * Deff = 13.350, \text{ siendo,}$$

k = corresponde al percentil de la normal asociado a un intervalo de confianza bilateral del 95% (1,96).

s^2 = varianza de la variable de interés.

e = es el error absoluto requerido para la precisión deseada.

$Deff$ = es el factor de ajuste llamado Efecto del Diseño.

Para alcanzar este tamaño se propone un adicional de 2.670 viviendas (20%) para cubrir los efectos de la no respuesta (actualización del marco, rechazos, etc.).

ERRORES DE MUESTREO

Nivel	13.350 viviendas			
Estimación	P = 0.7 E.A (%)	E.R (%)	P = 0.12 E.A (%)	E.R (%)
Nacional	0,6	8,6	0,8	6,4
Nac. urbano	0,7	9,4	0,8	6,9
Nac. rural	1,5	21,5	1,9	16,0
Región 1	2,9	41,9	3,7	31,2
Región 2	3,2	45,0	4,0	33,5
Región 3	3,1	43,8	3,9	32,5
Región 4	2,6	37,2	3,3	27,6
Región 5	2,0	28,4	2,5	21,1
Región 6	2,5	35,9	3,2	26,6
Región 7	2,3	33,1	3,0	24,6
Región 8	1,6	22,9	2,0	17,0
Región 9	2,4	34,6	3,1	25,7
Región 10	2,1	29,7	2,6	22,0
Región 11	3,2	45,1	4,0	33,5
Región 12	2,9	41,8	3,7	31,1
Región 13	1,0	14,7	1,3	10,9

Distribución de la Muestra

Región	Estrato	Muestra			
		Población Censo'92	Viviendas particulares Censo'92	Número de viviendas	Número de secciones
I	Urbano	378.528	72.228	430	43
	Rural	20.160	5.302	120	12
II	Urbano	453.344	88.724	450	45
	Rural	14.819	1.987	90	9
III	Urbano	248.824	50.309	430	43
	Rural	24.568	5.171	120	12
IV	Urbano	426.097	80.723	540	54
	Rural	151.462	33.335	140	14
V	Urbano	1.430.595	302.733	1.000	100
	Rural	130.454	30.957	140	14
VI	Urbano	534.425	102.499	510	51
	Rural	253.978	54.726	170	17
VII	Urbano	578.598	113.480	560	56
	Rural	336.312	72.197	230	23
VIII	Urbano	1.558.599	299.156	1.300	130
	Rural	376.840	76.384	300	30
IX	Urbano	577.317	107.527	550	55
	Rural	296.547	66.111	230	23
X	Urbano	695.910	124.773	680	68
	Rural	365.142	75.315	280	28
XI	Urbano	70.975	13.580	420	42
	Rural	24.007	4.243	120	12
XII	Urbano	145.803	34.165	430	43
	Rural	11.918	2.311	90	9
RM	Urbano	5.848.776	1.212.055	3.700	370
	Rural	249.776	50.107	320	32
Total				13.350	1.335

Descripción del cuestionario

Las preguntas del cuestionario, son en su mayoría precodificadas, esto implica que las alternativas de respuestas están preestablecidas para cada pregunta. La forma de responder es marcando una cruz en el recuadro que se ubica frente a la alternativa que corresponda a la respuesta del informante o completando la respuesta con un código que se incluye en el área que abarca la pregunta.

El cuestionario está organizado en tres partes, a saber:

Primera parte: Registro de Personas en el Hogar (RPH)

Segunda parte: Correspondiente al Módulo I de Identificación de Dificultades

Tercera parte: Correspondiente al Módulo II de Caracterización de Discapacidad

La primera de ellas, el Registro de Personas del Hogar (RPH), contiene preguntas para hacer una caracterización adecuada del hogar y para determinar las características sociodemográficas de sus integrantes, de acuerdo con los requerimientos de información de esta investigación.

MÓDULO I: IDENTIFICACIÓN DE LAS DIFICULTADES (34 PREGUNTAS)

El Módulo I de Identificación de Dificultades esta destinado a investigar y caracterizar diversas dificultades de cada uno de los integrantes del hogar, sin distinción de edad, por lo que se aplica a TODOS los miembros del Hogar, permite detectar a las personas que tienen algún tipo y grado de discapacidad, a través de 9 secciones:

- Sección 1.1** Revisión de salud general (cinco preguntas)
- Sección 1.2** Visión y audición (seis preguntas)
- Sección 1.3** Cuidado personal (tres preguntas)
- Sección 1.4** Capacidad para moverse (cuatro preguntas)

- Sección 1.5** Comprensión y comunicación (cinco preguntas)
- Sección 1.6** Actividades de la vida diaria (una pregunta)
- Sección 1.7** Dificultades para el desempeño de actividades vitales (una pregunta múltiple, que en la practica es equivalente a un total posible de 54 variables)
- Sección 1.8** Dificultades en aparatos y sistemas orgánicos vitales (una pregunta)
- Sección 1.9** Preguntas generales del hogar (ocho preguntas)

Esta última sección incluye preguntas cuyo objetivo principal es generar una clasificación socioeconómica de las personas encuestadas. Este conjunto de preguntas remite, fundamentalmente, a características de la vivienda y del jefe de hogar, sea este una persona con o sin discapacidad.

MÓDULO II: CARACTERIZACIÓN DE LAS PERSONAS CON DISCAPACIDAD (32 PREGUNTAS)

Se consulta a la(s) persona(s) seleccionadas aleatoriamente a partir del Módulo I de acuerdo a una tabla de selección previamente elaborada para este estudio por el INE.

- Sección 2.1** Capacidad para relacionarse con otras personas (cinco preguntas)
- Sección 2.2** Participación y aceptación social (siete preguntas)
- Sección 1.3** Deficiencias y dispositivos de apoyo (nueve preguntas)
- Sección 2.4** Servicios utilizados y necesitados (tres preguntas)
- Sección 2.5** Interrelaciones y derechos (ocho preguntas)

Selección del personal de terreno

Supervisores: Los supervisores fueron seleccionados del banco de supervisores y encuestadores que habitualmente trabajan en los proyectos especiales del INE, con experiencia en la Encuesta Nacional del Empleo y probada capacidad para manejar actividades operativas de terreno. El total de supervisores fue de 18 personas para trabajar en todas las regiones del país quienes se designaron previo análisis de antecedentes y entrevista personal.

Encuestadores: Se seleccionaron 41 encuestadores principalmente entre personas con experiencia en trabajos anteriores relacionados con la Encuesta Nacional del Empleo (ENE) y otras encuestas de hogares realizadas por el INE. El proceso de selección se realizó a través de una evaluación de antecedentes curriculares que cumplieran con el perfil establecido para desempeñarse como encuestador y una entrevista personal.

Capacitación

La capacitación se organizó en 4 sedes regionales para las 13 regiones del país.

- Sede Norte, en la Serena, que comprendió las regiones I, II, III y IV
- Sede Sur, en Puerto Montt, que comprendió las regiones X, XI y XII
- Sede Centro Sur, en Concepción, que abarcó las regiones VIII y IX; y
- Sede Central, en Santiago, que comprendió las regiones V, VI, VII y RM.

En cada una de las sedes se aplicó un programa de

trabajo de dos días de duración, estructurado en dos partes:

A los supervisores, además, se les instruyó en forma especial, sobre los siguientes temas:

- Etapas fundamentales de la supervisión: planificar, organizar, dirigir, controlar y evaluar.
- Áreas que comprende la supervisión: calidad, cobertura y oportunidad.
- Tareas que deben realizar los supervisores: control de desempeño, control de calidad, control de visitas, control de cobertura, evaluación de entrevista y revisión de cuestionarios.

Parte I	Parte II
1. Presentación y bienvenida	1. Uso de cuestionarios: Trabajo a realizar en oficina antes de salir a terreno.
2. Presentación participantes	2. Llenado del RPH. Uso de códigos. Informante
3. Programa de trabajo de las jornadas	3. Orden de anotación de los miembros del hogar
4. Presentación de la Encuesta de Discapacidad (FONADIS). Objetivos, antecedentes, conceptos, tipo de discapacidad, etc.	4. Módulo 1. Pregunta a pregunta
5. Conceptos y definiciones generales que se ocupan en las encuestas de hogares.	5. Informante con probabilidad de selección para aplicar módulo II.
6. Miembros y no miembros del hogar	6. Uso de la columna N° de orden seleccionado
7. Marco legal	7. Uso de tabla aleatoria
8. Organización del trabajo en secciones y viviendas seleccionadas	8. Módulo II. Pregunta a pregunta
9. COCA: Cobertura, Oportunidad y Calidad, en el diligenciamiento de encuestas y en la supervisión	9. Repaso general. Ejercicios
10. Función del supervisor	10. Comentarios y cierre
11. Descripción general del cuestionario	

Operativo de terreno

El operativo de terreno se realizó entre el 2 de Agosto y el 30 de Septiembre, abarcando un total de 43 días hábiles. La organización operativa en las regiones fue la siguiente:

1. Supervisor jefe de operaciones de terreno
2. Supervisores
3. Encuestadores

Para la Región Metropolitana, la organización fue diferente a las otras regiones del país en atención al tamaño de la muestra seleccionada, estructurándose de la siguiente forma:

1. Jefe de operaciones de terreno de la región metropolitana
2. Supervisores
3. Encuestadores

En total el número de supervisores fue de 18 quienes

tuvieron a su cargo un total de 41 encuestadores, lográndose una relación de supervisión nacional de 0,44 y en la RM de 0,25. Estos estándares se consideraron adecuados a los requerimientos de calidad del trabajo de relevamiento de información.

Esta dotación posibilitó contar con 1.763 días/persona de terreno, en días hábiles. A ello se agregan 287 días/persona correspondientes a trabajo realizado durante los fines de semana del período de levantamiento. Consecuentemente, estas cifras permiten determinar que para visitar las 16.020 viviendas seleccionadas en la muestra y lograr la cobertura final de 14.638 viviendas encuestadas, se trabajó un promedio de 7,8 viviendas por día/persona. La duración promedio de las entrevistas fue entre 50 y 60 minutos.

La cobertura final de la encuesta se ilustra en el siguiente cuadro:

Región	Secciones	Tamaño inicial de la muestra (viviendas)	Viviendas seleccionadas	Encuestables									No encuestables	
				Total	%	Logradas						No logradas		
						N° de Viv.	Cobertura		Viv./ Secc. (promedio)	N° de hogares	Viv.	%		
							% a encuest.	% a selec.						
I	55	550	660	592	99,5	589	99,5	89,2	10,7	589	3	0,5	68	
II	54	540	648	592	99,3	588	99,3	90,7	10,9	588	4	0,7	56	
III	55	550	660	603	98,7	595	98,7	90,2	10,8	595	8	1,3	57	
IV	68	680	816	740	95,3	705	95,3	86,4	10,4	705	35	4,7	76	
V	114	1140	1.368	1237	98,2	1.215	98,2	88,8	10,7	1.215	22	1,8	131	
VI	68	680	816	751	89,2	670	89,2	82,1	9,9	670	81	10,8	65	
VII	79	790	948	868	91,1	791	91,1	83,4	10,0	791	77	8,9	80	
VIII	160	1600	1.920	1748	99,1	1.732	99,1	90,2	10,8	1.732	16	0,9	172	
IX	78	780	936	842	89,0	749	89,0	80,0	9,6	749	93	11,0	94	
X	96	960	1.152	1050	94,0	987	94,0	85,7	10,3	987	63	6,0	102	
XI	54	540	648	556	98,4	547	98,4	84,4	10,1	547	9	1,6	92	
XII	52	520	624	579	72,4	419	72,4	67,1	8,1	419	160	27,6	45	
XIII	402	4020	4.824	4480	93,3	4.182	93,3	86,7	10,4	4.182	298	6,7	344	
Total	1.335	13.350	16.020	14.638	94,1	13.769	94,1	85,9	10,3	13.769	869	5,9	1.382	

En cuanto a la **tasa de logro** respecto del total de viviendas seleccionadas en la muestra, a nivel nacional alcanzó a un **85,9%**, es decir, de **16.020** viviendas seleccionadas en la muestra, se definieron como encuestables 14.638 viviendas de las que finalmente se encuestaron **13.769**. La diferencia corresponde a viviendas NO encuestables por alguno de los siguientes motivos, propios del marco muestral: viviendas inaccesibles, desocupadas temporalmente, en reparación, ampliación o remodelación, de temporada,

demolidas, incendiadas, erradicadas, destruidas o trasladadas y viviendas destinadas a usos distintos al de su finalidad original. Considerando esto, la tasa de logro se establece en un 94,1% valor que incrementa sustancialmente al 85,9% calculado sobre las viviendas seleccionadas originalmente.

Cabe destacar que el tamaño inicial determinado para la muestra fue de 13.350 viviendas las que se incrementaron en un 20% para corrección de no respuestas.

Gestión operativa de los supervisores

Los supervisores se abocaron a la realización integral de las siguientes acciones orientadas a conseguir calidad y cobertura en la encuesta

- **Control de desempeño** consistió en la revisión de la tarea diaria del encuestador en cuanto a: diligenciamiento oportuno de los formularios en cada uno de los hogares seleccionados y al logro de las entrevistas de todos los miembros del hogar correspondiente.
- **Control de calidad** cuyo objetivo se orientó a la verificación de la calidad y consistencia de la información recolectada por los encuestadores respecto a: miembros del hogar, población de 15 años y más, educación, actividades que realizan, uso del tiempo, etc.
- **Control de visitas** que tuvo como finalidad la verificación de la concurrencia directa y personal del encuestador a las viviendas que debió visitar y encuestar. Además, que cumpliera con las visitas concertadas para ubicar al informante indicado.
- **Control de cobertura** que se realizó conjuntamente con el control de visitas y tuvo como propósito comprobar

que las viviendas encuestadas correspondieran exactamente a las viviendas seleccionadas en la muestra. Además, de asegurar la efectiva determinación y clasificación de las viviendas no logradas según categoría A, B o C.

- **Evaluación de entrevistas** que tuvo como objetivo verificar el manejo del cuestionario por parte del encuestador. Específicamente en el conocimiento conceptual de éste, observancia de las secuencias, planteamiento de las preguntas a los informantes, determinación e inclusión de todos los miembros del hogar.
- **Revisión de cuestionarios** que implicó examinar en gabinete todos los formularios diligenciados para detectar y corregir en terreno, si era necesario, inconsistencias, omisiones, legibilidad y grado de detalle de ciertas preguntas. Comprendió verificación de las correcciones hechas por el encuestador en terreno generadas en la revisión de los formularios diligenciados, y comprobación que las viviendas y hogares a encuestados correspondieran efectivamente al directorio de la selección muestral.

Procesamiento de datos

PROCESAMIENTO MANUAL DE LA INFORMACIÓN (REVISIÓN Y CODIFICACIÓN)

El procesamiento de datos se realizó en la Sede Central del INE, en Santiago, donde se recibió el material de campo. Para la ejecución de estas actividades se trabajó de acuerdo a las normas internas del INE para las actividades de Ajuste, Revisión y Codificación que determinan las pautas necesarias para lograr excelencia en la calidad de los resultados finales. Además, se trabajó con el Manual C.I.U.O – 88, para clasificar todas las ocupaciones declaradas en la encuesta

Las actividades que se ejecutaron como parte del procesamiento manual, fueron las siguientes:

Recepción de cuestionarios

Cada carpeta de sección que se recibía para ser procesadas, contó con un formulario de control con el detalle de lo enviado. Esto, una vez revisado fue verificado computacionalmente con la base de datos de las secciones de la muestra, así cuando no fue enviado desde terreno el material correspondiente, rápidamente era detectado y solicitado lo que permitió controlar la entrega de material de acuerdo a la planificación.

Ajuste y codificación

Recibido el material de trabajo, se procedió a verificar cuidadosamente los antecedentes sobre “Identificación geográfica de la carpeta de sección”. Estos antecedentes debían coincidir en todos y cada uno de los cuestionarios que fueron ocupados en la encuesta, luego se verificó que las viviendas, logradas o no, estuvieran debidamente registradas e identificadas en la contratapa de la carpeta de sección. También, se revisó que efectivamente fueran devueltos de terreno los documentos utilizados como: directorios, mapas, croquis, formulario 61 – 62 y el informe de la sección.

Una vez que terminó esta primera parte del ajuste, se procedió a analizar y codificar el material con apoyo de los manuales e instructivos en la segunda parte del ajuste del material, efectuándose el proceso de foliación.

Revisión y foliación

Desde un comienzo del proceso se asignó a una persona con experiencia para que revisara el material ya codificado, además la Jefe de la Unidad, procedió a revisar aleatoriamente el material, para asegurar la calidad del producto. Después que el material era ajustado (en su primera parte), codificado y revisado, era necesario concluir con el proceso de ajuste y foliar los cuestionarios, todo esto debidamente controlado en los formularios diseñados para ello.

INGRESO DE DATOS Y VALIDACIONES PRELIMINARES

La digitación se realizó simultáneamente con la codificación y revisión, a través de 3 equipos, en los cuales estaba instalada la aplicación para el Ingreso de Datos.

A los digitadores se les asignaba diariamente una Unidad de Trabajo. Esta consistía en una cantidad de cuestionarios debidamente revisados, foliados y codificados, correspondientes a una o varias secciones de trabajo de terreno. Del trabajo diario de cada digitador, se efectuaban respaldos.

La aplicación de Ingreso de Datos contó entre otras cosas con sistemas antivirus, generación diaria de copias de seguridad, además de un registro de transferencia que permitió llevar un control estricto de lo que se digitaba, individualizando a cada uno de los equipos. El programa fue dotado de capacidad para detener el avance si se digitaba algún código que no correspondía, o no demostraba consistencia, (omisión de marcas o datos faltantes) o secuencia en las respuestas.

PROCESAMIENTO DE LA ENCUESTA

El sistema computacional para el procesamiento de la encuesta se diseñó en base al documento términos de referencia que es el documento básico con todas las especificaciones y necesidades iniciales consideradas para el proyecto. Posteriormente en base a reuniones se precisan las especificaciones y se refuerza el diseño

considerando la necesidad de contar con aplicaciones específicas orientadas a cada una de las necesidades del proceso.

Algunas de las aplicaciones del Sistema para el procesamiento de Datos de la Encuesta Nacional sobre Discapacidad fueron:

- Ingreso de Datos y Validaciones on line
- Validación de Datos Batch
- Reportes de errores en formato Cristal Report y Planilla Excel
- Generación de Variables agregadas
- Expansión de Datos
- Generación de Base de Datos en SPSS
- Tabulaciones
- Calculo Coeficientes de Variación
- Informe de Total de Hogares y Personas por Estrato y acumulados

Ingreso de Datos y Validaciones on line

Permitió el ingreso y la modificación de los datos de las encuestas físicas a la base de datos, previamente se aplicaron las codificaciones y revisiones respectivas a los formularios.

Se confeccionó un programa para el ingreso de datos. A este se le incorporó reglas que permitía detectar algunas inconsistencias en el momento del ingreso de la información, que básicamente consideró validaciones de rango y de consistencia para algunas variables.

La aplicación no fue necesaria instalarla en los PC para la digitación, ya que por la tecnología realizada se podía acceder a la aplicación de entrada de datos por medio de una dirección URL. Los datos digitados se grabaron directamente en los Servidores centrales de la institución y fueron respaldados diariamente aplicando las políticas de respaldo existente en el INE.

Validación de datos Batch

La validación de datos tiene por finalidad depurar la base de datos según los criterios establecidos para que existan datos fidedignos y sin errores en la base de datos unitaria. Esto permite obtener resultados óptimos en los procesos posteriores de tabulación, expansión, etc.

El programa de validación Batch (fuera del proceso de ingreso de datos) incorporó adicionalmente a las reglas básicas otro conjunto de criterios de consistencia de datos. La aplicación validó la base de datos central y emitió un informe de errores para facilitar el proceso de corrección de información.

Reportes de errores en formato Cristal Report y Planilla Excel

La aplicación de validación de datos Batch exploró errores en los datos centrales y al encontrar errores los informaba por medio de un informe en Cristal Report en WEB y permitió la posibilidad de generar el mismo informe en una planilla Excel, para dar las mejores alternativas para la corrección de datos del sistema.

Generación de variables agregadas

Los datos originales por si mismo, no necesariamente permiten trabajar adecuadamente con la información, por esta razón, por necesidades del sistema y para permitir mayor facilidad en los siguientes procesos intermedios y finales es que se generaron un conjunto de variables agregadas, entre los que se cuentan tramos de edad, clasificador socio económico, factor de expansión, categoría ocupacional, región, área, entre otras.

Expansión de Datos

Se construyó una aplicación que calcula los factores de expansión, utilizando la metodología de factores ajustados al total de población en el estrato, con corrección y ajuste por sexo, lo que implica que se calcula un factor de expansión por sección.

Para implementar la metodología de cálculo de los factores, se requirieron tablas adicionales a los datos de la encuesta, como proyecciones de población por sexo, con totales por estratos y número de viviendas por sección.

También se aplicó el algoritmo que permite ajustar el factor de expansión a un valor entero, para evitar las distorsiones que el valor inicial, que contiene una extensión decimal, provoca al ser aplicado directamente. En este caso, se ocupó la misma metodología de la Encuesta Nacional de Empleo, que permite asignar a cada persona un factor de expansión entero.

Generación de Base de Datos en SPSS

Por requerimientos del Sistema, se debía entregar la Base de datos final en SPSS. Para esto se preparó la base integral de datos de la encuesta y se aprovecharon funciones y facilidades de este software estadístico para entregar la información en esta herramienta. También se utilizó este software para analizar los datos y comprobar resultados de los tabulados.

Tabulaciones

Para la generación de tabulados de la encuesta se utilizó la aplicación del software SPSS. Este permite que los cuadros generados sean traspadados a planillas en formato Excel.

Calculo de Coeficientes de Variación

Para calcular algunos tabulados con coeficientes de variación se programaron varias aplicaciones que permitieron obtener los formatos y resultados esperados.

Las aplicaciones permitieron la acumulación de valores necesarios a nivel de muestra, estimación y varianza por las distintas agrupaciones de interés.

Informe de Total de Hogares y Personas por Estrato y acumulados

Para mantener el mejor control de los datos digitados y gestionar adecuadamente esta crucial etapa de producción, se preparó una aplicación Web que permitía que en cualquier instante por medio de una dirección URL se conociese el total de datos digitados por estratos y el total general de información depositada en la base de datos central de la encuesta.

Metodología de estratificación socio-económica

La clasificación socioeconómica busca categorizar a cada hogar y, considerando el conjunto de características que lo describen, asignarlo a alguna clasificación que ha sido previamente definida de acuerdo a una serie de atributos.

Para esta encuesta, se determinó implementar una metodología que permitiera clasificar los hogares de acuerdo a un conjunto de variables socioeconómicas, teniendo presente que el óptimo era utilizar información de variables relacionadas con la vivienda, nivel educacional del jefe de hogar y ocupación de este.

Es importante señalar que una clasificación objetiva de los hogares debe situarse en el contexto que existen otras variables que permiten una clasificación más rigurosa, y que permite distinguir a dos hogares, que presentando el mismo nivel de ingresos por ejemplo, pueden ser clasificados en distintas categorías, ya sea por el tipo de vivienda, el nivel educacional del jefe de hogar, el número de personas integrantes del hogar, etc.

Variables seleccionadas: prácticamente en casi to-

dos los estudios sobre clasificación socioeconómica las variables que realmente permiten discriminar a los hogares son las mismas: las relacionadas con la vivienda (tipo de vivienda, número de piezas, número de piezas exclusivas para dormir, etc.), y las relacionadas con el hogar y jefe de hogar (años de estudio, nivel cursado, ocupación, nivel de ingresos, hacinamiento, número de personas del hogar, etc.).

MÉTODO UTILIZADO PARA LA ESTRATIFICACIÓN

El método utilizado fue el de COMPONENTES PRINCIPALES NO LINEAL –PRINCALS–. Pueden existir muchos modelos que pueden aplicarse para clasificar estadísticamente los hogares, pero el modelo a elegir dependerá exclusivamente de la información disponible, de encuestas complementarias, etc. Más aún, en el caso de esta encuesta, dado el tipo de variables, y el no contar con una clasificación a priori para modelar la variable respuesta, se utilizó el método de componentes principales no lineal, que se explica.

PRINCALS es un acrónimo para designar los “Principal components analysis by means of Alternating Least Squares”. (Principales componentes de análisis por medio de alternar los mínimos cuadrados). El input de este procedimiento lo constituye una matriz de datos en la que las filas están formadas por objetos/sujetos y las columnas por variables. La diferencia respecto al análisis de componentes principales tradicional, radica en que este procedimiento puede llevar a cabo un análisis en que las diferentes variables pueden tener una relación no lineal y estar medidas en diferentes escalas (nominales, ordinales, de intervalo y de cuociente), en tanto que en el estándar la relación es lineal y las variables deben estar medidas en una escala de intervalo o de cuociente.

El proceso que sigue el PRINCALS y la interpretación de sus resultados es conseguir cuantificaciones de los objetos/sujetos y, por tanto, de las categorías de las variables que sean óptimas en el sentido que las categorías estén lo más separadas unas de otras en las dimensiones estudiadas, y a su vez, dentro de cada categoría los sujetos estén lo más próximos unos a otros, es decir, con puntuaciones cuanto más homogéneas entre sí.

Dentro de este procedimiento existen varias alternativas como el número de dimensiones, que hemos dejado para asignarle puntuaciones óptimas a los sujetos/objetos, a su vez hemos elegido la opción de normalización simétrica, en el sentido que maximiza la distancia entre las categorías de las variables, al mismo tiempo que maximiza las distancias entre los sujetos (hogares) disímiles.

RESULTADOS

Una vez obtenidos los puntajes, se ajustaron de modo que el mínimo fuera cero, con lo que se obtiene finalmente un coeficiente para cada variable utilizada en la estratificación.

La cuantificación de cada categoría de las variables se multiplica por estos coeficientes o ponderadores de variables, para así obtener un puntaje que se asigna a un sujeto de acuerdo a si pertenece o no a la categoría respectiva. Luego, la suma sobre todos los puntajes de las variables asociadas a un sujeto en particular, es el puntaje final del sujeto, en un espacio multidimensional único,

que en nuestro caso, al trabajar con una dimensión, es la proyección sobre un único eje, tanto de las variables como de los sujetos (hogares) de modo que podemos ver las relaciones entre las variables y sujetos de una manera sencilla.

CATEGORIZACIÓN

Para la categorización se utilizaron los siguientes niveles, independientes o agrupados de tal manera que constituyan un reflejo fiel de la realidad socioeconómica nacional.

Nivel socioeconómico	Categoría
Bajo	1
Medio bajo	2
Medio	3
Medio alto	4
Alto	5

Los puntajes citados anteriormente son consistentes en cuanto a la fiabilidad de clasificación. Más aún el índice de la escala de estos puntajes, denominado Índice de Crombach corresponde aproximadamente a 0,8. Los valores de este coeficiente oscilan entre 0 y 1 y su valor depende tanto de las variables involucradas en la escala como de la correlación entre las mismas variables o de sus varianzas/covarianzas, de modo que se puede aspirar a tener un coeficiente de fiabilidad muy elevado aún partiendo de los mismos valores promedio de correlaciones, varianzas o covarianzas y simplemente incrementando el número de variables o ítems con los que se genera la escala o puntaje. Se considera una escala o puntaje o índice consistente, si el coeficiente de alpha de Crombach es de por lo menos 0,5 en el rango de 0 a 1, por lo tanto, a valores más cercanos a uno le corresponde una fiabilidad mejor.

Este Coeficiente alpha de Crombach es uno de los más utilizados a la hora de establecer la fiabilidad de una escala, y está basado en la consistencia interna de la misma. Se obtiene como promedio de los coeficientes de correlación de Pearson entre todos los ítems (variables) de la escala si las puntuaciones de los mismos están estandarizadas, o como promedio de las covarianzas, si no lo están.

Factores de expansión

El factor de acuerdo con el diseño, corresponde al aplicado a una muestra de dos etapas, asignándole un factor de expansión a cada hogar seleccionado y a cada persona seleccionada que depende del número de personas que tiene la vivienda y del número de viviendas que tiene la sección por estrato.

El factor de expansión se puede interpretar como la cantidad de personas en la población que representa una persona en la muestra. Este factor puede ser distinto para cada persona, y esa diferencia proviene del hecho que la muestra no es autoponderada.

Por otra parte, se incluye un ajuste de población de personas, por sexo, en cada una de las áreas para aumentar la precisión de las estimaciones.

Factores a utilizar:

$$FE_i^{(1)} = \frac{M_h}{M_{hi} * n_h} * \frac{M'_{hi}}{m_{hi}} =$$

Factor de Expansión para estimar personas del estrato h en la sección i , sin ajuste de población total. Siendo:

M_h = Número de viviendas en el estrato h al Censo de Población y Vivienda de 1992.

M_{hi} = Número de viviendas en la sección i del estrato h al Censo 1992.

M'_{hi} = Número de viviendas actualizadas en la sección i del estrato h .

m_{hi} = Número de viviendas en la muestra en la sección i del estrato h .

n_h = Número de secciones en la muestra del estrato h .

$$FE_{hi}^{(2)} = \frac{M_h}{M_{hi} * n_h} * \frac{M_{hi}}{m_{hi}} * \frac{P_h}{\hat{P}_h} =$$

Factor de expansión de hogares y personas del hogar en el estrato h en la sección i , con ajuste de población total. Donde:

$$\hat{P}_h = \sum FE_i^{(1)} * \frac{P_{hij}}{p_{hij}} =$$

Población estimada de personas en el estrato h . Siendo:

P_{hij} = Número de personas en la vivienda j de la sección i en el estrato h que tienen 15 años y más.

p_{hij} = Una Persona en la muestra de vivienda $j \geq 15$ años.

P_h = Proyección de población total del estrato h .

$\hat{P}_h$ = Estimación de la población total del estrato h .

La estimación de un total dado para una variable, se obtiene, primero, multiplicando el valor de la variable en cada persona por su factor de expansión y luego sumando sobre todas las personas de la muestra.

Estimadores y varianza de los estimadores

ESTIMADORES

Para obtener el valor estimado a nivel de área geográfica, regiones o total país, se multiplica el Factor de Expansión por persona, FE, por el valor de la variable en la persona y luego sume esos valores para todas las personas que tienen el atributo en estudio.

Nomenclatura utilizada para los estimadores:

FE_{hij} = Factor de Expansión por persona en la vivienda “ j ”, sección “ i ”, en el estrato “ h ”.

Si se estima personas y hogares entonces $FE_i^{(2)}$.

Y_{hijc} = Valor de la característica medida en la persona de la vivienda “ j ”, sección “ i ”, estrato “ h ”, que cumple con la característica “ c ”.

P_{hijc} = Persona de la vivienda “ j ”, sección “ i ”, estrato “ h ”, que cumple con la característica “ c ”.

Total estimado de la variable “y”, para el nivel de estimación requerido

$$\hat{Y}_c = \sum_h \sum_i \sum_j FE_{hij} \cdot [y_{hijc}], \text{ siendo,}$$

y_{hijc} , el valor de la variable en la persona.

Total estimado de las personas que tienen el atributo en estudio para el nivel de estimación requerido

$$\hat{P}_c = \sum_h \sum_i \sum_j FE_{hij} \cdot [p_{hijc}], \text{ siendo,}$$

p_{hijc} , la persona que cumple con la condición “c” en la vivienda “j”.

Tasa o promedio por persona para el nivel de estimación requerido

La estimación de razones se obtiene a partir del cociente entre los totales estimados para cada variable.

$$\hat{T}_c = \frac{\hat{Y}_c}{\hat{P}_c} = \frac{\text{Total Estimado de la variable “y” con la condición “c”}}{\text{Total Estimado de personas de 15 años y más que cumplen con la condición “c”}}$$

Proporción de personas con el atributo para el nivel de estimación requerido

La estimación de proporciones se obtiene a partir del cociente entre el total estimado y el parámetro para cada nivel.

$$\hat{P}_{\text{prop}} = \frac{\hat{P}_c}{\hat{P}} = \frac{\text{Total Estimado de personas de 15 años y más que cumplen con la condición “c”}}{\text{Total Estimado de personas de 15 años y más}}$$

VARIANZA DE LOS ESTIMADORES

Las fórmulas indicadas a continuación corresponden a la expresión matemática de cada varianza según sea la estimación propuesta: totales, promedios, tasas o proporciones.

Nomenclatura utilizada para la varianza de los estimadores:

FE_{hi} = Factor de expansión por sección “i”, del estrato “h”.

y_{hijc} = Valor de la característica medida en la persona de la vivienda “j”, sección “i”, estrato “h”.

p_{hijk} = Persona de la vivienda “j”, sección “i”, estrato “h” que tiene de 15 años y más que cumple con la característica “c”.

p_{hij} = Personas de la vivienda “j”, sección “i”, estrato “h” que tiene de 15 años y más.

n_h = Número de secciones en la muestra del estrato “h”.

R_h = Corresponde a la razón entre el valor estimado de la variable (o el número estimado de personas que tienen el atributo) y el número de estimado de personas del estrato “h”.

$$R_h = \frac{\hat{P}_{hc}}{\hat{P}_h} =$$

El factor de expansión FE_{hi} es un algoritmo aplicado como ponderador para determinar varianzas

$${}^{(1)}FE_{hi} = \frac{M_h \cdot M'_{hi}}{M_{hi} \cdot m_{hi}}$$

y se aplica sobre personas y hogares personas de 15 años y más en la sección “i” del estrato “h”.

$${}^{(2)}FE_{hi} = \frac{M_h \cdot M'_{hi}}{M_{hi} \cdot m_{hi}} \cdot p_{hij}, \text{ siendo,}$$

M_h = Número de viviendas en el estrato “h” según el censo de población y viviendas de 1992.

M_{hi} = Número de viviendas en la sección “i” según el censo 1992.

M'_{hi} = Número de viviendas actualizadas en la sección “i”, estrato “h”.

m_{hi} = Número de viviendas en la muestra de la sección “i”, estrato “h”

p_{hij} = Personas de la vivienda en el estrato “h”, sección “i” en la vivienda “j”.

Varianza del total estimado de la variable “y” para el nivel requerido

La varianza de una estimación de totales se obtiene sumando el algoritmo de las personas que tienen el atributo.

La expresión matemática es:

$$\hat{V}(\hat{Y}) = \sum_h \sum_i \sum_j \frac{1}{n_h \cdot (n_h - 1)} \cdot [F_{hi} \cdot (y_{hijc} - \hat{R}_h \cdot P_{hij})]^2$$

Varianza del total estimado de las personas que tienen el atributo en estudio para el nivel requerido

$$\hat{V}(\hat{P}) = \sum_h \sum_i \sum_j \frac{1}{n_h \cdot (n_h - 1)} \cdot [F_{hi} \cdot (y_{hijc} - \hat{R}_h \cdot P_{hij})]^2$$

Varianza de la tasa o promedio por persona para el nivel requerido

$$\hat{V}(\hat{T}) = \frac{1}{p^2} [\hat{V}(Y) + \hat{T}^2 \cdot \hat{V}(P) - 2 \cdot \hat{T} \cdot \text{COV}(Y, P)]$$

Varianza de proporciones para el nivel requerido

La varianza de proporciones es un caso particular de la varianza de tasas, donde:

$$\hat{V}(P) = 0 \text{ y } \text{COV}(\hat{Y}, \hat{P}) = 0$$

De esta forma la expresión matemática es:

$$V(\text{PROP}) = \frac{V(\hat{Y})}{p^2}$$

Siendo en este caso, $Y = P_c$, las personas que tienen el atributo y $P =$ El total de personas que tienen de 15 años y más en el nivel estimado.

Coefficiente de variación

El coeficiente de variación de una estimación se obtiene a del cociente entre la desviación estándar de la estimación (raíz cuadrada de la varianza) y el valor de la estimación.

La expresión matemática es:

$$C\hat{V}(Y) = \frac{\sqrt{V(\hat{Y})}}{\hat{Y}}$$

Conocido este valor, es posible determinar con la estimación y el coeficiente de variación de ésta, el intervalo de confianza en que se mueve el parámetro. Para ello, se sugiere considerar un intervalo con 95% de confianza ($k=1.96$).

La expresión matemática es:

$$\hat{Y} - 1.96 \cdot \sqrt{\hat{V}(Y)} \leq Y \leq \hat{Y} + 1.96 \cdot \sqrt{\hat{V}(Y)}$$

Construcción del índice de prevalencia de discapacidad

Hasta antes de este estudio, la prevalencia de la discapacidad en Chile se calculaba de manera directa en las encuestas con una sólo pregunta (CENSO y CASEN), consignando como porcentaje nacional la frecuencia simple resultante de las personas que respondían positivamente a la presencia de alguno de los déficit consultados. En la Encuesta de Calidad de Vida del año 2000 la prevalencia de la discapacidad correspondió al total de personas que presentaba a lo menos una dificultad en cualquiera de las preguntas que contenía la sección de discapacidad.

Sin embargo, dado lo dinámico del concepto de discapacidad presente en este estudio, fue necesario establecer un Índice de Discapacidad, cuyos pasos se explican a continuación.

Para determinar la “**Prevalencia de discapacidad en Chile**”, es decir poder establecer el número de personas

con discapacidad que hay en Chile y el porcentaje de la población total del país que representa, se realizaron los siguientes pasos:

Paso 1

Se estableció qué es discapacidad:

“**Discapacidad** es un término genérico, que incluye deficiencias de las funciones y/o estructuras corporales, limitaciones en la actividad y restricciones en la participación, indicando los aspectos negativos de la interacción entre un individuo (con una “condición de salud”) y sus factores contextuales (factores ambientales y personales)”⁵”.

En este sentido **persona con discapacidad** sería aquella persona que presenta deficiencias de sus funciones y/o estructuras corporales, limitaciones en sus actividades

y restricciones en su participación, como resultado de la interacción negativa de su condición de salud y los factores contextuales (ambientales y personales) en los que se desarrolla”.

Paso 2

Se estableció cuáles son los conceptos centrales sobre los que la CIF define la discapacidad, concluyéndose:

- 1 Estado negativo de salud
- 2 Limitaciones o dificultades en la realización de actividades
- 3 Restricciones en la participación
- 4 Barreras ambientales (físicas y actitudes)

Paso 3

Se escogió de entre las preguntas del Módulo 1, aplicadas a todos los miembros del hogar sin distinción de edad, todas aquellas preguntas que se referían a estos conceptos centrales.

Se seleccionaron las siguientes:

1 Estado negativo de salud:

Preguntas N° 1, 2, 3, 5, 26.

Preguntas	Contenido
1	Estado de salud general físico y mental
2	Presencia de dolor
3	Problema de salud que ha causado dificultad para realizar actividades diarias incluyendo laborales
5	Presencia de diagnóstico de un especialista sobre alguna discapacidad
26	Dificultades graves en aparatos y sistemas orgánicos vitales (respiratorios, cardiovasculares, endocrinos, genitourinarios)

2. Limitaciones o dificultades en la realización de actividades:

- Dificultades para ver: 6 y 8,
- Dificultades para oír 9, 11,
- Dificultades para el cuidado personal 12, 13, 14,
- Dificultades para desplazarse: 15, 16, 17 y 18,
- Dificultades para la comprensión y la comunicación: 19, 20, 21, 22, 23.

Preguntas	Contenido
6	Dificultad para ver
8	Impacto de este problema
9	Dificultad para oír
11	Impacto de este problema
12	Dificultad para bañarse sólo
13	Dificultad para vestirse o ponerse los zapatos sólo
14	Dificultad para comer por sí mismo
15	Dificultad para estar de pie 30 minutos o más
16	Dificultad para caminar distancias largas
17	Dificultad para ponerse de pie desde posición sentado
18	Dificultad para moverse dentro de la casa
19	Dificultad para aprender cosas nuevas
20	Dificultad para concentrarse
21	Dificultad para recordar cosas importantes
22	Dificultad para resolver imprevistos
23	Dificultad para comprender en general lo que dice la gente

3. Restricciones en la participación:

Pregunta N°25.

Esta pregunta contiene 54 posibles respuestas, pues apunta a establecer el como diferentes dificultades o problemas han afectado para recrearse, participar socialmente, lograr el cuidado personal, realizar el trabajo doméstico, estudiar y trabajar remuneradamente (ver cuadro de página siguiente).

Paso 4

Se aplicó un puntaje a cada una de las respuestas, dado que en algunos casos las respuestas iban de 0 a 3:

- 0 Ninguna dificultad
- 1 Poca o leve dificultad
- 2 Mucha dificultad
- 3 Extrema dificultad

Y en otros casos fue 0 ó 2, dependiendo si la respuesta fue dicotómica.

Paso 5

Suma de puntajes y división en 4 tramos: Este fue sumado resultando una tabla con recorrido de puntajes el que se dividió en 4 partes similares (intervalos homo-

généos considerando el valor promedio de la variable y su desviación estándar aplicada a partir de la media para la determinación de los intervalos).

Paso 6

De este modo surgen 4 tramos o intervalos que representan la presencia o no de discapacidad y en grados leve, moderada y severa.

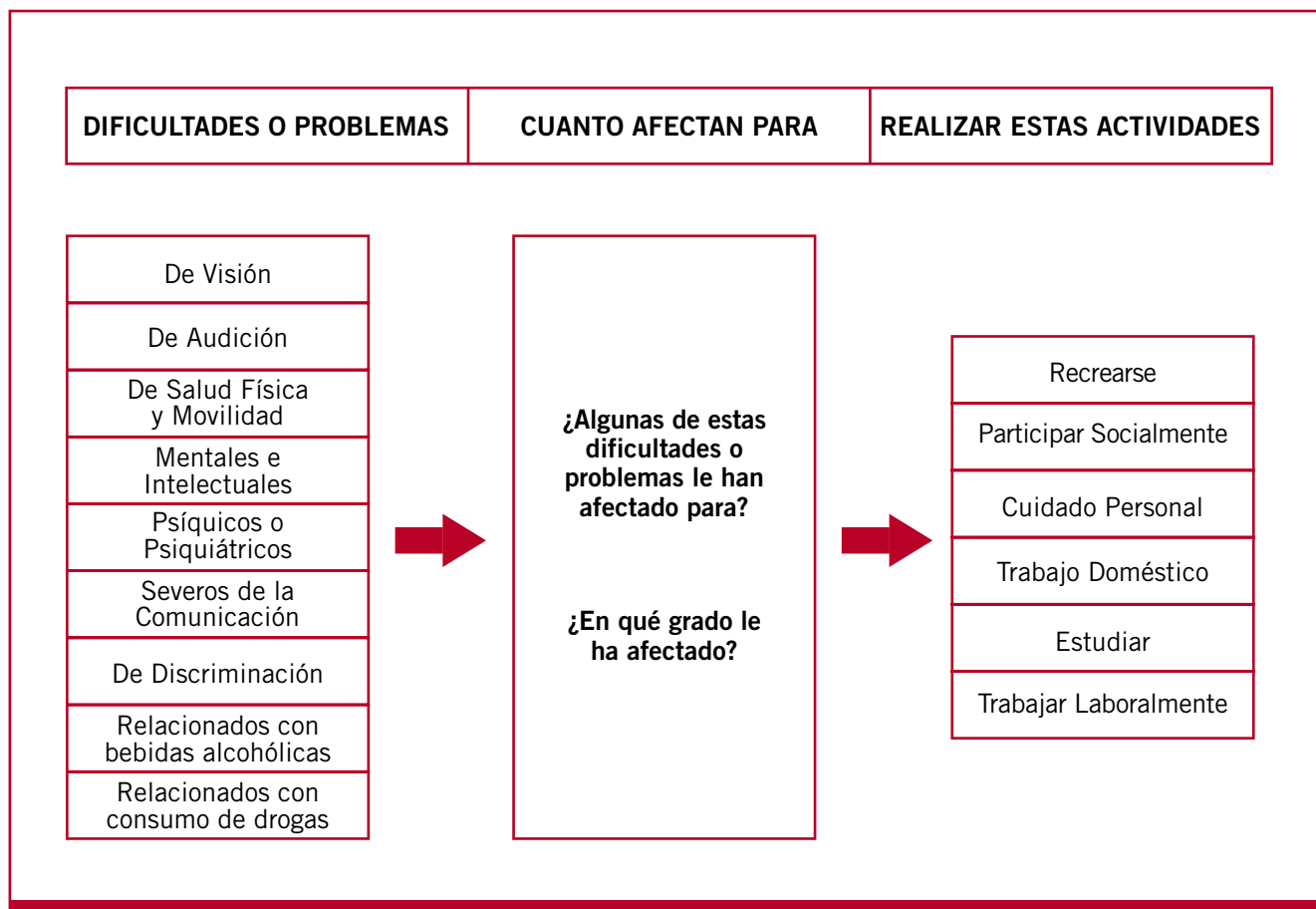
- 0 Ninguna discapacidad
- 1 Discapacidad leve
- 2 Discapacidad moderada
- 3 Discapacidad severa

Paso 7

Finalmente la suma de los resultados de los tramos 1,2 y 3 es equivalente al total personas con discapacidad en sus grados leve, moderado y severo

Esto permitió establecer el número y porcentaje de personas con discapacidad en Chile. En este caso resultó en un 12,93%, representando 2.068.072 personas. El tramo "0" (cero) corresponde a personas sin discapacidad. En este tramo existen muchas personas con algunas limitaciones en sus actividades o ciertas restricciones en su participación, pero que no alcanzan la combinación y exigencia establecida por el Índice de Discapacidad para ubicarlas de manera cierta como personas con discapacidad. Es decir, existe un nivel de exigencia para definir persona con discapacidad que es mayor que la sola presencia de una dificultad o de un estado negativo de salud.

Variables contenidas en pregunta N°25 del cuestionario



Definición de los grados de discapacidad

Discapacidad Leve: personas que presentan alguna dificultad para llevar a cabo actividades de la vida diaria sin embargo la persona es independiente y no requiere apoyo de terceros y puede superar barreras del entorno.

Discapacidad Moderada: personas que presentan una disminución o imposibilidad importante de su capacidad para realizar la mayoría de las actividades de la vida diaria, llegando incluso a requerir apoyo en labores básicas de auto cuidado y supera con dificultades sólo algunas barreras del entorno.

Discapacidad Severa: personas que ven gravemente dificultada o imposibilitada la realización de sus actividades cotidianas, requiriendo del apoyo o cuidados de una tercera persona y no logra superar las barreras del entorno.

⁵ Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud. OMS/OPS,2001.